



# Modulation pharmacologique du raisonnement et de la prise de décision : apports pour la psychiatrie

Alexandre Salvador

## ► To cite this version:

Alexandre Salvador. Modulation pharmacologique du raisonnement et de la prise de décision : apports pour la psychiatrie. Psychiatrie et santé mentale. Université Sorbonne Paris Cité, 2017. Français. NNT : 2017USPCB007 . tel-02109169

**HAL Id: tel-02109169**

**<https://theses.hal.science/tel-02109169>**

Submitted on 24 Apr 2019

**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.

UNIVERSITÉ PARIS DESCARTES

École doctorale n°158 Cerveau Cognition Comportement

*Inserm U894 Centre de Psychiatrie et Neurosciences*

# Modulation pharmacologique du raisonnement et de la prise de décision : apports pour la psychiatrie

Par Alexandre Salvador

Thèse de doctorat de Biologie

Dirigée par Raphaël Gaillard

Présentée et soutenue publiquement le 25 avril 2017

Devant un jury composé de :

|                        |                                                        |
|------------------------|--------------------------------------------------------|
| Raphaël GAILLARD       | Professeur - Université Paris Descartes                |
| Stefano PALMINTERI     | PhD - Ecole Normale Supérieure                         |
| Léon TREMBLAY          | Directeur de Recherche - Université Lyon 1             |
| Denis DAVID            | Professeur - Université Paris Sud                      |
| Jean-Christophe CORVOL | Professeur - Université Paris 6, Pierre et Marie Curie |
| Régis BORDET           | Professeur - Université Lille 2                        |



Except where otherwise noted, this work is licensed under  
<http://creativecommons.org/licenses/by-nc-nd/3.0/>



## **Résumé :**

L'innovation thérapeutique est limitée en psychiatrie. De nombreux médicaments sélectionnés sur la base de résultats encourageants dans les essais chez l'animal se révèlent décevants lors des essais cliniques. La validité limitée des modèles animaux, et leur utilisation pour tenter de mimer des pathologies définies de façon catégorielle sur la base de regroupement de symptômes de surface sans lien clair avec les processus cérébraux, les mécanismes biologiques ou la génétique, participent à ces difficultés.

Une branche des neurosciences cognitives, l'étude de l'apprentissage par renforcement, associée à l'utilisation d'interventions pharmacologiques ciblées chez le sujet malade ou le sujet sain, représente une opportunité de mieux caractériser les processus cérébraux sous-tendant certaines dimensions cardinales des pathologies psychiatriques.

Nous illustrons l'utilisation de l'étude de l'apprentissage par renforcement avec intervention pharmacologique dans deux études expérimentales. La première cherche à caractériser l'effet de l'aripiprazole, un antipsychotique atypique, chez des patients atteints du syndrome Gilles de la Tourette, en utilisant une tâche d'apprentissage contrefactuel, évaluant la capacité à apprendre non seulement des conséquences de ses actions, mais également des conséquences hypothétiques d'actions alternatives possibles. La seconde étude, randomisée contrôlée et en double aveugle, étudie l'effet de deux classes différentes d'antidépresseurs, l'escitalopram et l'agomélatine, chez le sujet sain. L'effet de leur administration est évaluée à court terme (3 jours) et à long terme (8 semaines) dans deux tâches probabilistes de sélection de stimulus, l'une simple, l'autre avec renversements occasionnels.

L'utilisation de cette approche pourrait participer à la définition d'endophénotypes et, en collaboration avec la recherche préclinique, aider à la création de nouveaux modèles animaux pour en améliorer la valeur prédictive.

**Mots-clés : Psychiatrie, Psychopharmacologie, Neurosciences cognitives, Apprentissage par renforcement , Modèles animaux**



# **Title : Pharmacological challenge of cognition and decision-making : implications for psychiatry**

## **Abstract :**

Successful new drug development has declined in psychiatry in the last decades. This is in part the result of a high failure rate in translating positive preclinical efficacy results to positive clinical trials. Limitations in the validity of animal models and shortcomings in the usefulness of the current categorical diagnostic system.

Cognitive neurosciences and particularly reinforcement learning and its computational analysis might provide biomarkers required to develop new ways of classifying mental disorders on the basis of both observable behaviour and neurobiological measures. Used in conjunction with pharmacological challenges, it may bring new insights into the pathophysiology and brain mechanisms underlying psychiatric disorders. It may also help design new animal models with improved predictive validity for the development of medications relying on innovative mechanisms of action.

We illustrate the use of reinforcement learning and pharmacological challenge in two experimental studies.

In the first experiment, we administered a reinforcement learning task that involves both direct learning from obtained outcomes and indirect learning from forgone outcomes to two groups of Gilles de la Tourette patients, one receiving aripiprazole, one unmedicated and to a group of healthy subjects.

In the second experiment, we administered two probabilistic stimulus selection learning tasks (one simple, one with occasional reversals) to healthy subjects randomly and blindly allocated to either escitalopram, a typical serotonin reuptake inhibitor, agomelatine, an antidepressant with a different mechanism of action, or placebo. The experiment compared the effect of these two classes of antidepressants to placebo after both short term (3 days) and long term (8 weeks) treatment.

These experiments bring insights into the understanding of the clinical condition studied, and the effects of the drugs tested. Implications of this approach for the translational approach to drug development is discussed.

**Keywords : Psychiatry, Psychopharmacology, Cognitive neuroscience, Reinforcement learning, Animal models**



## Remerciements

Je remercie tout particulièrement

Raphaël Gaillard pour avoir accepté de diriger cette thèse malgré un emploi du temps surnaturel, pour ses nombreuses qualités, humaines autant que professionnelles, la finesse de son enseignement clinique et la constance de son soutien ;

Stefano Palminteri pour avoir accepté de co-diriger cette thèse, pour sa formation et ses conseils multiples sur les passations IRM, la modélisation ... et les doubles dissociations<sup>TM</sup> ;

Régis Bordet, Jean-Christophe Corvol, Denis David et Léon Tremblay qui me font l'honneur de participer au jury de cette thèse ;

L'équipe du Centre de Psychiatrie et Neurosciences, au sein duquel j'ai réalisé ce travail ;

Etienne Koechlin pour avoir permis ma présence au sein du Laboratoire de Neurosciences Cognitives pour le travail avec Stefano Palminteri ;

Valentin Wyart pour la collaboration autour de la kétamine et sa formation accélérée à l'EEG et son analyse. Également à Luc Arnal pour son aide dans ce même projet ;

Mathias Pessiglione et son équipe, Jean Daunizeau, Sébastien Bouret et toute l'équipe Motivation Brain Behaviour de l'ICM pour les échanges et feedback précieux lors de la mise en place de projets ou l'analyse des données ;

Yulia Worbe pour nous avoir permis la mise en place du protocole Contrefactuel chez les patients GTS, Cécile Delorme pour l'aide apportée sur ce projet ;

L'équipe du Centre de Recherche Clinique de l'hôpital Sainte Anne, en particulier Macarena Cuenca, Cecile Bergot, Emilie Grasset pour leur aide et pour les pauses bavardes ;

La direction de la recherche de l'hôpital Sainte Anne, en particulier Marie Godard pour son aide pour rédiger le CPP ;

Les professeurs Henri Lôo, Jean Pierre Olié et Marie Odile Krebs pour la formation au métier de psychiatre dispensée dans l'enceinte prestigieuse du SHU de Sainte Anne ;

Mes co-chefs et assistants Pierre Abdel Ahad, Matthieu Gasnier, Florian Lejuste, Fabrice Rivollier, Sarah Smadja, Fabien Vinckier ainsi que les nouvellement nommés PH Claire Gauthier et Gilles Martinez pour leur participation à l'ambiance de travail amicale et solidaire précieuse dans l'exercice de notre profession ;

Fabien Vinckier en particulier pour son amitié pour les échanges fréquents et les discussions animées, au travail ou devant une bière ;

Mes amis, Max, Jean-Charles, Luc, Benoit, Laurent, Etienne, Yves, Marine, Jean-No, Perrine, Stéphane et Stéphanie, Aurore, Benoit, Arthur pour les moments d'amitié partagés ;

Jolie maman pour les heures de relecture, et pour les accents circonflexes ;

Ma famille, pour son amour et tout ce qu'elle m'a transmis, mon admiration et ma gratitude sont croissants à mesure que je découvre le difficile rôle de parent ;

Ma femme, pour cette vie riche et rythmée, pour les voyages passés et à venir, pour l'équilibre que m'apporte notre famille ... et d'avoir tout géré à la maison les derniers temps de cette thèse...

# Table des matières

|                                                                                                                         |           |
|-------------------------------------------------------------------------------------------------------------------------|-----------|
| Table des matières                                                                                                      | ix        |
| Table des figures                                                                                                       | xiii      |
| Introduction                                                                                                            | 1         |
| <b>I Partie théorique</b>                                                                                               | <b>3</b>  |
| <b>1 Limites dans les modèles actuels de l'innovation en psychiatrie</b>                                                | <b>5</b>  |
| 1.1 L'innovation thérapeutique a ralenti en psychiatrie . . . . .                                                       | 5         |
| 1.1.1 Des découvertes thérapeutiques majeures il y a un demi-siècle                                                     | 5         |
| 1.1.2 Peu d'innovation thérapeutique majeure depuis . . . . .                                                           | 6         |
| 1.2 Des perspectives d'avenir sombres . . . . .                                                                         | 7         |
| 1.2.1 Des traitements disponibles loin d'être parfaits . . . . .                                                        | 7         |
| 1.2.2 Les groupes pharmaceutiques qui se désengagent de la psy-<br>chiatrie . . . . .                                   | 9         |
| 1.3 Une crise non limitée à la psychiatrie . . . . .                                                                    | 10        |
| 1.3.1 Une faible probabilité de traverser avec succès les tests cliniques                                               | 10        |
| 1.3.2 Des causes économiques et techniques . . . . .                                                                    | 12        |
| 1.3.3 Des causes scientifiques . . . . .                                                                                | 14        |
| 1.4 Une crise plus marquée en psychiatrie . . . . .                                                                     | 15        |
| 1.4.1 Une probabilité de succès en phase clinique encore plus faible                                                    | 15        |
| 1.4.2 Un problème translationnel plus important . . . . .                                                               | 16        |
| 1.5 Résumé et conclusion . . . . .                                                                                      | 17        |
| <b>2 Les limites des études animales en psychiatrie</b>                                                                 | <b>21</b> |
| 2.1 Critères d'évaluation des modèles animaux . . . . .                                                                 | 21        |
| 2.1.1 Validité prédictive . . . . .                                                                                     | 22        |
| 2.1.2 Validité apparente . . . . .                                                                                      | 23        |
| 2.1.3 Validité de construit . . . . .                                                                                   | 24        |
| 2.1.4 Dimensions supplémentaires prenant en compte certaines<br>particularités des pathologies psychiatriques . . . . . | 24        |
| 2.2 Versatilité des modèles animaux . . . . .                                                                           | 26        |
| 2.2.1 Modélisation de facteurs d'environnement . . . . .                                                                | 26        |

|          |                                                                                        |           |
|----------|----------------------------------------------------------------------------------------|-----------|
| 2.2.2    | Modélisation de facteurs développementaux . . . . .                                    | 27        |
| 2.2.3    | Modélisation de mécanismes cérébraux . . . . .                                         | 27        |
| 2.2.4    | Modélisation de facteurs génétiques . . . . .                                          | 28        |
| 2.2.5    | Des combinaisons possibles . . . . .                                                   | 28        |
| 2.3      | Des améliorations aux résultats insuffisants . . . . .                                 | 29        |
| 2.3.1    | Les différents types de validité de ces modèles sont connus .                          | 29        |
| 2.3.2    | Des améliorations sont régulièrement proposées pour en améliorer la validité . . . . . | 30        |
| 2.3.3    | Mais la validité globale des modèles animaux est critiquée .                           | 32        |
| 2.3.4    | Exemple d'échec de transposition des résultats précliniques à l'homme . . . . .        | 32        |
| 2.4      | Des limites dans l'amélioration possible . . . . .                                     | 33        |
| 2.4.1    | Limites dans la validité homologique . . . . .                                         | 33        |
| 2.4.2    | Limites dans la validité ontopathogénique et mécanistique .                            | 36        |
| 2.4.3    | Limite dans la validité éthologique . . . . .                                          | 38        |
| 2.4.4    | Limites dans la validité prédictive . . . . .                                          | 38        |
| 2.5      | Résumé et conclusion . . . . .                                                         | 40        |
| <b>3</b> | <b>Limites des études cliniques, l'apport des neurosciences cognitives</b>             | <b>43</b> |
| 3.1      | Limites liées à l'étude des patients . . . . .                                         | 43        |
| 3.1.1    | Hétérogénéité . . . . .                                                                | 43        |
| 3.1.2    | Caractère neurodéveloppemental de certaines pathologies psychiatriques . . . . .       | 45        |
| 3.2      | Limites liées à la classification des pathologies psychiatriques . . . .               | 45        |
| 3.2.1    | Nosographie et hétérogénéité . . . . .                                                 | 46        |
| 3.2.2    | Nosographie et neurodéveloppement . . . . .                                            | 48        |
| 3.2.3    | Nosographie et génétique . . . . .                                                     | 49        |
| 3.3      | Les endophénotypes et les biomarqueurs . . . . .                                       | 50        |
| 3.3.1    | Les endophénotypes . . . . .                                                           | 50        |
| 3.3.2    | Distinction entre endophénotypes et biomarqueurs . . . . .                             | 52        |
| 3.3.3    | Le bénéfice attendu des endophénotypes . . . . .                                       | 53        |
| 3.3.4    | Les Research Domain Criteria (RDoC) . . . . .                                          | 55        |
| 3.3.5    | Variété des endophénotypes et biomarqueurs . . . . .                                   | 56        |
| 3.4      | L'apport des neurosciences cognitives . . . . .                                        | 57        |
| 3.4.1    | Définition et application à la psychiatrie . . . . .                                   | 57        |
| 3.4.2    | Biomarqueurs neuro-cognitifs de pathologies psychiatriques .                           | 58        |
| 3.5      | Couplage avec la psychopharmacologie . . . . .                                         | 59        |
| 3.5.1    | Étude de l'effet de traitements . . . . .                                              | 59        |
| 3.5.2    | Modélisation de pathologies psychiatriques . . . . .                                   | 60        |
| 3.6      | Résumé et conclusion . . . . .                                                         | 63        |
| <b>4</b> | <b>Focus sur l'apprentissage par renforcement</b>                                      | <b>65</b> |
| 4.1      | Historique et définition . . . . .                                                     | 65        |
| 4.1.1    | Historique . . . . .                                                                   | 65        |
| 4.1.2    | Définition . . . . .                                                                   | 66        |

|                                            |                                                                                                                                                                                             |            |
|--------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 4.2                                        | Formalisation et modélisation computationnelle . . . . .                                                                                                                                    | 68         |
| 4.2.1                                      | Formalisation . . . . .                                                                                                                                                                     | 68         |
| 4.2.2                                      | Algorithme “Temporal difference” . . . . .                                                                                                                                                  | 69         |
| 4.2.3                                      | Algorithme “Q-learning” . . . . .                                                                                                                                                           | 71         |
| 4.3                                        | Bases cérébrales de l’apprentissage par renforcement . . . . .                                                                                                                              | 72         |
| 4.3.1                                      | Cerveau et récompense . . . . .                                                                                                                                                             | 72         |
| 4.3.2                                      | Cerveau et erreur de prédiction . . . . .                                                                                                                                                   | 73         |
| 4.3.3                                      | L’apprentissage par punition . . . . .                                                                                                                                                      | 76         |
| 4.3.4                                      | L’apprentissage contrefactuel . . . . .                                                                                                                                                     | 78         |
| 4.3.5                                      | L’apprentissage model-based et model-free . . . . .                                                                                                                                         | 82         |
| 4.4                                        | Résumé et conclusion . . . . .                                                                                                                                                              | 84         |
| <br><b>II Partie expérimentale</b>         |                                                                                                                                                                                             | <b>87</b>  |
| Liste de publications et articles en cours |                                                                                                                                                                                             | 89         |
| 5                                          | Article 1 : Specific effect of a dopamine partial agonist on counter-factual learning : evidence from Gilles de la Tourette syndrome                                                        | 91         |
| 6                                          | Article 2 : Effects of acute and chronic antidepressant treatment on reinforcement learning and cognition : a 8 week randomized double blind placebo controlled study in healthy volunteers | 117        |
| <br><b>III Discussion générale</b>         |                                                                                                                                                                                             | <b>143</b> |
| 7                                          | Discussion générale                                                                                                                                                                         | 145        |
| 7.1                                        | Critique du choix des tâches d’apprentissage . . . . .                                                                                                                                      | 145        |
| 7.1.1                                      | Le choix des tâches . . . . .                                                                                                                                                               | 145        |
| 7.1.2                                      | Cohérence de deux tâches proches dans une même population                                                                                                                                   | 147        |
| 7.2                                        | Nécessité de connaissances fondamentales supplémentaires . . . . .                                                                                                                          | 149        |
| 7.2.1                                      | Sur notre connaissance des tâche . . . . .                                                                                                                                                  | 150        |
| 7.2.2                                      | Sur l’interaction des tâches avec la pathologie et les challenges pharmacologique . . . . .                                                                                                 | 152        |
| 7.2.3                                      | Sur la validité des modèle pharmacologiques de pathologies psychiatriques . . . . .                                                                                                         | 156        |
| 7.3                                        | Apports pour la psychiatrie et limites de l’approche . . . . .                                                                                                                              | 158        |
| 7.3.1                                      | Apports théoriques pour la psychiatrie . . . . .                                                                                                                                            | 158        |
| 7.3.2                                      | Une caractérisation plus proche des mécanismes génétiques?                                                                                                                                  | 161        |
| 7.3.3                                      | Une caractérisation spécifique d’une population clinique? . .                                                                                                                               | 164        |
| 7.3.4                                      | Une évaluation complémentaire en population non clinique?                                                                                                                                   | 166        |
| 7.4                                        | Approche translationnelle . . . . .                                                                                                                                                         | 167        |
| 7.4.1                                      | Translation et translation inverse . . . . .                                                                                                                                                | 168        |
| 7.4.2                                      | Translation et mesures neurophysiologiques . . . . .                                                                                                                                        | 172        |

|                                                                                                                         |     |
|-------------------------------------------------------------------------------------------------------------------------|-----|
| Conclusion                                                                                                              | 175 |
| Bibliographie                                                                                                           | 179 |
| Annexe 1 - Enhanced habit formation in Gilles de la Tourette syndrome                                                   | 217 |
| Annexe 2 - Illusion of agency in patients with Gilles de la Tourette Syndrome                                           | 229 |
| Annexe 3 - Confidence and psychosis : a neuro-computational account of contingency learning disruption by NMDA blockade | 239 |

# Table des figures

|     |                                                                                                                                                               |    |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Évolution du nombre de médicaments de mécanisme d'action différent entre 1950 et 2010 . . . . .                                                               | 7  |
| 1.2 | Innovation thérapeutique dans les différents champs médicaux, évaluée par le nombre de nouveaux médicaments commercialisés . . . .                            | 8  |
| 1.3 | Probabilité de succès aux différentes étapes des essais cliniques d'un candidat médicament entre la phase I et l'autorisation de mise sur le marché . . . . . | 11 |
| 1.4 | Augmentation de la complexité d'un essai clinique de phase III typique entre 2002 et 2012 . . . . .                                                           | 13 |
| 1.5 | Probabilité de succès en phase II des essais cliniques, par spécialité                                                                                        | 16 |
| 1.6 | Voies de développement de nouveaux médicaments, comparaison de la psychiatrie et des autres spécialités médicales . . . . .                                   | 18 |
| 2.1 | Modèle animal classique pouvant correspondre aux trois critères de validité principaux énoncés par Willner . . . . .                                          | 22 |
| 2.2 | Modèle animal prenant en compte l'aspect neurodéveloppemental des pathologies psychiatriques . . . . .                                                        | 25 |
| 2.3 | Version modifiée du test de nage forcée améliorant sa valeur prédictive                                                                                       | 31 |
| 2.4 | Anatomie comparée du cortex frontal chez l'homme, le singe et le rat                                                                                          | 35 |
| 2.5 | Délétions génétiques affectant des phénotypes agressifs chez la souris                                                                                        | 37 |
| 2.6 | Durée des traitements administrés dans les études précliniques publiées entre 2010 et 2015 . . . . .                                                          | 39 |
| 3.1 | Recouvrement de symptômes entre différentes catégories de diagnostic en psychiatrie . . . . .                                                                 | 47 |
| 3.2 | Concordance inter-juges des principaux diagnostics psychiatriques .                                                                                           | 48 |
| 3.3 | Analyse en réseau des gènes associés à la schizophrénie . . . . .                                                                                             | 51 |
| 3.4 | Utilisations potentielles des biomarqueurs en psychiatrie . . . . .                                                                                           | 53 |
| 3.5 | Justification de l'intérêt d'une approche par endophénotype pour l'analyse génétique de pathologies ayant un substrat génétique complexe                      | 54 |
| 3.6 | Probabilité de succès aux différentes étapes des essais cliniques selon qu'un biomarqueur est utilisé ou non . . . . .                                        | 55 |
| 3.7 | Comparaison des doses de kétamine utilisées dans les études cognitives versus en anesthésiologie . . . . .                                                    | 62 |

|     |                                                                                                                                                                                                   |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.1 | Différents mécanismes d'apprentissage par renforcement et de contrôle de l'action . . . . .                                                                                                       | 67  |
| 4.2 | La réponse dopaminergique encode l'erreur de prédiction de la récompense . . . . .                                                                                                                | 75  |
| 4.3 | Deux axes du contrôle du comportement, valence (récompense versus punition) et action (activation motrice ou approche versus inhibition motrice) . . . . .                                        | 79  |
| 4.4 | Proposition de vision globale des différentes aires cérébrales participant au raisonnement contrefactuel et des différentes fonctions cognitives s'y rapportant . . . . .                         | 81  |
| 4.5 | Démonstration du phénomène d'apprentissage latent dans un labyrinthe complexe . . . . .                                                                                                           | 83  |
| 7.1 | Comparaison du comportement essai par essai de deux sujets dans une tâche probabilistique de sélection de stimulus . . . . .                                                                      | 148 |
| 7.2 | Exemple de description de l'interaction dynamique de plusieurs sous-systèmes cérébraux dans une tâche d'apprentissage par renforcement chez le rat . . . . .                                      | 153 |
| 7.3 | Multiplés niveaux dans le contrôle inhibiteur à l'intérieur des ganglions de la base et l'hypothèse physiopathologique récente du syndrome Gilles de la Tourette . . . . .                        | 155 |
| 7.4 | Comparaison de la capacité de différentes substances psychotropes à déclencher des phénomènes subjectifs en lien avec des symptômes positifs . . . . .                                            | 157 |
| 7.5 | Illustration du principe de l'IRM fonctionnelle model-based . . . . .                                                                                                                             | 160 |
| 7.6 | Comportement variable entre sujets et, pour un même sujet, au cours du temps dans une tâche de Markov évaluant le comportement model-based versus model-free . . . . .                            | 163 |
| 7.7 | Méthodologie permettant un apport de l'approche computationnelle au problème de classification en psychiatrique . . . . .                                                                         | 165 |
| 7.8 | Illustration du dispositif à écran tactile permettant l'étude de l'apprentissage par renforcement chez le rat et la souris dans des conditions proches de celles utilisées chez l'homme . . . . . | 172 |

# Introduction

La pratique de la psychiatrie est une activité passionnante, mais également frustrante. Malgré notre volonté de soigner les patients, les armes médicamenteuses dont nous disposons sont pleines d'insuffisances. Leur efficacité est limitée, leurs effets indésirables importants et la proportion de patients qui décident d'arrêter leur traitement malgré les conseils de leur médecin est plus importante que dans bien d'autres spécialités médicales.

Le psychiatre regarde parfois avec jalousie ces autres spécialités dans lesquelles des innovations thérapeutiques sont annoncées régulièrement comme en cardiologie ou en cancérologie.

Dans la partie théorique de cette thèse, nous examinons tous d'abord dans quelle mesure ce défaut d'innovation thérapeutique est réel, dans quelle mesure il concerne plus la psychiatrie que les autres spécialités médicales, et à quel étape du processus d'innovation se situe le problème (Chapitre 1). Ceci nous amène à nous intéresser à notre capacité à inférer l'efficacité chez l'homme de molécules candidates à partir de l'expérimentation animale. Nous examinons donc spécifiquement les limites des modèles animaux, leur validité et leur capacité à modéliser des aspects variés et pertinents pour les pathologies psychiatriques (Chapitre 2). Puis nous nous intéressons aux limites de l'expérimentation chez l'homme et aux problèmes posés par les systèmes de classification nosographique catégorielle actuels. L'apport possible des endophénotypes est exposé et la contribution des neurosciences cognitives et de la psychopharmacologie est illustrée (Chapitre 3). L'étude de l'apprentissage par renforcement est une spécialité des neurosciences cognitives. Nous exposons son histoire, sa formalisation et certains de ses développements qui pourraient permettre de mieux caractériser les troubles psychiatriques et le mode d'action des traitements psychotropes (Chapitre 4).

Dans la partie expérimentale, nous illustrons par deux études l'utilisation d'une modulation pharmacologique de tâches d'apprentissage par renforcement.

La première est l'étude de l'effet de l'aripiprazole, un antipsychotique atypique, chez des sujets atteints du syndrome Gilles de la Tourette, syndrome caractérisé par des tics moteurs et verbaux et associé à une dysfonction dopaminergique à la fois sous-corticale et préfrontale. Nous caractérisons l'effet du traitement en termes d'apprentissage contrefactuel, qui est la capacité à apprendre non seulement des

conséquences d'actions effectuées, mais également de conséquences hypothétiques d'actions, prenant en compte le résultat que le sujet aurait obtenu s'il avait choisi une autre action (Chapitre 5 - Article 1).

La seconde est l'étude des effets de deux classes différentes d'antidépresseurs, l'escitalopram, un inhibiteur de la recapture de la sérotonine typique, et l'agomélatine, un antidépresseur au mécanisme d'action différent, contre placebo chez le sujet sain. Nous caractérisons ces effets, à la fois en administration subaigüe (à 3 jours) et en administration chronique (à 8 semaines), dans deux tâches d'apprentissage par renforcement de type tâche probabiliste de sélection de stimulus, l'une simple et l'autre avec renversement occasionnel (Chapitre 6 - Article 2).

Nous ouvrons ensuite une discussion sur les promesses d'une telle approche, sur les difficultés qu'elle rencontre, et les changements qu'elle réclame dans les liens entre recherche chez l'animal et recherche chez l'homme afin de favoriser à terme l'innovation thérapeutique (Discussion générale).

Première partie

Partie théorique



# Chapitre 1

## Limites dans les modèles actuels de l'innovation en psychiatrie

### 1.1 L'innovation thérapeutique a ralenti en psychiatrie

#### 1.1.1 Des découvertes thérapeutiques majeures il y a un demi-siècle

Les principales classes thérapeutiques que nous utilisons au quotidien en psychiatrie sont le fruit de découvertes fortuites effectuées il y a plusieurs décennies.

C'est le cas par exemple des antipsychotiques. La chlorpromazine était initialement un antihistaminique étudiée par un chirurgien, Henri Laborit, qui cherchait un stabilisateur neurovégétatif (Laborit et al., 1952). Son efficacité contre les productions hallucinatoires et délirantes, son action sur l'agitation maniaque, l'angoisse et l'agressivité sont des découvertes fortuites dues à la perspicacité et à la curiosité transdisciplinaire de deux psychiatres, Pierre Deniker et Jean Delay entre 1950 et 1955.

C'est le cas également des antidépresseurs, à la même période. Selikoff et Robitzek, deux médecins travaillant sur deux nouveaux antituberculeux, l'iproniazide et l'isoniazide, remarquent une amélioration clinique au plan infectieux, mais pas seulement. Ils notent dans la partie observation toxicologique de leur publication (Selikoff et Robitzek, 1952) :

“Alterations of Mood : With higher dosage ranges, especially of Marsilid, there is a subtle general stimulation to be noted.” (Selikoff et Robitzek, 1952)

et également que

“All patients are now ambulatory. Strength and energy rapidly returned ; lethargy disappeared. The patients exhibited renewed vigor and indeed this occasionally served to introduce disciplinary problems.” (Selikoff et Robitzek, 1952)

L'efficacité antituberculeuse des molécules amène une large diffusion de la publication de Selikoff et Robitzek et les effets “collatéraux” sur l'humeur sont notés avec intérêt par plusieurs psychiatres comme Max Lurie (Salzer et Lurie, 1953) ou Jean Delay (Delay et Buisson, 1958). L'amélioration de l'humeur induite par l'iproniazide est testée et confirmée chez des patients déprimés, ce qui amène la création du terme antidépresseur. Ce n'est que plus tard que le mécanisme d'action de l'iproniazide est découvert, il s'agit d'un inhibiteur de la monoamine oxydase (IMAO).

Loin d'être anecdotiques, ces modes de découverte ont été prépondérants dans les grandes innovations initiales en psychiatrie. Ils concernent également d'autres classes médicamenteuses, comme le lithium (Cade, 1949), ou les benzodiazépines (Preskorn, 2010). Les conséquences de ces innovations ont été majeures. Une intervention pharmacologique a enfin pu être efficace dans les psychoses par exemple. Leur histoire naturelle a pu être modifiée, des patients ont pu ressortir de l'asile psychiatrique et se réinsérer dans la société.

Ces découvertes sont une illustration du bénéfice de la sérendipité, qui est l'exercice d'une sagacité fortuite et qui fournit parfois des avancées majeures dans la découverte thérapeutique.

### 1.1.2 Peu d'innovation thérapeutique majeure depuis

Le nombre de classes médicamenteuses reposant sur des mécanismes d'action différents en psychiatrie aujourd'hui est comparable à ce qui existait à la fin des années 50. En comparaison, d'autres champs de la médecine, comme la cardiologie, ont vu dans le même temps le nombre de médicaments reposant sur des mécanismes d'action différents augmenter de façon spectaculaire (Figure 1.1) (NAMHC, 2010).

Contrairement à l'oncologie qui voit une augmentation depuis 1990 du nombre de nouvelles molécules arrivant sur le marché, la situation ne semble pas s'améliorer en psychiatrie avec par exemple un nombre de nouveaux antidépresseurs moins élevé durant la décade 2000 - 2010 que durant la décade 1990 - 2000 (Figure 1.2). Il semble donc y avoir un ralentissement dans le développement de nouvelles thérapeutiques pour le système nerveux central depuis les années 1990 (Kesselheim et al., 2015).

De plus, parmi les nouveaux médicaments commercialisés, nombreux sont ceux qui constituent des modifications mineures de molécules connues (NAMHC, 2010).

Les nouvelles molécules commercialisées pour le traitement de la schizophrénie par

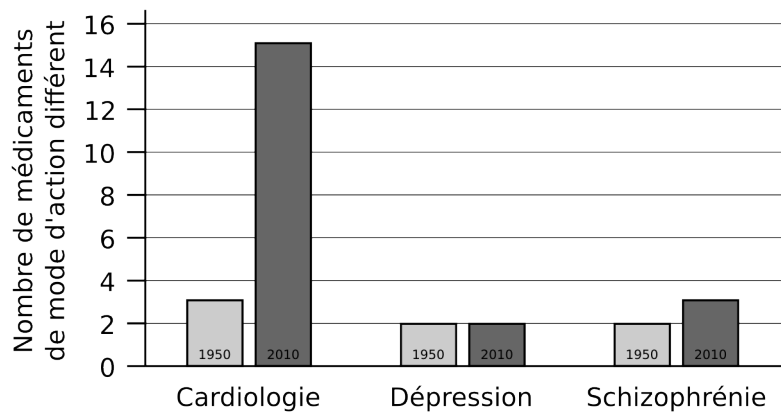


FIGURE 1.1 – Évolution du nombre de médicaments de mécanisme d'action différent entre 1950 et 2010. Comparaison pour les pathologies cardiovasculaires et pour deux pathologies du champ de la psychiatrie, la dépression et la schizophrénie. D'après (NAMHC, 2010)

exemple ont gagné en affinité pour les récepteurs D2. Elles étaient plus “puissantes” en ce sens qu’une moindre dose de substance active donnait les mêmes effets (Creese et al., 1976). Cependant, l’efficacité du médicament sur la maladie restait le même car le mécanisme d’action restait inchangé (Freedman, 2005). La découverte des antipsychotiques “atypiques” a vu l’utilité d’un mécanisme d’action complémentaire, comme l’antagonisme 5HT2A, qui vient diminuer les effets indésirables moteurs du blocage dopaminergique D2 dans la voie motrice nigro-striée. Mais ce progrès est là encore une découverte fortuite liée à l’observation par les cliniciens de l’effet inhabituel (“atypique”) d’un traitement avec une moindre incidence des syndromes extra-pyramidaux. Il ne découle pas de la définition d’une cible et de l’élaboration d’un médicament psychotrope aux caractéristiques désirées en termes réceptoriel. L’affinité antagoniste 5HT2A de ces antipsychotiques “atypiques” n’a été découvert que bien postérieurement à leur commercialisation.

## 1.2 Des perspectives d'avenir sombres

### 1.2.1 Des traitements disponibles loin d’être parfaits

La découverte des médicaments psychotropes a révolutionné la prise en charge des maladies psychiatriques. Ceux dont nous disposons aujourd’hui restent cependant encore insuffisamment efficaces. C’est le constat fait par des études d’efficacité à grande échelle, en conditions proches de la pratique clinique et avec une méthodologie solide, comme l’étude CATIE pour les antipsychotiques et l’étude STAR\*D pour les antidépresseurs.

Pour les antipsychotiques, le marqueur d’efficacité retenu dans l’étude CATIE était

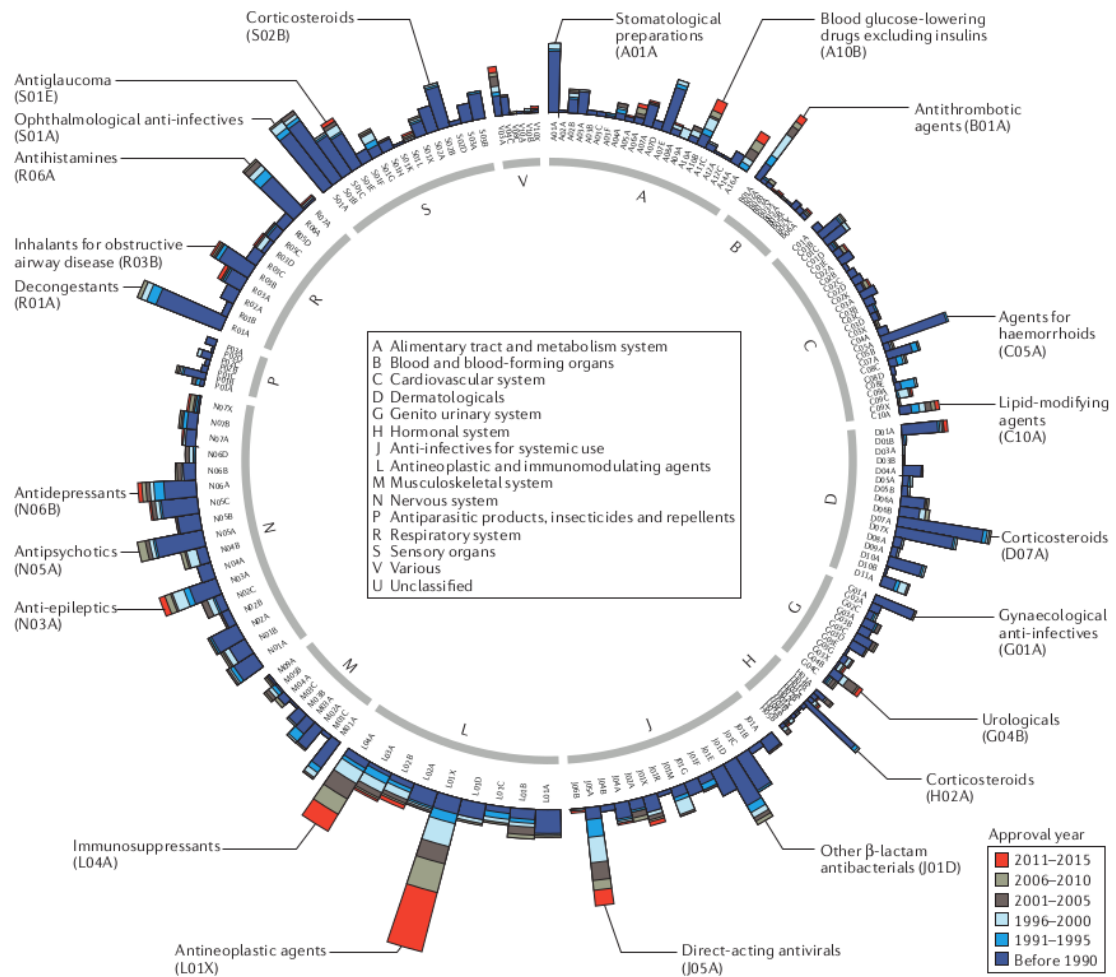


FIGURE 1.2 – Innovation thérapeutique dans les différents champs médicaux, évaluée par le nombre de nouveaux médicaments commercialisés. Chaque point dans le cercle central correspond à un type de médicament repéré par son code de classification (Organisation Mondiale de la Santé - code ATC). Chaque histogramme représente le nombre de médicaments commercialisés selon l'année de la première autorisation de mise sur le marché, regroupés tous les 5 ans avec la dernière période (2011-2015) en rouge (Santos et al., 2017).

la durée avant arrêt du traitement, une mesure écologique et qui reflète aussi bien des problèmes d'efficacité que d'observance. Les résultats ont pu être qualifiés de “décourageants” (Freedman, 2005) : avec tous les antipsychotiques étudiés, la majorité des patients ont arrêté leur traitement en moins de 18 mois, durée de l'étude. Les patients sous quetiapine ou risperidone ne montraient pas de différence de durée par rapport au traitement plus ancien servant de comparateur, la perphénazine. Seule l'olanzapine se démarquait avec 36% des patients toujours sous traitement à l'issue des 18 mois, mais elle le faisait au prix d'effets indésirables, notamment métaboliques, plus importants (Lieberman et al., 2005).

Pour les antidépresseurs, l'étude correspondante, STAR\*D a inclus 4041 patients et étudié la prescription séquentielle pour le traitement de la dépression. Le taux de rémission après la première ligne de traitement n'était que de 36.8%. Le taux de rémission cumulé après 4 lignes de traitement n'était que de 67% (Rush et al., 2006). Le taux de réponse (diminution de plus de 50% des scores de dépression) étaient de l'ordre de 47% pour la première ligne de traitement par citalopram (Rush et al., 2006). Les patients nécessitant plus de lignes de traitements étaient d'autre part plus à risque de rechute dans la période de suivi. Les méta analyses plus récentes confirment des taux de réponse du même ordre, bien que des différences puissent être mises en évidence entre les différentes molécules (Cipriani et al., 2009), avec pour la molécule considérée comme la plus efficace un taux de réponse d'environ 59%.

### 1.2.2 Les groupes pharmaceutiques qui se désengagent de la psychiatrie

Nous voyons donc qu'après une période d'innovation thérapeutique intense, fruit d'une sérendipité féconde mais non reproductible, nous sommes dans une situation où nous disposons d'armes thérapeutiques qui, si elles ont radicalement modifié la prise en charge des patients en psychiatrie, nous laissent avec une marge d'amélioration très importante.

Or les connaissances sur les bases neuroscientifiques des pathologies psychiatriques sont en plein essor. Les résultats expérimentaux sont nombreux chez l'animal comme chez l'homme. Des circuits cérébraux commencent à être identifiés grâce à l'avancée des techniques d'imagerie fonctionnelle. La connaissance des mécanismes neurochimiques en jeu dans le cerveau progresse. De nombreux loci génétiques ont été identifiés en lien avec certaines pathologies psychiatriques (Insel et al., 2013).

On peut d'ailleurs noter que la psychiatrie se situe extrêmement bien en termes de découvertes scientifiques par rapport aux autres disciplines médicales. En effet, plusieurs revues en psychiatrie et en neurosciences possèdent des facteurs d'impact parmi les plus élevés des revues médicales (Geoffroy et al., 2014).

Cependant dans le même temps, plusieurs grandes entreprises pharmaceutiques

ont annoncé qu'elles réduisaient leurs opérations de recherche et développement dans le secteur psychiatrique. C'est le cas pour GlaxoSmithKlein qui en 2010 a fermé deux centres de recherche en neurosciences et pour Astra Zeneca qui a également fermé deux centres de recherche. Ce dernier a de plus annoncé l'arrêt de son activité de recherche et développement sur la schizophrénie, les troubles bipolaires, la dépression et l'anxiété (Miller, 2010).

Plus récemment, le groupe pharmaceutique Servier, pour lequel l'activité "neuro-psychiatrie" était l'un des cinq axes de développement principaux, semble avoir pris le même virage. Lors de sa conférence annuelle de décembre 2016, un de ses directeurs indiquait que le groupe possédait 23 candidats médicaments en développement clinique, dont 17 nouvelles entités moléculaires. Sur ces 17 nouvelles entités moléculaires, huit concernaient l'oncologie, trois les maladies cardio-vasculaires, trois les maladies neurodégénératives, deux les maladies immuno-inflammatoires et une le diabète (Bour, 2016). Aucune ne concernait la psychiatrie. Le groupe Servier a par ailleurs annoncé mettre fin à son activité de recherche en psychiatrie pour se concentrer, pour les pathologies concernant le cerveau, sur les maladies neurodégénératives comme la maladie d'Alzheimer, la maladie de Parkinson ou la sclérose en plaques (Bour, 2016).

Un des dirigeants de GlaxoSmithKlein indiquait que les champs de recherche comme la dépression et l'anxiété étaient des champs de recherche où "nous pensons que la probabilité de succès est relativement basse, [et] nous pensons que le coût pour atteindre le succès est exagérément haut" (Miller, 2010).

Pourquoi cette crise dans le processus d'innovation et de découverte de nouvelles armes thérapeutique en psychiatrie ?

## **1.3 Une crise non limitée à la psychiatrie**

### **1.3.1 Une faible probabilité de traverser avec succès les tests cliniques**

La difficulté pour une entreprise pharmaceutique de développer une nouvelle molécule qui deviendra effectivement un traitement disponible sur le marché ne se limite pas au seul champ de la psychiatrie. Ainsi, le "likelihood of approval", c'est-à-dire la probabilité qu'une molécule ayant franchi les étapes précliniques et entrant en phase I des essais cliniques deviennent un jour un médicament possédant une autorisation de mise sur le marché n'est que de 9.6% toutes disciplines confondues (Figure 1.3).

La phase I est essentiellement conduite pour vérifier la sécurité d'utilisation et déterminer les doses maximales qui peuvent être utilisées chez l'homme avant que des effets indésirables intolérables ou dangereux n'apparaissent. Le taux de succès

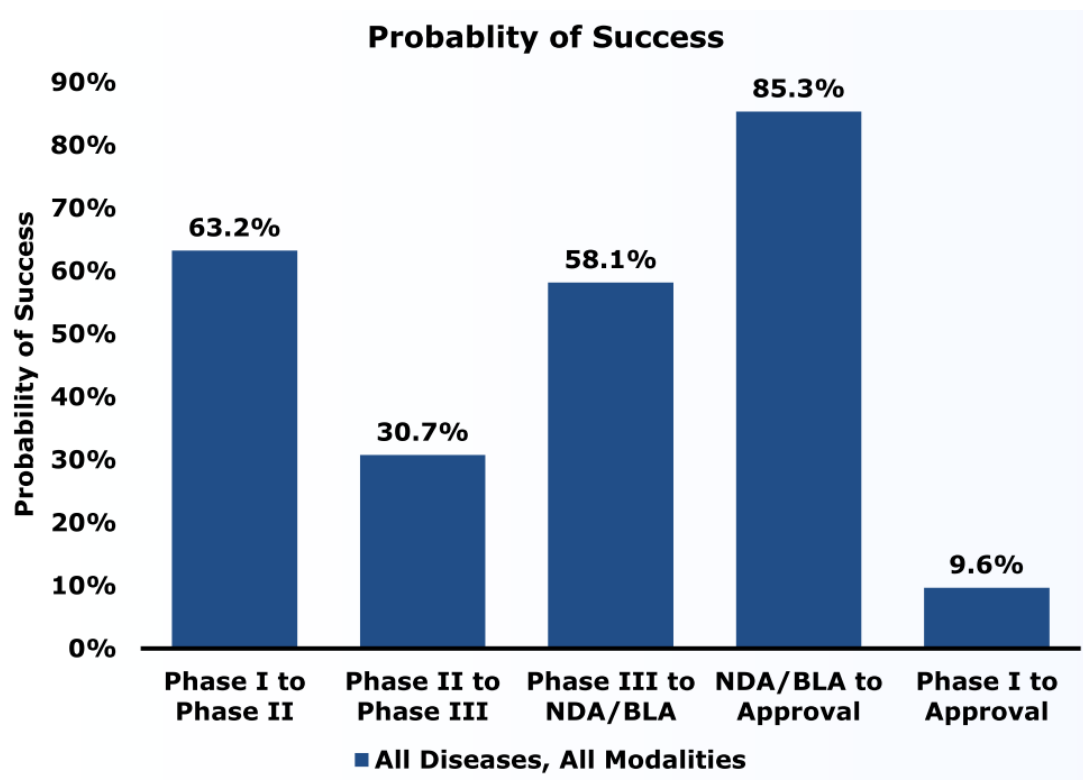


FIGURE 1.3 – Probabilité de succès aux différentes étapes des essais cliniques d'un candidat médicament entre la phase I et l'autorisation de mise sur le marché. NDA/BLA = New Drug Application / Biologic Licence Application. Approval = équivalent de l'Autorisation de Mise sur le Marché (Hay et al., 2014; Thomas et al., 2016)

à cette étape est relativement élevé, à 63.2%. C'est en phase II que le taux de succès est le plus faible avec 30.7% (Hay et al., 2014 ; Thomas et al., 2016). Or c'est la première phase où la démonstration de faisabilité est réellement testée chez l'homme. Elle détermine si la molécule testée a un effet chez l'homme comme attendu d'après les expérimentations précliniques.

Un non succès à passer de la phase II à la phase III peut signifier deux choses. Soit la molécule testée avec succès sur l'animal en préclinique ne montre pas d'effet apparent sur l'homme, c'est un échec d'ordre translationnel. Soit la molécule montre un effet sur l'homme, mais la firme pharmaceutique décide de ne pas poursuivre les essais.

### 1.3.2 Des causes économiques et techniques

Pourquoi une firme pharmaceutique déciderait-elle de ne pas poursuivre en phase III le développement d'une molécule présentant un effet sur l'homme en phase II ? La complexification des études de phase III et l'augmentation importante du coût qui lui est attaché peut-être un facteur explicatif.

Dans les quinze dernières années, la complexité des essais cliniques de phase III s'est en effet beaucoup accrue. Comme illustré dans la figure 1.4, le nombre de procédures auquel doit se soumettre un participant est passé de 106 en 2002 à 167 en 2012, le nombre de données enregistrées pour un essai de phase III atteint 929 000 en 2012, soit près d'un million d'enregistrements (Rosenblatt, 2017).

De plus, un nombre croissant de pays réclame, comme condition à l'obtention d'une autorisation de mise sur le marché, qu'une cohorte émanant de leur pays soit incluse dans les essais ; les candidats médicaments doivent se comparer le plus souvent à un traitement déjà actif et non un placebo, ce qui nécessite des cohortes plus grandes et un temps de suivi plus long pour montrer un effet ; les contrôles et organes de régulation sont plus nombreux (Rosenblatt, 2017).

Tout ceci concourt à rendre les essais de phase III plus complexes, plus hasardeux et plus coûteux. Environ 35% de tous les coûts de recherche et développement sont maintenant dépensés dans les essais de phase III. Les essais de phase III représentent 60% des coûts de tous les essais cliniques (Hay et al., 2014).

Cette augmentation importante du coût des essais cliniques de phase III fait qu'un groupe pharmaceutique n'aura pas les moyens de lancer simultanément un grand nombre d'études et devra donc faire des arbitrages. Un non succès de transition entre la phase II et la phase III ne signifie donc pas nécessairement que la molécule ne montre pas d'effet chez l'homme. Elle peut résulter d'un choix stratégique fait par la firme pharmaceutique de favoriser la poursuite du développement d'autres molécules, en fonction de considérations de marché par exemple.

| <b>Table 1. Increasing Complexity of a Typical Phase 3 Clinical Trial.*</b> |             |             |
|-----------------------------------------------------------------------------|-------------|-------------|
| <b>Trial Design Characteristic</b>                                          | <b>2002</b> | <b>2012</b> |
| Total no. of end points                                                     | 7           | 13          |
| Total no. of procedures                                                     | 106         | 167         |
| Total no. of eligibility criteria                                           | 31          | 50          |
| Total no. of countries                                                      | 11          | 34          |
| Total no. of investigative sites                                            | 124         | 196         |
| Total no. of patients undergoing randomization                              | 729         | 597         |
| Total no. of data points collected                                          | NA          | 929,203     |

FIGURE 1.4 – Augmentation de la complexité d'un essai clinique de phase III typique entre 2002 et 2012. Comparaison du nombre de procédures, critères d'inclusion, pays participant, sites participant, patients et données collectées. NA = not available. (Rosenblatt, 2017)

### 1.3.3 Des causes scientifiques

Une partie du faible succès de transition entre phase II et phase III est cependant attribuable au manque d'effet retrouvé chez l'homme alors que les études précliniques étaient prometteuses.

En dehors du champ de la psychiatrie, de nombreuses explications à ce phénomène ont été avancées. Certaines concernent le processus de la recherche de façon générale. Des études avancent qu'en cumulant les différentes causes d'inefficience dans le processus de recherche, on arrive à 85% de gâchis (Chalmers et Glasziou, 2009). Le design et la méthodologie des études seraient faibles, avec en particulier une insuffisance des mesures mises en place pour contrôler les biais dans 50% des études et une puissance inadéquate, menant à un faible taux de répliation (Macleod et al., 2014). L'accès aux résultats de l'intégralité des études serait insuffisant, il existerait un biais de publication pour les résultats positifs. Ainsi, parmi les résultats présentés sous forme de résumé lors de communications ou de congrès, seuls 53% étaient effectivement publiés sous forme d'article après 9 ans (Scherer et al., 2007). Même en se restreignant aux essais randomisés contrôlés, ce taux reste de seulement 63%. Les publications ne décriraient pas suffisamment l'intervention, ce qui gênerait la reproductibilité. Des biais dans la façon dont les résultats sont rapportés existeraient. 50% des variables d'intérêt décrites dans le projet initial ne seraient pas rapportées dans les articles finaux (Chalmers et Glasziou, 2009).

Ces critiques quoique assez générales n'en ont pas moins une portée pratique extrêmement importante. Prenons quelques exemples.

Dans la recherche sur le cancer, des auteurs ont étudié les 1915 publications de 2005 dans deux bases de données, portant sur les marqueurs pronostiques de cancer. Ils ont noté qu'entre 90% et 96% de ces études rapportaient des résultats significatifs (Kyzas et al., 2007), soit littéralement plus d'un millier en un an. Cependant, dans la pratique la proportion de ces cibles jugées significatives qui deviennent les cibles de médicaments effectivement développés puis finalement approuvés par les autorités de santé est extrêmement bas (Anderson, 2010), avec pour l'ensemble des tests basés sur des protéines au sens large approuvé par la FDA jusqu'en 2008 une estimation à 205.

Autre exemple, des scientifiques de Amgen ont tenté de répliquer des études précliniques très novatrices dans le champ de l'hématologie et de la cancérologie. Sur 53 papiers présentant des résultats positifs novateurs publiés, seulement 6 (11%) ont pu être répliqués (Begley et Ellis, 2012). Certaines de ces recherches non confirmées avaient pourtant déclenché un grand nombre de recherches secondaires, plusieurs centaines parfois, qui développaient des aspects de la publication princeps, mais sans chercher à confirmer ou infirmer les résultats initiaux. Au total, le nombre de publications secondaires n'était pas moins important pour les recherches non répliquées que pour les recherches répliquées (248 versus 231 pour les 21 études initialement publiées dans des journaux avec facteur d'impact supérieur à 20)

(Begley et Ellis, 2012).

La faible probabilité qu'un candidat médicament devienne un jour un médicament mis sur le marché est donc en partie due à des éléments économiques et méthodologiques qui affectent toutes les branches de la médecine et non seulement la psychiatrie. Voyons maintenant si des particularités affectent le développement de nouveaux médicaments pour les pathologies psychiatriques en particulier.

## 1.4 Une crise plus marquée en psychiatrie

### 1.4.1 Une probabilité de succès en phase clinique encore plus faible

Il semble effectivement exister des particularités dans le champ de la psychiatrie qui augmente encore la difficulté à obtenir en bout de ligne de nouveaux médicaments efficaces mis sur le marché. Une molécule candidate entrant en phase I toute spécialité confondue, a 9.6% de chances de devenir un jour un médicament disposant d'une autorisation de mise sur le marché. Cette probabilité s'amenuise pour les molécules candidates du champ de la psychiatrie, à seulement 6.2% (Hay et al., 2014; Thomas et al., 2016). Des chiffres différents existent selon les études, mais le plus faible taux de succès en psychiatrie y reste présent (Kola et Landis, 2004).

L'analyse étape par étape montre que la moindre efficacité en psychiatrie à amener une molécule candidate sélectionnée par la recherche préclinique au stade de médicament commercialisé provient de faiblesses à deux étapes clef. La phase II où la psychiatrie se classe en dernière position avec une probabilité de succès de seulement 24% pour une moyenne toutes spécialités confondues de 31% (Figure 1.5). La phase III où la psychiatrie se classe dans les trois dernières places, avec une probabilité de succès de 56% pour une moyenne de 58% (Hay et al., 2014; Thomas et al., 2016).

Lorsque les causes de non poursuite du développement du médicament sont identifiées, la cause principale pour les candidats médicaments en psychiatrie est l'échec à démontrer son efficacité (Kesselheim et al., 2015; Kola et Landis, 2004), avant la découverte de problèmes de sécurité, de toxicité ou d'effets secondaires non anticipés (Kola et Landis, 2004).

On découvrirait donc plus fréquemment en psychiatrie que dans d'autres spécialités que le candidat médicament est insuffisamment efficace seulement au stade des essais cliniques. On ferait de plus cette désagréable découverte dans des phases tardives des essais cliniques, en phase II et en phase III, plus fréquemment en psychiatrie.

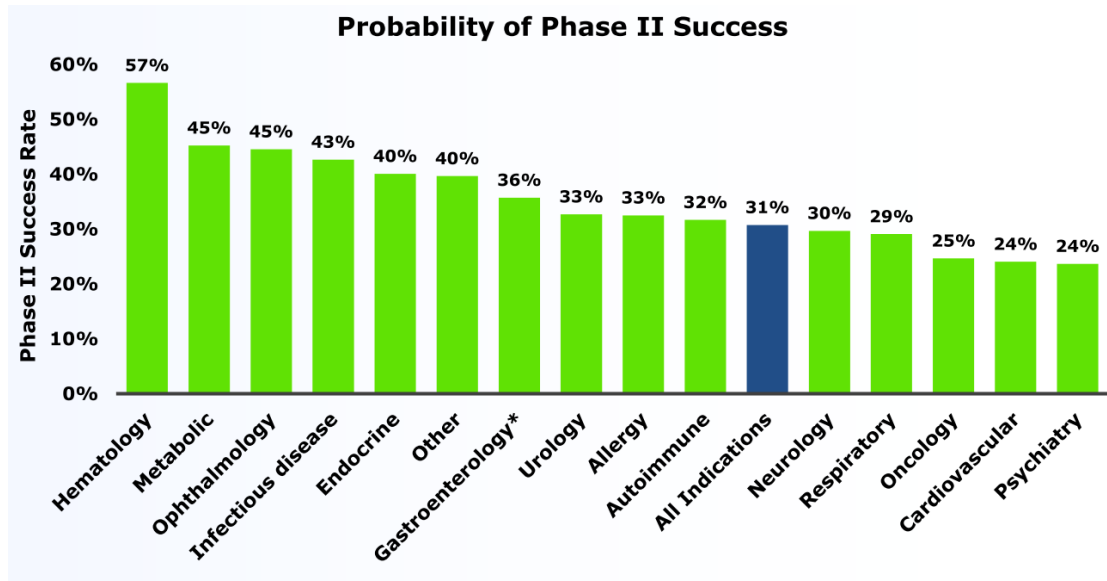


FIGURE 1.5 – Probabilité de succès en phase II des essais cliniques, par spécialité. Les spécialités sont classées par ordre décroissant de la plus haute à la plus basse probabilité de transition de la phase II à la phase III. (Hay et al., 2014 ; Thomas et al., 2016)

### 1.4.2 Un problème translationnel plus important

Dans la plupart des autres spécialités médicales, le développement d'un traitement commence par la définition de la physiopathologie et d'une cible.

Cette cible peut être définie par la biologie cellulaire. Un exemple historique et paradigmatique est la leucémie myéloïde chronique, pour laquelle l'observation initiale d'une anomalie chromosomique, le chromosome Philadelphie qui est un chromosome 22 raccourci sur son bras long du fait d'une translocation, a permis d'identifier que cette anomalie avait pour résultat la synthèse d'une protéine ayant une activité thyrosine kinase aberrante (Hehlmann et al., 2007). Ceci a conduit au développement d'une thérapie ciblée, l'imatinib, qui a révolutionné le pronostic de la maladie (Hehlmann et al., 2007).

Mais il n'est pas nécessaire d'avoir une compréhension parfaite de la physiopathologie. Une cible peut être définie sur la base d'associations observées entre un facteur et la maladie par le biais d'études qui peuvent être épidémiologiques. L'étude Framingham a ainsi permis de faire un lien entre taux de cholestérol et mortalité cardio-vasculaire (Kronmal et al., 1993). Ceci a conduit au développement de thérapeutiques spécifiques, agissant sur ce facteur de risque cardiovasculaire particulier.

En psychiatrie au contraire, les causes et mécanismes des pathologies psychiatriques sont mal compris et les cibles thérapeutiques sont souvent mal définies (Insel

et Landis, 2013). Bien pire, lorsqu'une cible hypothétique est définie, il existe parfois une incertitude sur le fait que la molécule candidate engage réellement cette cible. Les échecs de développement thérapeutique résultant d'un défaut d'efficacité constatée en phase II ont été analysés au sein du laboratoire pharmaceutique Astra Zeneca. Dans 21% des cas, il était impossible de déterminer si l'échec était du au fait que la molécule candidate n'avait pas atteint sa cible ou si elle avait atteint sa cible, mais n'avait pas eu d'action pharmacodynamique. Ce phénomène était plus marqué pour les pathologies du champ de la psychiatrie (Cook et al., 2014).

La mauvaise compréhension du mode d'action des molécules testées va encore plus loin. Le mécanisme même de transport de la molécule dans l'encéphale n'est pas toujours maîtrisé. Ainsi 98% des molécules découvertes ciblant le système nerveux central ne passeraient pas la barrière hémato-encéphalique (Pardridge, 2001). La mauvaise compréhension et la mauvaise modélisation des mécanismes permettant à une molécule de passer la barrière hémato-encéphalique aurait été la source de résultats négatifs dans les essais cliniques de certaines nouvelles molécules (Di et al., 2013).

En conséquence, le cheminement pour le développement d'un traitement en psychiatrie est différent de celui suivi pour le développement moderne des médicaments comme en cancérologie. Ceci est illustré Figure 1.6.

Le mode particulier de développement des médicaments en psychiatrie fait porter aux études chez l'animal une plus lourde responsabilité. Ils doivent déterminer quelles molécules candidates doivent poursuivre leur développement chez l'homme sans qu'un tri aussi efficace par l'analyse des mécanismes physiopathologiques ou le criblage à haut débit n'ait été fait au préalable.

Le fait que la psychiatrie ait un taux de succès en phase II et III parmi les plus faibles de toutes les spécialités médicales suggère qu'il existe un problème dans la confirmation chez l'homme des résultats obtenus chez l'animal.

## 1.5 Résumé et conclusion

L'innovation thérapeutique en psychiatrie, qui avait bénéficié de découvertes fortuites dans les années 50 est donc faible depuis cette date. Si ces médicaments ont initialement représenté une avancée majeure dans la prise en charge des patients, les innovations ultérieures n'ont bien souvent été que mineures et l'efficacité des médicaments actuellement disponibles est encore largement insuffisante.

Malgré ce constat et bien que le marché des psychotropes soit important en volume, les laboratoires pharmaceutiques se désengagent de la recherche de nouveaux traitements en psychiatrie.

Une des raisons en est la faible probabilité qu'une molécule candidate issue de la recherche préclinique passe avec succès les différentes étapes des études cliniques.

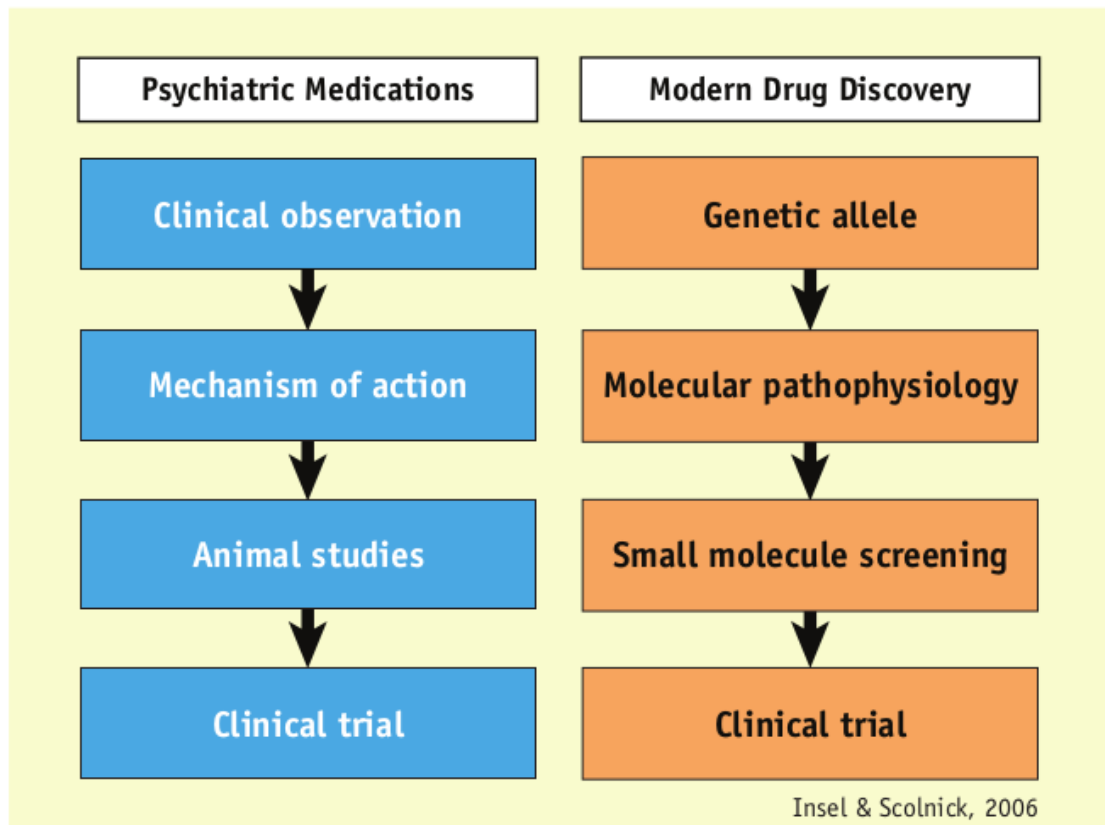


FIGURE 1.6 – Voies de développement de nouveaux médicaments, comparaison de la psychiatrie et des autres spécialités médicales. La voie traditionnelle de développement de médicaments en psychiatrie a commencé par la découverte fortuite de molécules efficaces, puis l'étude de leur mécanisme d'action. La sélection de molécules se fait ensuite par criblage chez l'animal puis évaluation clinique, sélectionnant ainsi des médicaments possiblement plus puissants, mais de même mécanisme d'action. A l'inverse, la voie moderne de développement de médicaments commence par l'identification d'une cible moléculaire ou cellulaire sur la base d'une physiopathologie. Un criblage à haut débit de molécules sur cette cible est fait sur des bibliothèques de molécules. Des molécules sélectionnées pourront être raffinées par génie chimique pour améliorer leur interaction avec la cible, avant d'entrer en phase de test préclinique puis clinique. (Insel et Scolnick, 2006)

C'est la démonstration de l'efficacité chez l'homme de molécules candidates validées chez l'animal qui fait souvent défaut. Le modèle animal joue en psychiatrie un rôle plus important que dans d'autres spécialités du fait de l'insuffisance actuelle des connaissances sur les causes et mécanismes des pathologies psychiatriques.

En conséquence, il convient de s'interroger sur les limites de l'expérimentation chez l'animal en psychiatrie et sur la validité des modèles animaux de pathologies psychiatriques, ce que nous ferons dans le prochain chapitre.



## Chapitre 2

# Les limites des études animales en psychiatrie

Les modèles animaux de pathologies psychiatriques consistent en des procédures appliquées à des animaux de laboratoire qui entraînent des changements de comportement dont on souhaite qu'ils soient *homologues* à certains aspects du trouble psychiatrique étudié. Ces animaux peuvent alors être utilisés comme outil expérimental pour l'étude de la pathologie chez l'homme ou de l'effet de médicaments (Willner, 2015).

Ces modèles se sont beaucoup développés au cours de vingt dernières années et ils sont très utilisés. Voyons d'abord de quelle façon évaluer la validité des modèles animaux, en particulier en psychiatrie.

### 2.1 Critères d'évaluation des modèles animaux

De nombreux critères différents ont été utilisés pour évaluer la qualité des modèles animaux, comme la réponse similaire aux traitements (McKinney et Bunney, 1969) ou la similitude des symptômes (Abramson et Seligman, 1977). Le concept d'homologie en particulier (plutôt que d'analogie) a été largement discuté et commenté (Robbins, 1998). Les changements induits par ces procédures doivent être *homologues* aux conditions psychiatriques, pas seulement analogues, en ce sens qu'on souhaite que le même processus soit à l'œuvre chez l'animal et chez l'homme.

Cependant, ces critères sont restés très peu consensuels jusque dans les années 1980. Willner, à cette époque, a proposé une analyse de la validité des modèles animaux sur trois dimensions qui sont depuis restées relativement consensuelles. Ces trois dimensions sont la validité prédictive, la validité apparente et la validité de construit (ou validité conceptuelle) (Willner, 1986, 1984). Ces dimensions montrent une conception du modèle animal dont on attend qu'il reproduise les mécanismes

pathologiques (modèle de la maladie), qu'il en mime les symptômes, que les liens entre mécanismes et symptômes y soient identiques et que les traitements y aient un effet comparable à ce qu'on observerait chez l'homme (Figure 2.1).

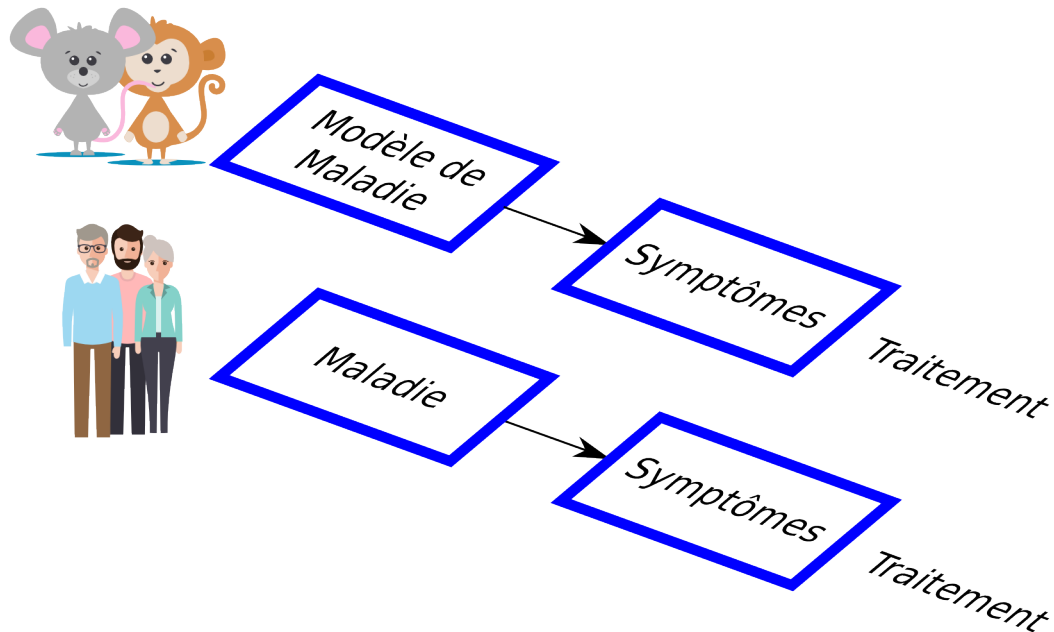


FIGURE 2.1 – Modèle animal classique pouvant correspondre aux trois critères de validité principaux énoncés par Willner. Dans ce modèle minimaliste, la validité de construit correspond à la modélisation du processus pathologique (maladie) et à la façon dont il produit les symptômes ; la validité apparente correspond à la similitude des symptômes ; la validité prédictive correspond à la capacité du modèle animal à prédire correctement l'effet des traitements chez l'homme. D'après Willner (1986)

Nous passons en revue ci-dessous ce que recouvrent ces trois grandes catégories de validité et introduisons ensuite un développement plus détaillé prenant en compte certains aspects spécifiques des pathologies psychiatriques.

### 2.1.1 Validité prédictive

La validité prédictive (“predictive validity” en anglais) indique dans quelle mesure une intervention qui modifie le comportement mesuré chez l'animal (en aggravation ou en amélioration) modifiera également, dans le même sens et dans la même proportion, ce qui est observé chez l'homme (Willner, 1986). En pratique, la validité prédictive indique dans quelle mesure la réponse aux médicaments dans le modèle animal est comparable à la réponse aux médicaments pour l'homme.

Ce type de validité est particulièrement important pour le problème qui nous intéresse : le faible rendement du développement de nouveaux médicaments, en

particulier du fait d'échecs en phase II des essais cliniques.

Dans cette optique, cette validité doit être double :

- d'une part la validité prédictive *positive* doit être bonne : un médicament qualifié d'efficace chez l'animal doit avoir une forte probabilité d'être efficace chez l'homme. A défaut, le risque est de poursuivre le développement du médicament, d'encourir les coûts pour finalement découvrir, à une phase tardive, que la molécule n'est pas efficace chez l'homme.
- d'autre part la validité prédictive *négative* doit également être bonne : un médicament qualifié d'inefficace dans le modèle animal doit avoir une forte probabilité de ne pas être efficace chez l'homme. A défaut, le risque est de disqualifier au stade préclinique des candidats médicaments qui auraient pu se révéler efficaces chez l'homme.

### 2.1.2 Validité apparente

La validité apparente ("face validity" en anglais) renvoie à la similitude phénoménologique entre le modèle et la pathologie étudiée (Willner, 1986).

Une difficulté particulière vient du fait que certains symptômes psychiatriques ne sont connus que par le rapport subjectif fait par le sujet. Ainsi en est-il par exemple du vécu de centralité ressenti par certains patients aux stades précoces de la schizophrénie, ou d'un sentiment de vacuité ou d'auto dépréciation rencontré chez certains patients déprimés. Ces symptômes semblent ne pas pouvoir être étudiés dans le cadre de modèles animaux. Dans certains cas cependant, on peut trouver chez l'homme des corrélats comportementaux de ces états subjectifs. C'est ce que fait le clinicien lorsqu'il note des particularités comportementales du sujet, comme des barrages, arrêts abrupts dans le discours, ou des fixations avec poursuite oculaire, interprétés comme des signes comportementaux en faveur d'hallucinations, alors même que le patient peut parfois nier présenter ces symptômes.

Si certains symptômes sont purement subjectifs, d'autres peuvent avoir des corrélats comportementaux qui font qu'une transposition dans le modèle animal est possible. Ainsi, différents tests comportementaux sont réputés reproduire tel ou tel symptôme du spectre dépressif, par exemple :

- le test de nage forcée (Porsolt et al., 1978) ou de la suspension par la queue (Steru et al., 1985) serait l'homologue de la résignation ;
- le splash test (Santarelli et al., 2003) serait l'homologue de la baisse de motivation ;
- le test de préférence au sucrose (Willner et al., 1987) serait l'homologue de l'anhédonie.

Cette homologie est cependant discutable et suspecte d'anthropomorphisme. Elle n'a de le plus souvent été établie qu'à posteriori, nombre de ces tests n'ayant pas

été retenus pour leur homologie avec des symptômes retrouvés chez l’homme mais pour leur capacité à identifier l’effet des premiers psychotropes efficaces découverts (Miczek, 2008).

### 2.1.3 Validité de construit

La validité de construit ou validité conceptuelle (“construct validity” en anglais) indique dans quelle mesure “le modèle mesure effectivement ce qu’il est censé mesurer” (Geyer et Markou, 1995). De façon plus explicite, ce critère requiert que les mécanismes biologiques perturbés dans la population clinique et chez l’animal soient similaire et que le lien entre ces mécanismes et l’expression sous forme de symptômes soit également similaire (Willner, 1986).

Prenons l’exemple de l’anhédonie, il faut que l’anhédonie soit centrale dans la pathologie et qu’elle soit présente dans le modèle animal. Si l’anhédonie est le résultat chez l’homme d’une dysfonction de circuits de la récompense incluant le striatum ventral, il faut alors de plus que la relation entre anhédonie et dysfonction dans le striatum ventral soit identique dans le modèle animal (Belzung et Lemoine, 2011).

Une validité de construit est plus facile à défendre lorsqu’on connaît les bases cérébrales des comportements mesurés. Si nous mesurons la réponse aux récompenses chez la souris dans un modèle de dépression. Mettre en évidence une diminution de cette réponse au plan comportemental seulement peut ne pas être suffisant. Encore faut-il montrer que cette diminution est liée à un effet spécifique du traitement des récompenses et non à un autre effet comme une augmentation de la sédation générale engendrée par le modèle.

La problématique de la validité de construit est forte en psychiatrie. En effet la capacité des modèles animaux à reproduire des pathologies bien délimitées, comme l’autisme versus de la schizophrénie, ou l’anxiété versus de la dépression, fait débat (Belzung et Lemoine, 2011).

### 2.1.4 Dimensions supplémentaires prenant en compte certaines particularités des pathologies psychiatriques

Les pathologies psychiatriques ont pour particularité d’être des pathologies complexes, avec des états prémorbides à risque, des interactions gène-environnement, des facteurs déclenchants. Belzung et Lemoine (2011) ont proposé un raffinement des trois grands types de validités, qui les détaille et y ajoute des dimensions pertinentes pour la psychiatrie.

Ils indiquent ainsi que l’animal ne doit pas seulement être un organisme qui mime la dysfonction chez l’humain : le processus même par lequel l’animal comme l’humain

arrivent dans cet état pathologique doit être similaire. L'interaction d'un organisme avec des facteurs environnementaux précoces le mettant en état d'organisme à risque puis, suite à l'exposition à des déclencheurs, en état d'organisme malade (Belzung et Lemoine, 2011). Cette approche précise et complexifie ce que doit être un modèle animal, mais en restant dans la lignée de la conception de Willner comme représenté Figure 2.2.

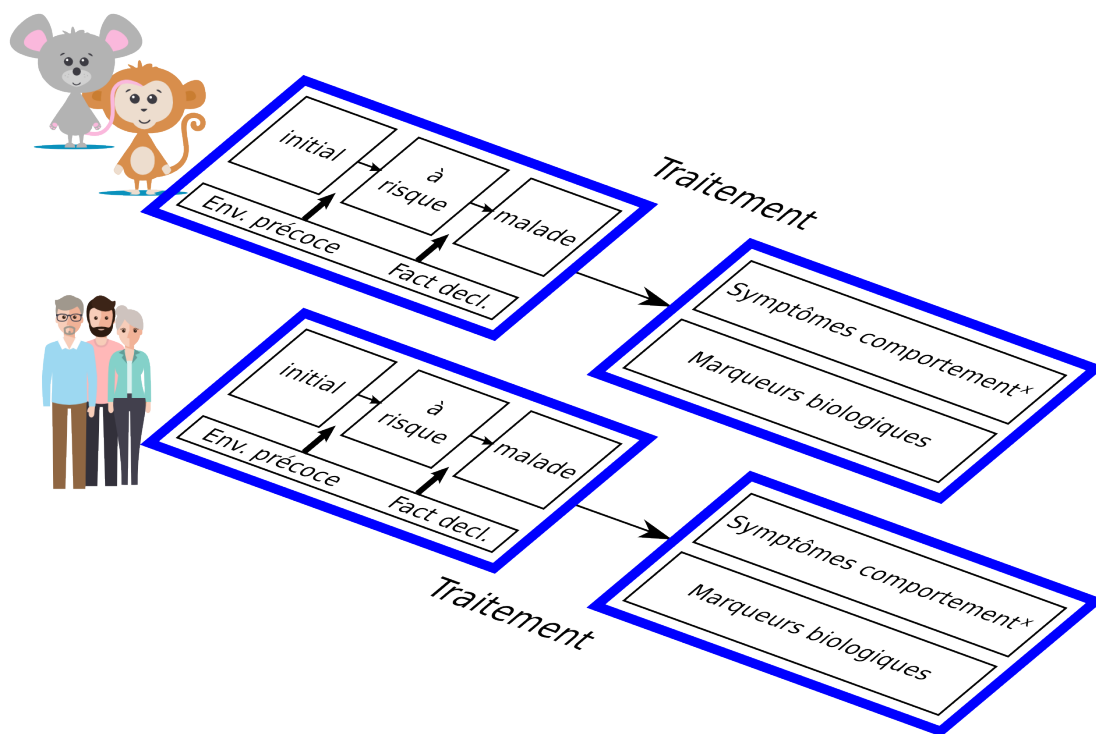


FIGURE 2.2 – Modèle animal prenant en compte l'aspect neurodéveloppemental des pathologies psychiatriques. Par rapport au modèle animal classique décrit plus haut, ce modèle explicite des dimensions particulièrement importantes pour la psychiatrie, en particulier le rôle de facteurs environnementaux précoces et de déclenchement, qui interagissent avec les variables génétiques (état initial) pour former un état prémorbide à risque pouvant basculer en état pathologique. D'après Belzung et Lemoine (2011)

Les auteurs distinguent neuf types de validités différentes, qu'ils regroupent en cinq groupes :

- validité homologique (espèce et lignée) qui décrit l'homologie de l'organisme initial ;
- validité pathogénique (ontopathogénique et de déclenchement) qui décrit la similitude dans la façon dont les facteurs environnementaux amènent cet organisme à l'état à risque puis à la maladie décompensée ;
- validité des mécanismes qui décrit la similitude dans les mécanismes qui sous-tendent l'expression symptomatique et qui répondent aux traitements ;

ces mécanismes peuvent être décrits aussi bien en termes neurobiologiques que cognitifs ;

- validité apparente (éthologique et biomarqueurs) qui décrit la similitude apparente dans l’expression de la maladie que ce soit en termes de comportement observable ou en termes de biomarqueurs (imagerie, biologie, . . . ) ;
- validité prédictive (induction et rémission) qui décrit la correspondance dans la relation entre les facteurs déclenchants et l’expression symptomatique d’une part et le traitement et la réponse au traitement d’autre part.

## 2.2 Versatilité des modèles animaux

De nombreux modèles animaux existent pour analyser les facteurs de risque de pathologies psychiatriques et les mécanismes physiopathologiques sous-tendant des symptômes particuliers (Millan, 2008). Nous ne prétendons pas faire ici un exposé complet de ceux-ci. Nous illustrons simplement la capacité des modèles animaux à simuler des aspects variés pertinents pour la psychiatrie.

### 2.2.1 Modélisation de facteurs d’environnement

Des facteurs d’environnement peuvent être modélisés de différente façon par des interventions de séparation maternelle, d’isolement social, par l’administration de substances comme les psychostimulants ou des dérivés du cannabis par exemple (Geyer et Ellenbroek, 2003).

L’isolement social chez l’animal de laboratoire, par exemple, déclenche des troubles du comportement. Les conséquences de cette manipulation, ont initialement été décrites simplement comme favorisant l’apparition de comportements agressifs. Puis un ensemble de modifications plus complet a été décrit avec des composantes comportementales et neurochimiques (Valzelli, 1973). Cependant, il semble que cette manipulation soit complexe à analyser. Sa signification et ses conséquences seraient par exemple très différentes chez des animaux territoriaux ou des animaux vivant en colonie (Miczek, 2008).

Des modèles de facteurs d’environnement plus complexes ont également été développés. C’est le cas du “stress chronique imprévisible modéré” (ou UCMS pour unpredictable chronic mild stress) dans lequel les animaux sont exposés, durant plusieurs semaines, à des facteurs de stress imprévisibles et variés. Cette procédure a été proposée comme modèle de symptômes dépressifs (Willner et al., 1987) avec une validité qui semble plus solide comme nous le verrons ci-dessous.

### 2.2.2 Modélisation de facteurs développementaux

Des facteurs développementaux peuvent être modélisés, par exemple par des lésions hippocampiques pratiquées chez les rats en période néonatale, en général à 1 semaine de vie. Ces rats développent par la suite une hypersensibilité à des tests par des psychostimulants ou des tests de stress (Lipska et al., 1993). On observe également chez ces rats des déficits dans le filtrage sensoriel (sensory gating) et dans les interactions sociales à l'âge adulte, qui sont des mesures d'intérêt dans la modélisation de la schizophrénie (Lipska et al., 1995 ; Sams-Dodd et al., 1997). Il a été démontré également que cette intervention avait pour conséquence à l'âge adulte des dysfonctionnements dans les circuits neuronaux dont on pense qu'ils sont également affectés dans la schizophrénie, à savoir le striatum et le cortex préfrontal (Lipska, 2004).

### 2.2.3 Modélisation de mécanismes cérébraux

Un des exemples les plus classiques dans ce domaine est l'identification de substances avec un potentiel antipsychotique qui agissent au niveau des récepteurs dopaminergiques D2, en cherchant dans quelle mesure elles corrigent ou non les effets de substances ayant un effet pro-dopaminergique.

Ce sont par exemple les tests de correction de l'hyperactivité locomotrice induite par les amphétamines, qui est en partie au moins médiée par une augmentation du relargage dopaminergique dans le circuit mésolimbique, particulièrement dans le nucleus accumbens (Costall et al., 1979), mécanisme que l'on pense à l'origine des symptômes positifs dans la schizophrénie.

Ce sont également les modèles de réversion des effets délétères sur le prepulse inhibition de l'apomorphine, ou des stéréotypies induites par cette même substance (Geyer et Markou, 1995).

Des modélisations d'hypothèses physiopathologiques plus modernes ont été faites, par exemple avec l'administration d'antagonistes NMDA à des rats, à partir de l'hypothèse qu'un dysfonctionnement des récepteurs NMDA existerait initialement dans le cours du développement des sujets qui développeront plus tard un tableau de schizophrénie avec dysfonction dopaminergique (Large, 2007). L'administration d'antagonistes NMDA chez le rat adulte produit des effets comportementaux qui sont considérés comme pouvant être homologues de symptômes positifs, négatifs et cognitifs, comme l'hyperactivité locomotrice, les stéréotypies, les modifications dans le prepulse inhibition, la réduction des interactions sociales, l'augmentation de l'immobilité dans les tests de nage forcée, les déficits cognitifs (Jones et al., 2008).

### 2.2.4 Modélisation de facteurs génétiques

De nombreuses manipulations génétiques ont été pratiquées, en particulier chez la souris pour évaluer différents aspects de pathologies psychiatriques (Willner et Belzung, 2015).

Dans le domaine de la dépression et de l'anxiété par exemple, des manipulations génétiques ont été effectuées sur des souris pour étudier la contribution de certains sous-types spécifiques de récepteurs à la sérotonine. Des souris chez lesquelles le récepteur à la sérotonine 5HT-1A n'était spécifiquement pas exprimé ont été créées et testées au plan comportemental. Les auteurs montrent que les souris dépourvues du récepteur 5HT-1A présentent une diminution des activités exploratoires et une augmentation de la peur des environnements aversifs. Ces résultats ont apporté un argument pour l'implication des récepteurs 5HT-1A dans les phénomènes d'anxiété (Ramboz et al., 1998).

Dans le champ de la schizophrénie, de nombreux facteurs génétiques peuvent être testés comme des délétions ou la surexpression de gènes de susceptibilité particuliers, en partant par exemple de ceux identifiés dans les études génétiques humaines. C'est ce qui a été fait avec la mutation DISC-1 (Disrupted In Schizophrenia). Les auteurs ont pu retrouver, chez des souris chez qui cette même mutation était provoquée, des anomalies comportementales comme une hyperactivité spontanée dans les tests en espace ouvert, une altération des interactions sociales ou un déficit en mémoire spatiale (Pletnikov et al., 2008).

### 2.2.5 Des combinaisons possibles

Nous redisons ici que les exemples donnés ci-dessus ne sont qu'un très faible échantillon des modèles existant. De nombreux autres modèles existent dans ce champ très actif, comme la modélisation d'interventions non pharmacologiques, la modélisation de l'action à long terme de médicaments, la modélisation de l'action des médicaments dans des populations particulières... (Millan, 2008).

De plus, du fait que certaines pathologies psychiatriques sont probablement le résultat d'une multitude de facteurs en interaction, des modèles particuliers où plusieurs facteurs sont manipulés simultanément sont possibles et probablement indispensables.

Un exemple d'une telle combinaison est l'exposition au stress de rats adolescents pour lesquels une lésion de l'hippocampe a été pratiquée en période néonatale (Millan, 2008). Un autre possibilité est la combinaison de facteurs génétiques avec des interventions environnementales comme par exemple l'administration de dérivés du cannabis à des souris génétiquement modifiées pour être knockout pour les récepteurs N-Méthyl-D-Aspartate (NMDA) (Millan, 2008). Des modèles plus complexes encore ont été développés plus récemment comme l'induction de réponses

immunes anormales durant la période périnatale chez des souris génétiquement modifiées pour exprimer des gènes de susceptibilité comme DISC-1 (Nagai et al., 2011)

Les possibilités en particulier des manipulations génétiques semblent grandes. A titre d'exemple, plus de 1200 modèles de souris transgéniques ont été sélectionnés et phénotypés (Insel et Landis, 2013). Ces modèles se sont progressivement raffinés. Il est ainsi maintenant possible de tester le rôle de certains gènes dans des cellules spécifiques ou à des moments précis du développement (Insel et Landis, 2013). C'est le cas par exemple pour la mutation DISC1 dont il est possible de limiter l'expression aux régions frontales du cerveau des souris (Pletnikov et al., 2008).

## 2.3 Des améliorations aux résultats insuffisants

### 2.3.1 Les différents types de validité de ces modèles sont connus

La validité apparente, la validité de construit et la validité prédictive sont devenues une formalisation consensuelle, qui n'est pas restée un outil théorique, mais est présente à l'esprit des chercheurs qui travaillent dans le champ de l'expérimentation animale. Willner se félicitait d'ailleurs que nombre d'auteurs dans le champ des modèles animaux commencent leur discussion par l'évaluation des différents types de validité de leur modèle (Willner et al., 2008).

L'hyperactivité induite par les amphétamines est-il un modèle valide de psychose induite ? Cette procédure manque sévèrement de validité apparente, mais en termes de validité prédictive pour la mesure de l'action de molécules à activité antipsychotique, elle représente un screening initial simple pour des substances pour lesquelles on recherche une activité antagoniste D2 (Miczek, 2008).

Une intervention comme le "stress chronique imprévisible modéré", permet d'observer des comportements de type anhédonie, dont on pense qu'ils sont centraux dans la dépression par exemple, et auraient ainsi une bonne validité apparente (Belzung et Lemoine, 2011). Cette intervention a également été associée à des modifications dans la transmission monoaminergique, et pourrait ainsi modéliser certains aspects des mécanismes de la dépression, ce qui est un élément de validité de construit (validité mécanistique) (Millan, 2008). Cependant, cette validité interagirait avec d'autres dimensions participant à la validité de construit, en particulier la race ou de la lignée d'animal sélectionnée (validité d'espèce ou de lignée) (Belzung et Lemoine, 2011). Enfin, l'effet de cette procédure en diminution de la préférence au sucrose est corrigé après une administration chronique d'antidépresseurs, mais pas après une administration aiguë, ce qui est un élément de validité prédictive (de rémission) (Miczek, 2008).

Le test de nage forcée à l'inverse a une validité apparente et de construit faible (Cryan et al., 2002). Le comportement mesuré n'a pas de similitude immédiate avec les symptômes humains dans la dépression, même si le terme anthropomorphe de résignation est parfois avancé. Le test s'effectuant le plus souvent sur l'animal sain, les mécanismes cérébraux en jeu ont peu de lien à priori avec la dépression. En revanche, ce test, particulièrement dans sa version modifiée (exposée plus bas), a une très bonne valeur prédictive, identifiant correctement comme efficaces un grand nombre de traitements antidépresseurs utilisés cliniquement, et rejetant des thérapeutiques inefficaces chez l'homme (Cryan et al., 2002).

### **2.3.2 Des améliorations sont régulièrement proposées pour en améliorer la validité**

Chaque test et modèle a donc des qualités et des insuffisances sur les différentes dimensions de validité de Willner. Ce fait est connu de chercheurs ce qui est un moteur poussant à l'amélioration de ces modèles, par le développement de variantes parfois subtiles, mais aux conséquences importantes.

Dans la procédure du "stress chronique imprévisible modéré" (Willner, 1997), les effets attendus qui sont l'obtention d'un modèle d'anhédonie objectivé par une diminution de la préférence au sucrose n'ont pas toujours été répliqués de façon fiable (Forbes et al., 1996 ; Harris et al., 1997 ; Phillips et Barr, 1997). Ceci a favorisé le développement de nombreuses variantes de cette procédure complexe pour tenter d'améliorer sa validité (Miczek, 2008).

Dans le test de nage forcée, Detke et collaborateurs ont proposé qu'il était possible de mesurer séparément les activités de nage et d'escalade des parois, la première mesure corrélant avec l'activité sérotoninergique des médicaments testés et la seconde avec l'activité noradrénergique (Detke et al., 1995). Les résultats ayant amené cette proposition de modification sont illustrés Figure 2.3.

Une autre mesure d'amélioration visant à palier la validité limitée de chaque test pris individuellement est d'utiliser plusieurs tests conjointement pour cribler des molécules candidates. C'est ce qui est fait en pratique, les qualités d'un modèle ou test venant compenser les faiblesses d'un autre test sur une dimension particulière de validité. Pour améliorer la valeur prédictive des tests précliniques, des auteurs recommandent d'utiliser un nombre "aussi grand que possible" de tests et modèles (Nestler et Hyman, 2010).

Malgré l'évaluation faite par les chercheurs de la validité de leurs outils précliniques et les mesures d'amélioration entreprises depuis plusieurs d'année, les faibles taux de succès en phase II et III des essais cliniques pour les médicaments du champ de la psychiatrie (Hay et al., 2014 ; Thomas et al., 2016), sont évocateur d'une valeur prédictive qui reste insuffisante.

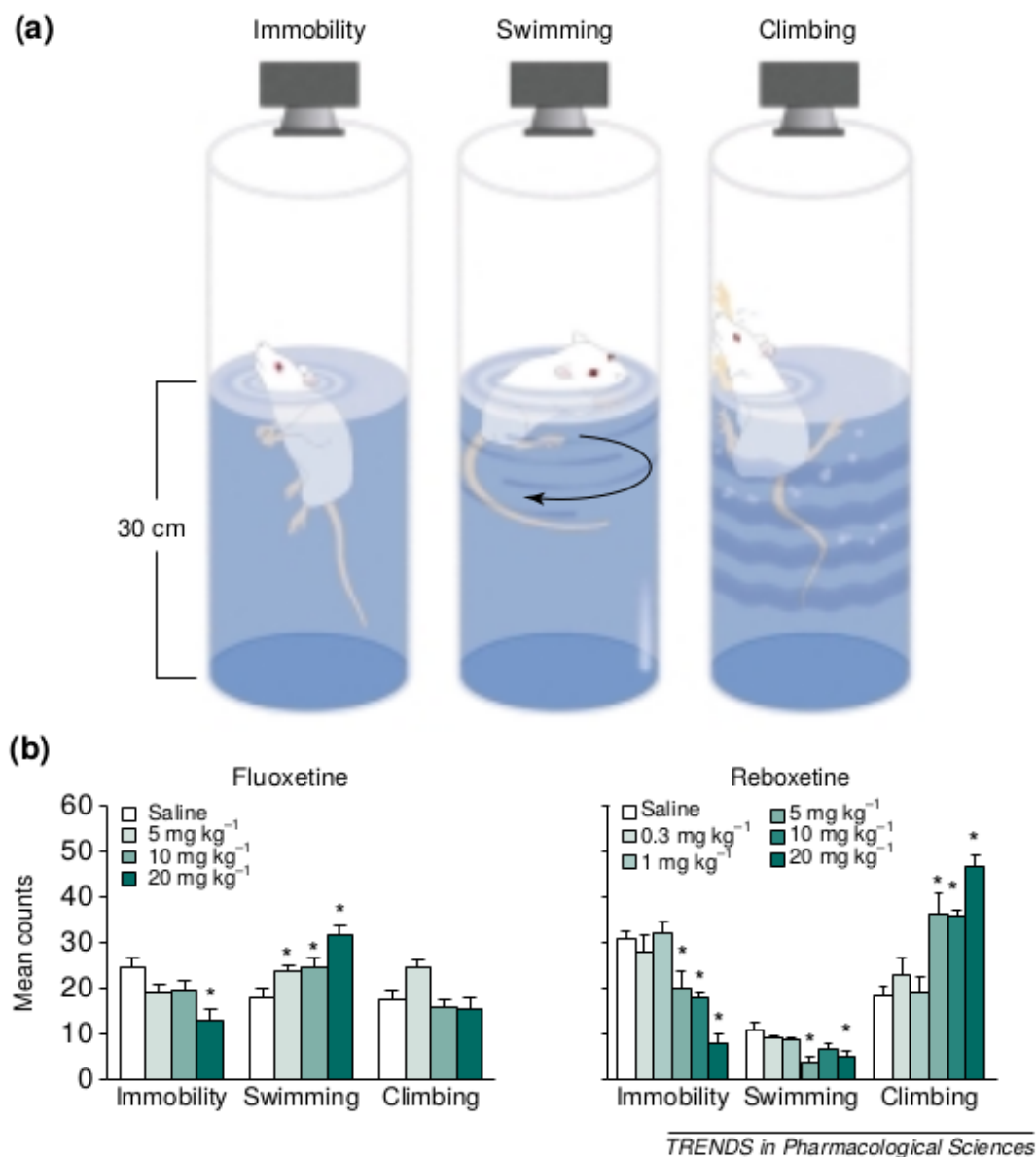


FIGURE 2.3 – Version modifiée du test de nage forcée améliorant sa valeur prédictive. Des rats dans une tâche de nage forcée présentent au moins trois types de comportement : rester immobile, nager et escalader. Différents types de molécules, un inhibiteur de recapture de la sérotonine, la fluoxétine, et un inhibiteur de la recapture de la noradrénaline, la reboxétine, ont des effets différents sur ces différents comportements. Les psychotropes qui potentialisent préférentiellement la transmission sérotoninergique augmentent le comportement de nage, tandis que ceux qui potentialisent préférentiellement la transmission noradrénergique augmentent le comportement d’escalade. (Cryan et al., 2002)

### **2.3.3 Mais la validité globale des modèles animaux est critiquée**

Si les études chez l'animal ont largement contribué à améliorer notre connaissance des pathologies en général, leur valeur prédictive pour l'efficacité des stratégies thérapeutiques reste discutée. Plusieurs revues systématiques reconnaissent aux modèles animaux une capacité limitée à prédire l'efficacité des molécules chez l'homme (Hackam, 2007 ; Hackam et Redelmeier, 2006 ; Knight, 2008 ; Perel et al., 2007).

Une de ces revues a consisté à reprendre systématiquement toutes les études thérapeutiques chez l'animal publiées entre 1980 et 2000 dans les sept journaux les plus prestigieux (Science, Nature, Cell, Nature Medicine, Nature Genetics, Nature Immunology, et Nature Biotechnology). Les auteurs se sont restreints aux articles ayant obtenu plus de 500 citations, avec l'idée que des articles aussi cités dans des journaux aussi prestigieux avaient plus de chance d'avoir été testés dans des études ultérieures chez l'homme. Or étonnamment, les auteurs n'ont trouvé une réplication publiée chez l'homme que pour 37% de ces études animales (Hackam et Redelmeier, 2006). Ils concluent qu'il faut être prudent sur l'extrapolation aux humains des découvertes chez l'animal même lorsqu'elles sont publiées dans les revues les plus sérieuses et sont très citées (Hackam et Redelmeier, 2006).

Ces revues sont une indication générale des limites dans la validité des modèles animaux, mais des exemples d'échecs identifiés avec des molécules précises est également informatif. Nous en exposons deux exemples ci-dessous.

### **2.3.4 Exemple d'échec de transposition des résultats pré-cliniques à l'homme**

Voyons un exemple d'échec dans la transposition à l'homme de résultats encourageants chez l'animal dans le champ de la psychiatrie.

La flibansérine est un médicament initialement proposé comme antidépresseur, et qui possède une activité agoniste 5HT-1A et antagoniste 5HT-2A. Cette molécule avait été jugée active dans dix des treize modèles animaux testés (Borsini et al., 2002). Les résultats étaient en effet positifs au learned helplessness test chez le rat, la bulbectomie chez le rat, le chronic mild stress test chez le rat et chez la souris, le comportement muricidal (comportement d'attaque de souris) chez le rat, le sevrage aux amphétamines chez le rat, la latence du sommeil REM chez le rat, le fixed ratio test chez la souris et le forced swim test chez la souris (Borsini et al., 2002). Ces tests montraient même que l'action antidépressive pouvait être plus rapide qu'avec l'imipramine ou la fluoxétine (Borsini et al., 2002). Malgré ces résultats positifs chez l'animal sur un grand nombre de modèles différents et dans des espèces animales différentes, les résultats des études cliniques de phase II ont été négatifs

(Cryan et al., 2008).

Un autre exemple est tiré de la recherche sur le développement de traitements non monoaminergiques de la dépression. Les antagonistes des récepteurs CRF1 avaient obtenu des résultats encourageants en préclinique. Les résultats étaient positifs dans plusieurs essais biologiques, le test de nage forcée (Jutkiewicz et al., 2005), le test de suspension par la queue (Nielsen et al., 2004) et le test DRL-72 (Louis et al., 2006), dans deux espèces, le rat et la souris. Les résultats étaient également positifs dans plusieurs modèles de dépression, le stress chronique imprévisible modéré (Dournes et al., 2013; Ducottet et al., 2003; Surget et al., 2011), le stress social (Wood et al., 2012), et dans un modèle génétique de dépression, la lignée de rats Flinder (Overstreet et Griebel, 2004). Si la première étude clinique, sur 24 patients a donné un résultat positif, les cinq études cliniques suivantes publiées n'ont pas retrouvé d'efficacité de ces antagonistes des récepteurs CRF1 (pour une revue, voir Griebel et Holsboer (2012)).

## 2.4 Des limites dans l'amélioration possible

Nous voyons donc qu'il existe un cadre théorique permettant d'évaluer la validité des modèles animaux, que ce cadre est effectivement utilisé, les modèles sont évalués et progressivement améliorés. Malgré tout, les échecs restent nombreux. Existe-t-il des limites fondamentales qui expliqueraient que malgré les précautions prise, la validité des modèles animaux sera toujours limitée ?

### 2.4.1 Limites dans la validité homologique

Les modèles animaux ont jusque là proposé une approche réductionniste des troubles mentaux, en s'intéressant à des parties du cerveau, et en s'intéressant à des modifications ciblées de type activation ou inhibition d'un récepteur particulier. Cette approche a été efficace et a permis de comprendre des médicaments essentiels comme les inhibiteurs de la recapture de la sérotonine. Cependant, les succès ne se sont pas poursuivis au même rythme, et il semble qu'il y ait au moins deux types de problèmes.

D'une part, même en restant au niveau très local de sous-parties de la cellule ou de récepteurs qui pourraient être communs entre l'humain et l'animal étudié, il apparaît maintenant que de très légères variations entre les séquences d'acides aminés chez le rat et l'humain peuvent produire des différences d'importance inattendue dans la pharmacologie des récepteurs. Ainsi, Hirst et collaborateurs ont montré que pour l'étude du système sérotoninergique, il est peu probable que les souris knockout de type 6 puissent être utiles pour valider les récepteurs de type 6 chez l'humain comme cible thérapeutique du fait d'une différence inattendue entre les espèces, à la fois en termes de distribution locorégionale de ces récepteurs mais aussi en

termes de profil pharmacologique (Hirst et al., 2003). Il est clair que la non prise en compte de ces différences fines, maintenant identifiées, peut être à l'origine de résultats très différents dans le profil d'efficacité des molécules candidates entre l'expérimentation chez l'animal et l'expérimentation chez l'homme.

Il est possible de mieux s'interroger sur l'adéquation de l'espèce choisie. Par exemple de nouveaux antidépresseurs agissant sur la famille des tachykinines, qui interagissent avec les récepteurs NK1, NK2 et NK3 ont été étudiées (Regoli et al., 1994). Or si les rats ont bien des récepteurs NK3 dans certaines régions cérébrales comme l'amygdale, l'aire tegmentale ventrale, il existe des différences importantes avec l'homme dans d'autres régions comme l'hypothalamus (Griebel et Beeské, 2012). En revanche le cochon d'Inde présente un profil réceptologique pour le NK3 qui est plus semblable à l'homme (Griebel et Beeské, 2012).

D'autre part, certains dénoncent maintenant cette approche réductionniste comme délétère, indiquant qu'elle passe à côté des propriétés émergentes qui caractérisent les systèmes complexes comme le système nerveux central (Van Regenmortel, 2004). Il faut alors se poser la question de l'homologie entre espèces à un niveau différent du niveau cellulaire ou récepteuriel.

Or si les différences de structure entre le cerveau de l'homme et du rat sont évidentes, elles restent notables avec celui du macaque rhésus, pourtant considéré comme un de ses plus proches apparentés. L'aire visuelle représente 10% du cerveau chez le macaque contre seulement 3% chez l'homme (Kreiman et al., 2002). Le cerveau humain est architecturalement et biochimiquement plus complexe, avec un temps de croissance plus long (Graaf-Peters et Hadders-Algra, 2006), un nombre de synapses par neurone plus important (Kreiman, 2007). Il y aurait ainsi plus de 90 gènes impliqués dans divers mécanismes neuronaux et qui seraient différents entre les singes et les humains (Kreiman et al., 2002).

Un exemple de différence entre espèces qui concerne pleinement les circuits impliqués dans la prise de décision est l'architecture du cortex orbito-frontal. Les études architectoniques initiales ne retrouvaient pas d'homologie claire entre les structures comme le vmPFC ou l'OFC entre l'homme et le macaque (Carmichael et Price, 1994).

Des études d'imagerie plus récentes ont évalué de façon quantitative le profil de densité de différentes aires du cortex comme moyen de définir de façon quantitative les différentes zones architectoniques. Et ces études ont montré un parallèle dans l'organisation de l'OFC entre l'homme et le singe comme illustré Figure 2.4. Il persiste cependant des différences, par exemple la densité des neurones pyramidaux est environ 70% plus importante chez l'homme que chez le singe (Wallis, 2012).

L'homologie entre l'OFC chez le singe et chez le rat est quant à lui moins clair. En effet, il y a dans le cortex préfrontal du singe et de l'homme un gradient postéro-antérieur avec à l'arrière du cortex agranulaire qui ne contient pas de couche IV, puis du cortex dysgranulaire contenant une couche IV rudimentaire, et enfin à

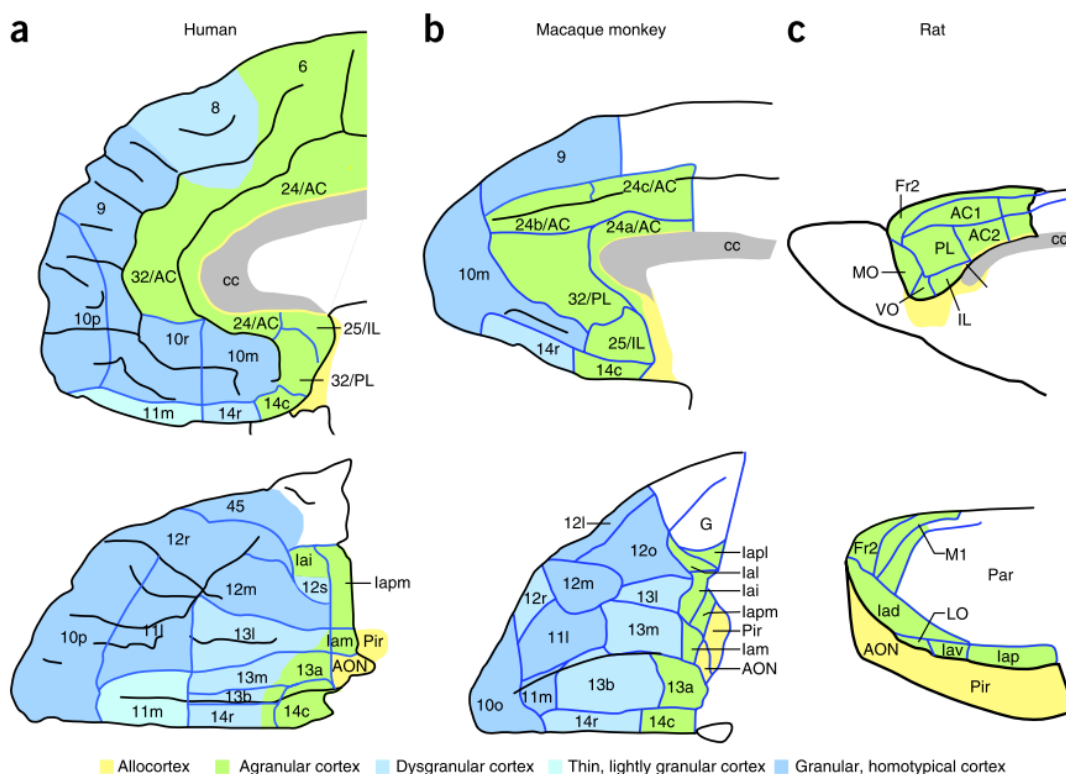


FIGURE 2.4 – Anatomie comparée du cortex frontal chez l'homme, le singe et le rat. (a,b) Carte architectonique des faces médianes (haut) et orbitales (bas) du lobe frontal chez l'homme (a) et chez le singe (b). (c) Face médiane (haut) et latérale (bas) du cortex frontal chez le rat. AC, anterior cingulate area; AON, anterior olfactory nucleus; c, caudal; cc, corpus callosum; Fr2, second frontal area; I, insula; i, inferior; Ia, agranular infralimbic cortex; IL, infralimbic cortex; l, lateral; LO, lateral orbital area; m, medial; M1, primary motor area; MO, medial orbital area; o, orbital; p, posterior; Par, parietal cortex; Pir, Piriform cortex; PL, prelimbic cortex; r, rostral; s, sulcal; v, ventral; VO, ventral orbital area. (Wallis, 2012)

l'avant du cortex granulaire, qui contient une couche IV bien développée. Alors que ce gradient est clair chez les primates et l'homme, il est absent chez le rat qui ne contient que du cortex agranulaire (Wallis, 2012).

### 2.4.2 Limites dans la validité ontopathogénique et mécanistique

Certains auteurs sont optimistes sur la capacité des modèles animaux à s'informer des avancées récentes en génétique pour "mimer" des phénomènes biologiques associés aux tableaux psychiatriques, en particulier les anomalies développementales du cortex, les anomalies de la transmission glutamatergique, et les caractéristiques génétiques de certains traits comportementaux (Lipska et Weinberger, 2000).

Cependant, pour reproduire un phénotype fiable chez l'animal, il faut que la mutation génétique ait une bonne pénétrance, comme c'est le cas par exemple pour le syndrome de l'X fragile (Moy et Nadler, 2008) ou le syndrome de Rett (Shahbazian et al., 2002). Or pour les troubles psychiatriques, aucun allèle à haute pénétrance n'a été identifié pour le moment. Ils sont associés à un grand nombre de variants génétiques communs et avec des variants génétiques rares, chacun ayant une valeur prédictive faible (Giegling et al., 2017).

Ainsi, différents patients avec un même diagnostic ont probablement un background génétique et une ontopathologie différents (Nestler et Hyman, 2010). À l'inverse, une mutation spécifique peut amener à des expressions phénotypiques variées, ce qui a été observé avec la mutation DISC-1 (Brandon et al., 2009).

Si l'on crée un modèle animal sur la base d'une de ces mutations, on aura au mieux un modèle d'un sous-groupe très restreint et d'une forme particulière de la pathologie cible. On n'aura pas d'assurance que les molécules efficaces pour ce modèle d'une forme particulière de la pathologie auront une efficacité qui s'étendra aux autres formes de la pathologie. La validité de construit et la validité apparente pour les modèles d'inactivation génique chez la souris comme modèle d'une pathologie psychiatrique telle que la dépression ou la schizophrénie sont donc loin d'être garanties (Belzung, 2014).

A titre d'illustration, la technique d'inactivation génique a été d'un apport limité pour améliorer la compréhension des mécanismes neurobiologiques sous-tendant les comportements agressifs. Des gènes ont été retrouvés sur chacun des chromosomes, leur manipulation avait un impact en augmentation ou en diminution sur le niveau d'agressivité mesuré dans le comportement des souris, comme illustré Figure 2.5.

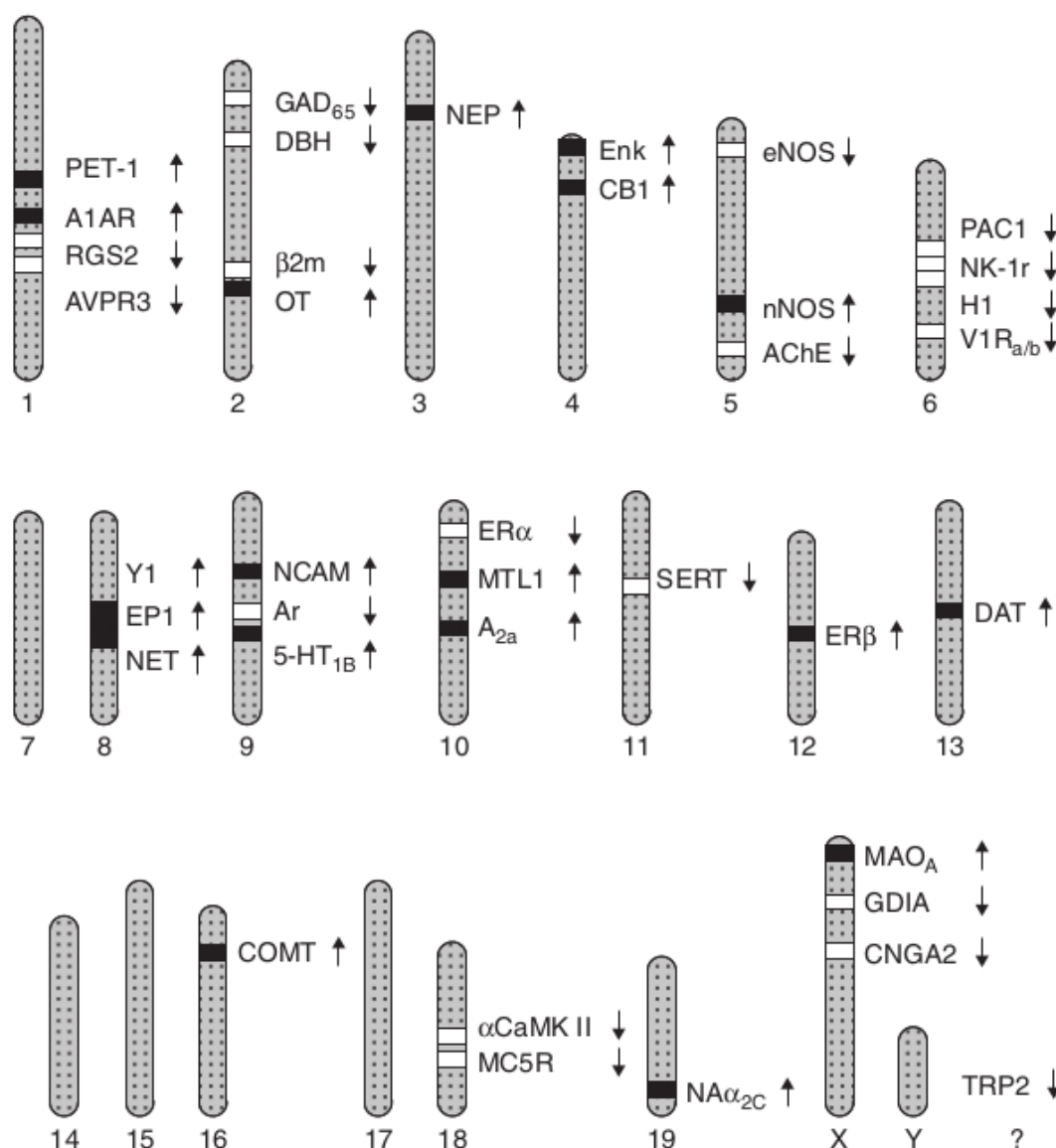


FIGURE 2.5 – Délétions génétiques affectant des phénotypes agressifs chez la souris. Représentation de 21 chromosomes et de la localisation approximative des gènes (barres horizontales) qui, inactivés, ont amené une augmentation (barre noire, flèche vers le haut) ou une diminution (barre blanche, flèche vers le bas) du niveau d'agressivité chez la souris. Les protéines cibles de la délétion sont indiquées sous forme abrégée. Abréviations : 5-HT<sub>1B</sub>, serotonin receptor 1B; A1AR, adenosine receptor A1; A2a, adenosine receptor 2A; AchE, acetylcholinesterase; alphaCaMK II, alpha-calcium-calmodulin kinase II; Ar, aromatase; AVPR3, arginine vasopressin 1B receptor; beta2m, beta2-microglobulin; CB1, cannabinoid receptor 1; CNGA2, cyclic nucleotide gated channel alpha 2; COMT, catechol-O-methyltransferase; DAT, dopamine transporter; DBH, dopamine beta hydroxylase; eNOS, endothelial nitric oxide synthase; ENK, enkephalin; EP1, prostaglandin E receptor 1; ERalpha, estrogen receptor alpha; ERbeta, estrogen receptor beta; GAD65, glutamic acid decarboxylase (65 amino acids); GDIA, guanosine diphosphate (GDP) dissociation inhibitor 1; H1, histamine receptor 1; MAOA, monoamine oxidase A; MC5R, melanocortin-5 receptor; MTL1, nuclear receptor subfamily 2, group E, member 1; Naalpa2C, adrenergic alpha receptor 2C; NCAM, neural cell adhesion molecule; NEP, neutral endopeptidase; NET, norepinephrine transporter; NK-1r, neurokinin receptor 1; nNOS, neuronal nitric oxide synthase; OT, oxytocin; PAC1, adenylate cyclase activating polypeptide 1 receptor 1; PET-1, ETS (E26 transformation specific) domain transcription factor; RGS2, regulator of G protein signaling; SERT, serotonin transporter; Trp2, transient receptor potential family 2; V1Ra/b, a cluster of vomeronasal receptor genes located on chromosome six, V1ra1-9 and V1rb1-4,7-9; Y1, neuropeptide Y receptor 2. (Miczek, 2008)

### 2.4.3 Limite dans la validité éthologique

Les limites de la validité apparente des modèles animaux ont déjà été évoquées plus haut.

Le problème existe non seulement au niveau des symptômes, mais également au niveau syndromique. Nous développons plus avant ce point dans l'évaluation des limites des études chez l'homme, mais l'évaluation critique des systèmes de diagnostic, qui classifient les pathologies psychiatriques en catégories sur la base de symptômes de surface, pose des limites sur la possibilité pour un modèle animal d'avoir une bonne validité dès lors qu'on cherche à modéliser l'ensemble du syndrome.

### 2.4.4 Limites dans la validité prédictive

Cette validité nous intéresse particulièrement puisqu'elle concerne la valeur prédictive de l'efficacité éprouvée chez l'animal d'un candidat médicament.

Un des problèmes principaux pour la validité prédictive vient de l'aspect circulaire qui existe entre le développement des modèles animaux et la validation des médicaments (Millan, 2008). En effet, les modèles animaux ont été créés pour caractériser l'effet des médicaments connus qui avaient été découverts de façon fortuite, les antipsychotiques et les antidépresseurs. Les molécules candidates sont passées au crible de ces modèles classiques ce qui a pour conséquence de sélectionner celles qui reproduisent les mêmes effets que les médicaments historiques, avec leur mécanisme connu et leurs insuffisances (Miczek, 2008). Ceci explique que les modèles animaux classiques actuels ont échoué dans la prédiction de l'efficacité chez l'homme de médicaments dont le mécanisme d'action était nouveau (Nestler et Hyman, 2010).

Des facteurs additionnels peuvent venir mettre un peu plus à mal la validité prédictive de la réponse à un médicament chez l'animal.

La pharmacocinétique peut être différente chez l'homme et l'animal et faire que les concentrations au site d'action soient notablement différentes chez l'animal et chez l'homme, ce qui peut rendre compte d'une différence d'efficacité.

Des problèmes de tolérance peuvent expliquer la faible validité prédictive dans certains cas. Une molécule sera parfois jugée non efficace chez l'homme car l'apparition d'effets indésirables lors des phases précoces des essais cliniques aura empêché que son administration à dose suffisamment élevée soit testée chez l'homme lors d'une phase plus tardive. Cette limitation des doses du fait d'effets secondaires inacceptables est moins fréquente chez l'animal (Willner et al., 2008).

Des phénomènes d'habituation seraient mal pris en compte dans la recherche préclinique (Bespalov, Müller, et al., 2016). Alors que les essais cliniques évaluant l'efficacité des médicaments sont conduits sur un minimum de plusieurs semaines,

la plupart des études précliniques testent l'efficacité des candidats médicaments en se limitant à des administrations aiguës ou subaiguës, comme illustré Figure 2.6. Or toute exposition répétée à une substance peut faire apparaître un phénomène d'habituation, c'est-à-dire une diminution de l'effet obtenu au cours du temps pour une dose de produit restant stable.

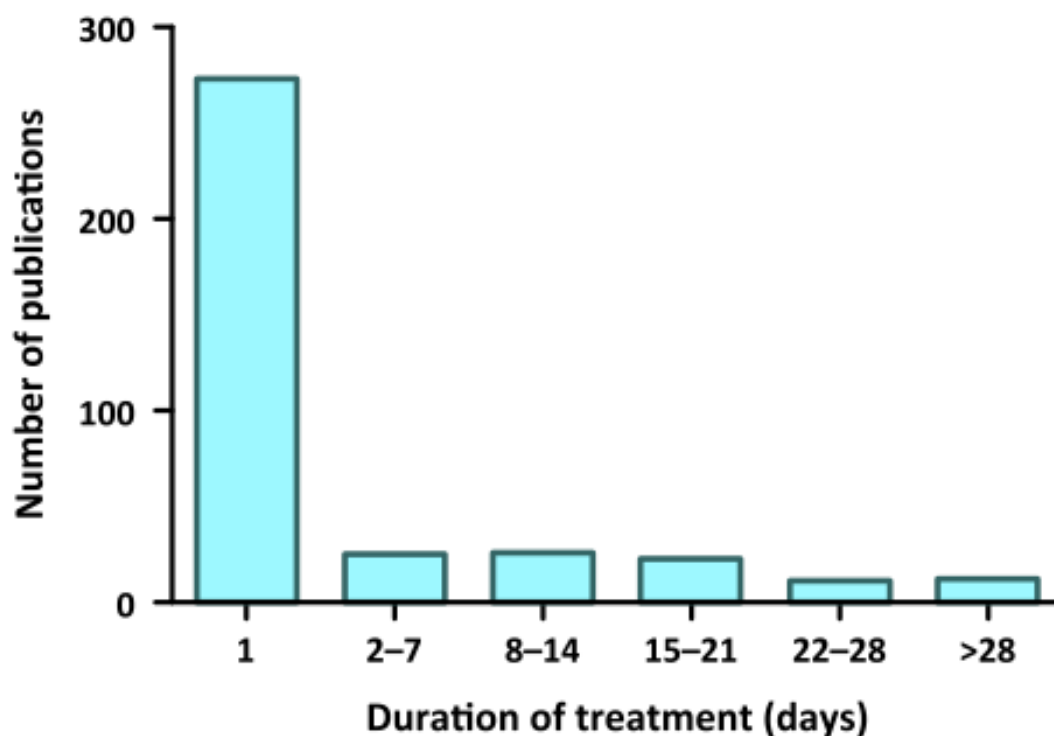


FIGURE 2.6 – Durée des traitements administrés dans les études précliniques publiées entre 2010 et 2015. L'analyse est extraite des publications effectuées dans les journaux suivants : Neuropsychopharmacology, Psychopharmacology, Journal of Pharmacology et Journal of Pharmacology and Experimental Therapeutics et se limitent aux publications concernant les rongeurs utilisés comme tests et modèles dans la recherche thérapeutique en psychiatrie (Bespalov, Müller, et al., 2016)

La standardisation des études précliniques avec des individus très proches (lignée, sexe, age, poids, ...) pourrait amener de faux positifs. Le résultat favorable pourrait être le fruit d'une conjonction particulière de facteurs confondants non contrôlés ayant eu une portée pratique sur l'échantillon étudié du fait de leur homogénéité, qui réduit la variabilité intra-expérience. Le résultat ne serait alors plus reproduit dès que les conditions de laboratoire seraient modifiées. Ce manque de généralisabilité pourrait être la cause d'un faible taux de réplication dans les études animales (Richter et al., 2009). Dans un travail ultérieur, ces auteurs ont expérimentalement renforcé leur thèse en montrant que l'introduction d'une variabilité même simple au sein d'une expérience améliore la reproductibilité inter-expériences (Richter et al., 2010).

## 2.5 Résumé et conclusion

Les modèles animaux bénéficient d'un cadre théorique clair permettant d'apprécier leur validité selon plusieurs axes, validité apparente, validité de construit et validité prédictive.

Les tests et procédures sont nombreux, variés, en constante évolution. Ils peuvent modéliser des aspects divers des pathologies psychiatriques comme certains comportements symptomatiques, des mécanismes supposés de pathologies, le rôle de facteurs d'environnement, l'aspect neurodéveloppemental... Les modèles explorant les anomalies génétiques en lien avec les pathologies psychiatriques est un domaine de recherche particulièrement intense.

Chacun de ces modèles possède ses propres forces et faiblesses selon le type de validité que l'on examine. Une améliorations progressives de ces modèles se fait par modification fine de leurs modalités pratiques. Pour palier la validité limitée de chaque test pris individuellement, plusieurs tests sont utilisés conjointement pour cribler des molécules candidates et tenter d'améliorer la valeur prédictive globale de la recherche préclinique.

Cependant, malgré ces efforts, la valeur prédictive des modèles animaux est critiquée et les exemples de transposition décevante à l'homme de résultats pourtant très prometteurs chez l'animal ne sont pas rares.

De nombreux facteurs, liés à la génétique, à l'anatomie, aux différences même fines dans la densité ou la répartition des récepteurs entre espèces peuvent limiter la validité des modèles animaux. A cela s'ajoutent des différences dans l'opérationnalisation des tests, dans les doses ou des durées de traitement non comparables aux essais chez l'homme.

Un problème important vient du fait que les modèles classiques utilisés en routine pour passer au crible les candidats médicaments ont été créés pour caractériser l'effet des premiers médicaments découverts. Les molécules candidates qui passent avec succès les tests préclinique sont donc celles qui ont des mécanismes d'action proches de ceux des médicaments connus, alors même que ces mécanismes sont insatisfaisants pour leur manque d'efficacité et leurs effets secondaires importants (Miczek, 2008). Cette circularité explique l'échec des modèles classiques à prévoir l'efficacité chez l'homme de médicaments au mécanisme d'action nouveau (Nestler et Hyman, 2010).

La conséquence de ces limites n'est pas seulement que des molécules sont considérées comme prometteuses de façon non fondée, sur la base des essais sur l'animal. Certaines molécules, qui pourraient être de bons médicaments, sont également écartées sur la base d'essais négatifs sur l'animal. De nombreuses molécules qui ont été écartées pourraient, pour certains auteurs, mériter une réévaluation (Bespalov, Steckler, et al., 2016).

Y a-t-il une alternative à cette approche ? Sydney Brenner, prix Nobel de physiologie / médecine en 2002 pour ses travaux sur *Caenorhabditis elegans*, comme modèle animal pour comprendre la régulation génétique de l'organogénèse et de la mort cellulaire programmée, a ainsi déclaré dans son discours du prix Nobel :

“I am convinced that we will make our significant discoveries in humans and that the mouse will be used to validate the human findings by genetic synthesis, much in the same way as the chemist confirms a structure analysis by chemical synthesis” (Brenner, 2003).

Cette position n'est pas isolée :

“It remains true that human disease is best studied in human subjects” (Van Regenmortel, 2004).

Ou encore :

“Finally, in some cases, the best”preclinical" models may be typical humans engaging in tasks that are disrupted by disorders" (NAMHC, 2010).

L'expérimentation chez l'homme serait-elle moins sujette à des limitations ?



# Chapitre 3

## Limites des études cliniques, l'apport des neurosciences cognitives

### 3.1 Limites liées à l'étude des patients

#### 3.1.1 Hétérogénéité

Les patients en psychiatrie présentent une hétérogénéité multidimensionnelle.

Il existe en effet une hétérogénéité entre patients, différents patients n'appartenant pas nécessairement au même sous-type de pathologie. Il existe également une hétérogénéité des symptômes, certains pouvant être plus spécifiques d'un diagnostic, d'autres appartenir à plusieurs catégories de diagnostic (Wardenaar et Jonge, 2013). Cette hétérogénéité est particulièrement importante en psychiatrie, avec une faible capacité des cliniciens à s'accorder sur les diagnostics posés (Freedman et al., 2013).

Il existe de plus une hétérogénéité temporelle. D'une part deux patients porteur d'un même diagnostic peuvent avoir des trajectoires d'évolution de la maladie très différentes (Wardenaar et Jonge, 2013). Ceci est particulièrement visible pour les pathologies schizophréniques ou pour les troubles dépressifs récurrents. D'autre part, un individu malade peut, sans changer de catégorie de diagnostic, présenter un tableau clinique très différent d'un moment à l'autre. Poursuivant sur l'exemple des pathologies schizophréniques, un patient a des périodes de rémission relative et d'aggravation symptomatique spontanées.

Les comorbidités sont également source d'hétérogénéité au sein d'une population pour laquelle le même diagnostic principal a été posé. Ces comorbidités sont la règle plus que l'exception en psychiatrie. Pour les patients présentant un diagnostic de schizophrénie, la probabilité qu'un diagnostic d'abus de substance est de 47% vie

entière, celle d'un trouble dépressif est de 50% (Buckley et al., 2009). Les troubles anxieux seraient également fréquents avec une prévalence estimée à 15% pour les troubles paniques, ou 23% pour les troubles obsessionnels compulsifs (Buckley et al., 2009).

Enfin, ces patients sont pour la plupart traités par des médicaments psychotropes qui peuvent influencer la mesure effectuée par le clinicien, que cette mesure soit une évaluation cognitive, une imagerie fonctionnelle, ou un prélèvement biologique. Ceci est là encore particulièrement vrai pour la schizophrénie (Lee, 2013).

Lorsqu'on tente de caractériser le fonctionnement cérébral chez un patient, il est donc difficile de dire si les particularités observées sont le résultat du processus pathologique sous-tendant la pathologie principale, des symptômes surajoutés durant la période de recrudescence, des pathologies comorbides, ou des traitements reçus.

L'étude des effets des traitement psychotropes se heurte elle aussi à cette hétérogénéité. Deux types d'études adoptent des stratégies différentes pour prendre en compte ce problème, les études naturalistiques et les essais thérapeutiques randomisés contrôlés.

Les études naturalistiques sont des études pragmatiques où il n'y a pas d'intervention randomisée, où les variables dépendantes sont souvent proches de la pratique clinique (nombre de ré-hospitalisation, temps avant rechute, ...). Les patients recrutés sont souvent peu sélectionnés et correspondent assez bien au profil des patients pris en charge dans la pratique clinique courante, y compris dans leur hétérogénéité. Ces études acceptent donc les différents types d'hétérogénéité et étudient l'effet des traitement en moyenne sur cette population représentative.

Les essais cliniques randomisés contrôlés (RCT) sont les études requises par les autorités de régulation autorisant la mise sur le marché des médicaments, elles sont considérées comme le gold standard. Elles tentent de s'affranchir d'une partie de cette hétérogénéité. Dans ces essais, la population est strictement sélectionnée avec une multiplication des critères d'exclusion (Correll et al., 2011). Cette stratégie peut permettre de mettre en évidence de façon plus univoque des effets significatifs des traitements, mais compromet la généralisabilité des résultats, la population étudiée étant parfois sensiblement différente de la population qu'aura à traiter un praticien (Ioannidis, 2008).

A l'extrême, les conclusions sur l'efficacité des traitements peuvent être différentes suivant que l'étude est un essai randomisé contrôlé ou une étude naturalistique. C'est le cas par exemple pour l'étude des traitements antipsychotiques retard par rapport aux antipsychotiques par voie orale. Les études randomisées contrôlées ne retrouvent pas de différence significative entre ces deux traitements, alors que les études naturalistiques retrouvent un bénéfice du traitement retard (Kirson et al., 2013; Kishimoto et al., 2013, 2014; Lafeuille et al., 2014). Une explication possible est que les RCT fonctionnant avec des évaluations régulières, des patients inclus

seront plus suivis que s'ils étaient en milieu naturel, ce qui aura tendance à réduire la différence entre traitement oral et traitement retard (Fagiolini et al., 2016), différence qui en l'occurrence est notamment liée à la question de l'observance (ou plutôt du défaut d'observance).

Au total, les essais cliniques présentent des limites et la méthodologie actuelle de la recherche ne peut garantir que les meilleurs patients sont inclus, ou que l'effet réel des traitements est effectivement mesurée (McEvoy et Freudenreich, 2008).

### 3.1.2 Caractère neurodéveloppemental de certaines pathologies psychiatriques

De nombreux éléments convergents indiquent que le phénomène pathologique à l'œuvre dans la schizophrénie par exemple débute bien avant l'apparition d'un tableau pathologique manifeste (Meyer-Lindenberg, 2010).

Les anomalies trouvées chez les patients pourraient refléter un état trop avancé pour qu'une compréhension des mécanismes initialement à l'œuvre soit possible. Dès lors, les progrès vers le développement d'un traitement précoce et préemptif pourraient ne pas découler naturellement de l'étude des patients eux-mêmes (Meyer-Lindenberg, 2010).

Les anomalies similaires trouvées chez deux patients porteurs du même diagnostic pourraient également refléter une voie commune finale, après l'évolution sur plusieurs années de deux processus pathologiques pourtant initialement différents chez ces deux patients (Meyer-Lindenberg, 2010).

L'étude des mécanismes sous-tendant les pathologies est importante pour le développement de nouveaux médicaments car elle permet de définir une cible thérapeutique. En psychiatrie le caractère neurodéveloppemental des troubles fait que le tableau clinique peut être peu représentatif des mécanismes initiaux qui sous-tendent la maladie.

## 3.2 Limites liées à la classification des pathologies psychiatriques

L'introduction d'une classification des pathologies mentales basée sur des critères diagnostiques athéoriques avec le DSM3 en 1980 a été une avancée importante. Elle proposait d'asseoir la classification diagnostique sur des éléments observables ce qui a eu deux conséquences majeures. Pour la clinique d'abord, elle a amélioré la reproductibilité des diagnostics, qui reposaient jusque là sur des cadres théoriques variés, psychanalytique, systémique, classique. Pour la recherche ensuite, elle a

permis que des groupes plus comparables d'un site de recrutement à l'autre soient constitués (McHugh, 2005).

Des critiques sont cependant apparues progressivement et ce système de classification a été remis en question sur la base de plusieurs lignes d'arguments.

### 3.2.1 Nosographie et hétérogénéité

Une classification basée sur des symptômes apparents groupées en syndromes fait que plusieurs patients très dissimilaires peuvent satisfaire les critères d'un même trouble (McHugh, 2005). Prenons l'exemple de l'épisode dépressif caractérisé. Un patient reçoit ce diagnostic lorsqu'il présente cinq des neuf symptômes listés dans le manuel. En conséquence, deux patients peuvent recevoir le même diagnostic en ayant 1 symptôme commun et 4 symptômes différents. Les patients regroupés au sein d'une même entité diagnostic peuvent donc être très hétérogènes du fait de ce mode de classification nosographique catégoriel par regroupement de symptômes (Goldberg, 2011).

De plus, certains syndromes psychiatriques sont non spécifiques et peuvent être présents dans les critères diagnostiques de plusieurs troubles psychiatriques différents (Owen et al., 2011). Les critères diagnostiques de l'épisode dépressif caractérisé dans le DSM-V incluent par exemple "Insomnie ou hypersomnie" (critère 4) ou "Difficulté à penser ou à se concentrer" (critère 8). Ils font également partie des critères du trouble anxieux généralisé avec "Perturbation du sommeil" (critère 6) et "Difficulté à se concentrer ou pensée vide" (critère 3) (American Psychiatric Association, 2013).

Ce recouvrement de symptômes est décrit et vérifié en pratique. Sur une population de patients répondant aux critères diagnostiques de trouble anxieux généralisé, 70.4% satisfaisaient également les critères diagnostiques de trouble dépressif caractérisé (Faravelli et al., 2012). Seulement deux des symptômes diagnostiques du trouble anxieux généralisé (préoccupation excessive et tension musculaire) étaient spécifiquement associés au diagnostic, les autres symptômes comme la tension, les problèmes de concentration ou la fatigue, étaient plus prévalents dans le trouble dépressif que dans le trouble anxieux généralisé (Faravelli et al., 2012).

Ce recouvrement de symptômes est également décrit de façon formelle. Une analyse de réseaux, illustrée Figure 3.1 permet de conclure qu'il existe effectivement des recouvrements significatifs de symptômes entre catégories de diagnostic (Borsboom et Cramer, 2013).

Au final, en dépit des objectifs de reproductibilité du diagnostic ayant présidé à l'élaboration du DSM, la concordance inter-juges pour les principaux diagnostics psychiatriques est faible, comme illustré Figure 3.2. La concordance pour les principaux diagnostics psychiatriques, mesurée par les coefficients kappa, ressortent à 0.46 pour la schizophrénie, 0.28 pour l'épisode dépressif caractérisé ou 0.20 pour

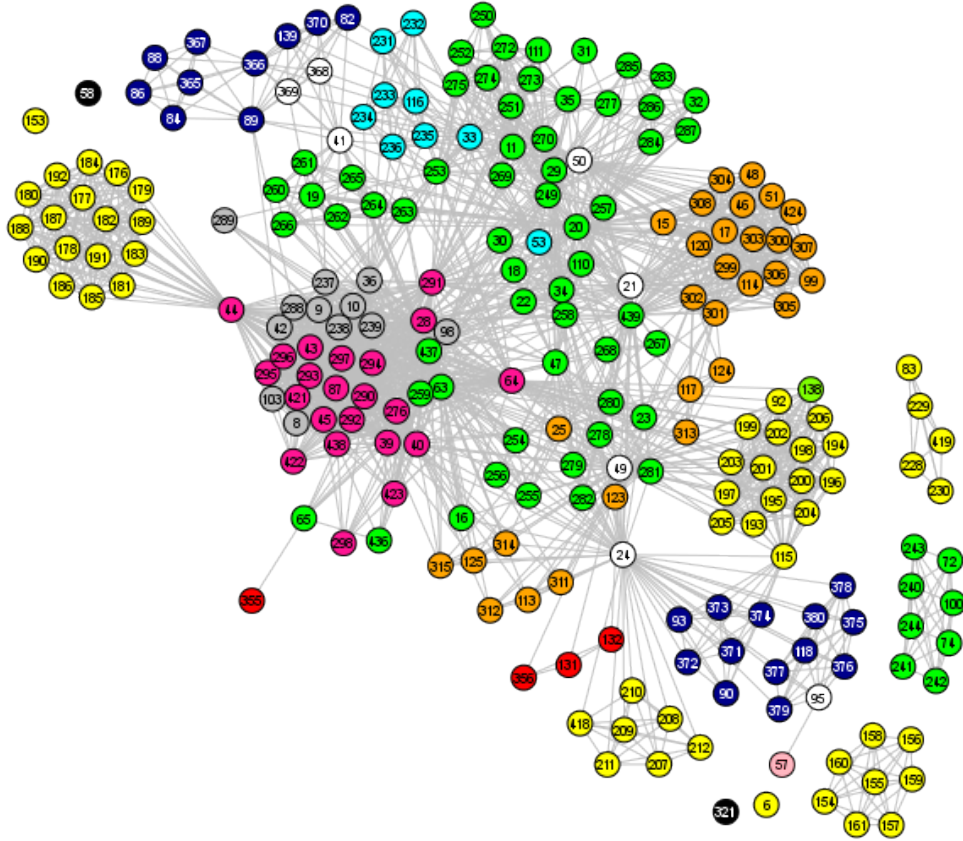


FIGURE 3.1 – Recouvrement de symptômes entre différentes catégories de diagnostic en psychiatrie. Une analyse de réseau extraite de Borsboom et Cramer (2013). Chaque point représente un symptôme. Un lien entre deux points signifie que les deux symptômes figurent dans le même trouble selon les critères du DSM-IV. Les couleurs représentent les chapitres du DSM-IV dans lequel ils apparaissent le plus souvent. Noter en particulier le recouvrement important entre les troubles de l'humeur (Rose) et les troubles du spectre schizophrénique (Gris). Légende : Jaune = Troubles habituellement diagnostiqués pendant la première enfance, la deuxième enfance ou l'adolescence, Bleu clair = Délirium, démence, trouble amnésique et autres troubles cognitifs, Vert = Troubles liés à une substance, Gris = Schizophrénie et Autres Troubles psychotiques, Rose = Troubles de l'humeur, Orange = Troubles anxieux, Violet = Troubles de la personnalité.

le trouble anxieux généralisé (Freedman et al., 2013), alors qu'un coefficient kappa est considéré comme acceptable habituellement à partir de 0.60 (Cohen, 1960).

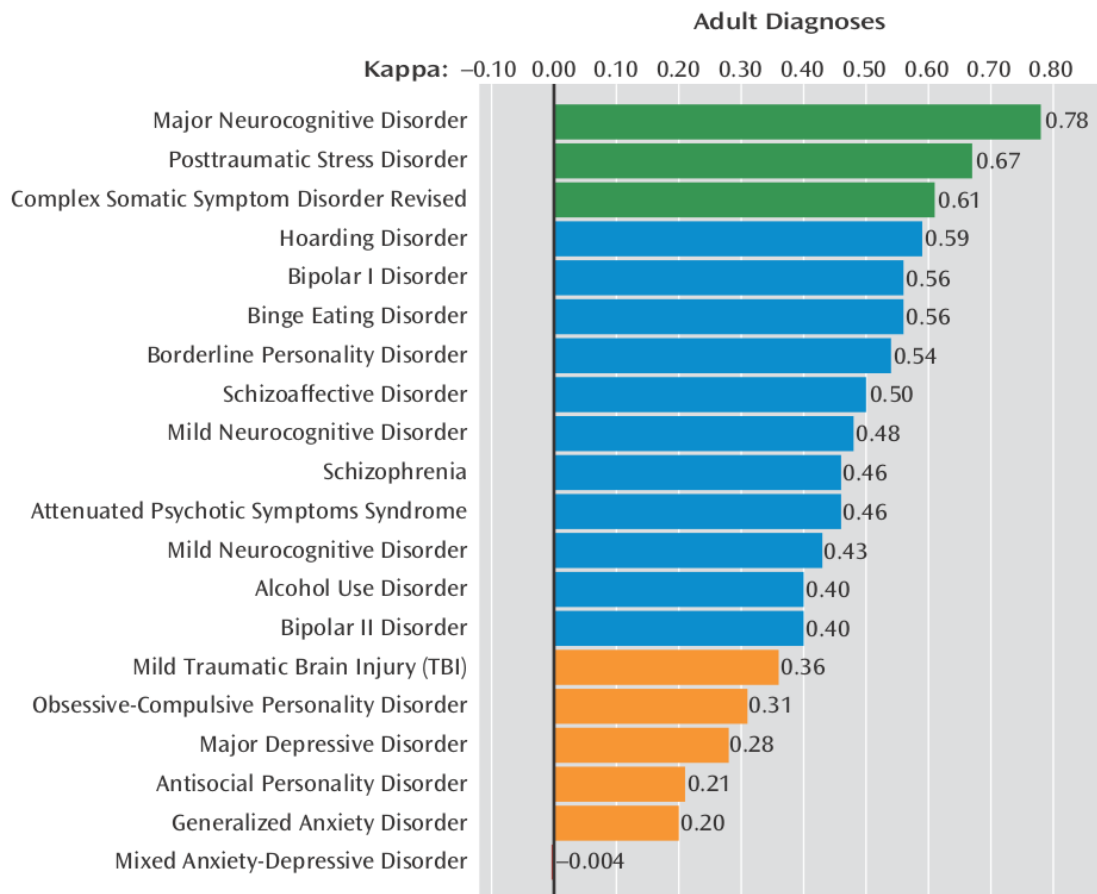


FIGURE 3.2 – Concordance inter-juges des principaux diagnostics psychiatriques. Cette concordance est issue d'études de terrain préparatoires à la rédaction du DSM-V. Kappa = coefficient kappa de Cohen. (Freedman et al., 2013)

Ces constats suggèrent que des mécanismes plus fondamentaux sont à l'œuvre, qui sous-tendent les troubles psychiatriques sans se conformer au découpage nosographique actuel (Cuthbert et Insel, 2010).

### 3.2.2 Nosographie et neurodéveloppement

La conceptualisation de certaines pathologies psychiatriques comme un trouble neurodéveloppemental (Meyer-Lindenberg, 2010) a amené à repenser les pathologies psychiatriques comme résultant du développement aberrant de circuits cérébraux spécifiques (Insel et Landis, 2013).

Un nouveau modèle de la dépression a par exemple émergé qui propose une étiologie en lien avec un dysfonctionnement spécifique de circuits cérébraux traitant les

émotions, la cognition, l'intéroception ou encore la conscience de soi (Drevets et al., 1997 ; Ressler et Mayberg, 2007). Ces nouvelles propositions suggèrent non seulement une nouvelle façon de stratifier la dépression, mais également la possibilité de nouvelles cibles thérapeutiques pour des médicaments comme la scopolamine, un agent anticholinergique de type muscarinique (Furey et Drevets, 2006), la kétamine, un antagoniste des récepteurs NMDA (Zarate et al., 2006), ou des techniques de stimulation profonde (Holtzheimer et Mayberg, 2011).

Cette nouvelle conceptualisation des troubles psychiatriques a également participé à la critique des classifications nosographiques utilisées (Cuthbert et Insel, 2010).

#### 3.2.3 Nosographie et génétique

La recherche en génétique fondait beaucoup d'espoir sur la capacité du système de classification reposant sur une critériologique objective à fournir des groupes de patients présentant un ensemble similaire d'anomalies génétiques qu'il aurait été aisé de caractériser.

L'héritabilité de pathologies psychiatriques est élevée, de l'ordre de 81% pour la schizophrénie, 75% pour les troubles bipolaires et 37% pour le trouble dépressif caractérisé (Sullivan et al., 2012). Concernant la schizophrénie, le sur-risque pour un individu ayant un parent malade croît de 2% pour un apparenté au 3ème degré à 9% pour un apparenté au premier degré, à 27% dans le cas d'un enfant dont les deux parents sont malades (Gottesman et al., 2010). Les études chez les jumeaux monozygotes montrent une concordance d'environ 50% (Lichtenstein et al., 2009) et pour les enfants adoptés, dont un des parents biologique est schizophrène, un risque 6 à 10 fois plus important que dans la population générale de développer une pathologie schizophrénique (Shih et al., 2004).

Cependant, malgré cette héritabilité bien établie, les recherches initialement optimistes espérant trouver une base génétique mendélienne aux catégories de diagnostic ont été un échec (Gottesman et Gould, 2003 ; Kendler, 2015). Les premières études évaluant des gènes candidats individuels n'ont pas permis de mieux comprendre les bases génétiques des pathologies (pour une revue de l'évaluation de 25 gènes candidats historiques pour la schizophrénie, voir Farrell et al. (2015)). Les études plus récentes, où le génome complet est couvert, permet de faire des études d'association pangénomiques. Mais les résultats sont là encore en dessous des espérances. Ces études d'associations retrouvent un grand nombre de gènes impliqués, chacun a un pouvoir explicatif faible. Pour la schizophrénie, plus de 100 loci différents ont ainsi été identifiés, correspondant à un total de 341 gènes codant des protéines (Lencz et Malhotra, 2015). A l'issue de ces études, une large partie de la variance héritable reste inexpliquée (Gershon et al., 2011).

Il est maintenant clair que la susceptibilité génétique à des pathologies psychiatriques comme la schizophrénie est complexe, hautement polygénique (Consortium et al.,

2009), avec une architecture consistant en un grand nombre de variants communs de faible pénétrance, et de quelques variants rares de pénétrance modérée (Giegling et al., 2017). De plus, un recouvrement important des facteurs de risque génétiques est retrouvé pour des catégories de diagnostic différentes comme la schizophrénie, l'autisme et le déficit intellectuel (Giegling et al., 2017 ; McCarthy et al., 2014).

Ces résultats décevants fournissaient une nouvelle ligne d'arguments pour remettre en cause le bien fondé de l'utilisation d'une nosographie psychiatrique fondée sur des regroupements syndromiques de symptômes.

Une critique fondamentale du DSM est donc que ses critères diagnostiques, le regroupement syndromique de symptômes de surface, ne rendent pas compte et ne suivent pas les contours des mécanismes biologiques à l'œuvre dans les pathologies psychiatriques (Robbins et al., 2012). Ceci a un effet néfaste sur l'efficacité de la recherche des corrélats génétiques des troubles psychiatriques. Ceci a également un effet néfaste sur la capacité d'essais cliniques chez l'homme à détecter un effet thérapeutique. Un nouveau médicament qui aurait une action sur un mécanisme physiopathologique particulier pourrait ainsi ne pas démontrer son efficacité s'il est testé sur un groupe de sujets porteurs du même diagnostic catégoriel mais dont un petit sous groupe seulement présente cette physiopathologie précise (Nutt et Need, 2014). Si la psychiatrie souhaite traiter les mécanismes sous-jacents en amont de l'apparition de symptômes dans une voie finale commune tardive (Howes et Kapur, 2009), ce problème sera de plus en plus présent. Dès lors il est souhaitable de pouvoir constituer des groupes de patients présentant une homogénéité des mécanismes physiopathologiques sous-jacents. Les endophénotypes sont une des solutions proposées dans ce but.

## **3.3 Les endophénotypes et les biomarqueurs**

### **3.3.1 Les endophénotypes**

A début du 20ème siècle, la distinction entre génotype et phénotype a été introduit par le botaniste Danois Wilhelm Johanssen (Gottesman et Gould, 2003). Il avait découvert que le phénotype est souvent un mauvais indicateur du génotype, le même génotype pouvant donner naissance à une grande variété de phénotypes, et le même phénotype pouvant apparaître à partir de génotypes différents (Gottesman et Gould, 2003). Les facteurs épigénétiques, l'interaction avec l'environnement (Petronis et al., 2000) ainsi que facteurs entièrement stochastiques (Rakyan et al., 2001) expliquent ce phénomène.

Les pathologies psychiatriques, qui impliquent un grand nombre de gènes, ont un phénotype déterminé par une interaction diachronique longue entre facteurs génétiques (illustrés Figure 3.3), facteurs environnementaux et facteurs épigénétiques (Lewis et Levitt, 2002).



La notion d'endophénotype n'est pas récente, puisque le terme est employé en 1973 dans la discussion sur le lien entre schizophrénie et génétique par Gottesman et Shields (1973). Il souligne :

“[the] useful distinction between exophenotype (external phenotype) and endophenotype (internal), with the latter only knowable after aid to the naked eye, e.g. a biochemical test or a microscopic examination of chromosome morphology.” (Gottesman et Shields, 1973).

L'idée est que si un phénotype décrit en association avec une pathologie est très spécifique et représente un mécanisme plus élémentaire, le nombre de gènes sous tendant le développement de ce phénotype est plus limité que ceux impliqués dans l'entité diagnostique psychiatrique globale (Gottesman et Gould, 2003).

Le concept s'est précisé et on considère aujourd'hui qu'un marqueur doit remplir un certain nombre de critères pour être considéré comme un endophénotype (Lenzenweger, 2013) :

1. l'endophénotype est associé avec la maladie dans la population ;
2. l'endophénotype est héritable (contrairement à un biomarqueur qui n'est pas nécessairement héritable) ;
3. l'endophénotype est idéalement indépendant de l'état, se manifestant chez l'individu que la maladie soit active ou non ;
4. l'endophénotype est plus prévalent chez les apparentés malades que chez les apparentés non malades des sujets malades ;
5. l'endophénotype présent chez un malade est trouvé de façon plus fréquente chez ses apparentés non malades que dans la population générale ;
6. l'endophénotype doit pouvoir être mesuré de façon fiable, et est idéalement plus associé avec la maladie étudiée qu'avec d'autres troubles psychiatriques.

### 3.3.2 Distinction entre endophénotypes et biomarqueurs

Un biomarqueur peut être n'importe quel indicateur mesurable d'une maladie. C'est une caractéristique mesurée objectivement, utilisée comme indicateur d'un processus biologique normal, d'un processus pathologique ou d'une réponse thérapeutique (Group, 2001).

Le biomarqueur est un concept plus large que celui d'endophénotype. Tous les endophénotypes sont des biomarqueurs, mais tous les biomarqueurs ne sont pas des endophénotypes (Lenzenweger, 2013). Les endophénotypes seraient un sous-groupe de biomarqueurs influencés par les facteurs génétiques qui sous-tendent la maladie étudiée (Glahn et al., 2014). Le biomarqueur n'est donc pas nécessairement héritable alors que c'est une caractéristique constitutive de l'endophénotype.

Les biomarqueurs se sont beaucoup développés dans des champs de la médecine comme la cancérologie. La surexpression du gène HER2 a par exemple été identifiée

initialement comme un facteur de mauvais pronostic dans le cancer du sein (Slamon et al., 1987). Cela a permis une meilleure compréhension de ce type particulier de cancer, et a guidé le développement de thérapies monoclonales avec le trastuzumab (Smith et al., 2007). Cette approche a permis l'émergence d'une médecine, qui, si elle n'est pas encore réellement personnalisée (Langreth et Waldholz, 1999), permet du moins de préciser de façon fine pour un petit sous-groupe de patients le diagnostic, le pronostic et les meilleures approches thérapeutiques.

De nombreuses applications sont possibles en psychiatrie comme illustré Figure 3.4. Elles incluent la confirmation diagnostique avec recherche de diagnostic différentiel, par exemple entre des troubles cognitifs dans le cadre d'un trouble dépressif et des éléments négatifs dans le cadre d'un trouble schizophrénique. Elles incluent une caractérisation pronostique par exemple de l'évolution d'un état mental à risque vers la schizophrénie ou non. Elles incluent une caractérisation pronostique de la réponse à un traitement. Actuellement le clinicien doit se contenter d'essayer un des traitements disponibles à l'aveugle ou en fonction de critères secondaires comme les effets indésirables, puis de changer d'option thérapeutique en l'absence de réponse à ce premier traitement. L'efficacité d'un traitement ne s'appréciant bien souvent en psychiatrie qu'avec un recul de plusieurs semaines, cette absence de marqueur de réponse rend la prise en charge des maladies psychiatriques longue et coûteuse.

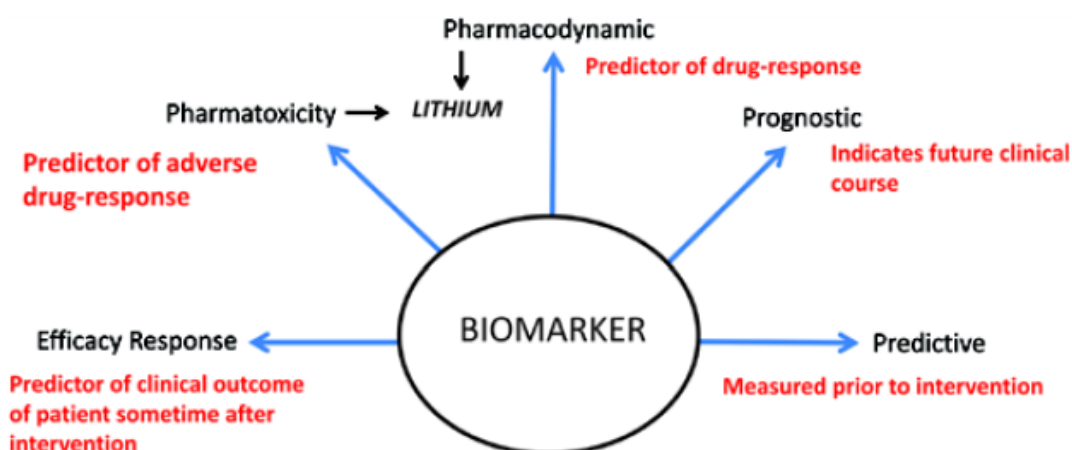


FIGURE 3.4 – Utilisations potentielles des biomarqueurs en psychiatrie. (Scarr et al., 2015)

#### 3.3.3 Le bénéfice attendu des endophénotypes

L'intérêt de l'approche par endophénotype est d'autant plus important que la pathologie étudiée a un substrat génétique particulièrement complexe (Figure 3.5). Plus le nombre de gènes impliqués dans la physiopathologie d'un trouble est important, plus grande sera la complexité phénotypique du tableau, et plus ardues seront les analyses génétiques (Gottesman et Gould, 2003).

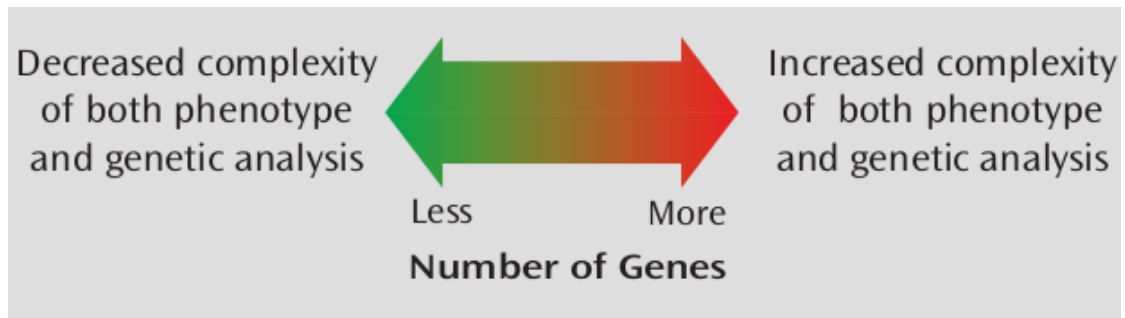


FIGURE 3.5 – Justification de l'intérêt d'une approche par endophénotype pour l'analyse génétique de pathologies ayant un substrat génétique complexe. Le nombre de gènes impliqués dans un phénotype particulier est considéré comme directement lié à la fois à la complexité du phénotype et à la difficulté des analyses génétiques. (Gottesman et Gould, 2003)

Les bénéfices des endophénotypes pourraient être multiples. Ils sont résumés par Lenzenweger en quatre points (Lenzenweger, 2013) :

1. Les endophénotypes pourraient être un reflet plus direct de l'activité cérébrale, synaptique, biochimique que les diagnostics catégoriels de maladies eux-mêmes, et il est possible qu'ils puissent être liés à la génétique avec des tailles d'effet plus importants.
2. A la fois les patients et leurs apparentés sains pourraient obtenir des scores variés dans ces mesures d'endophénotype, ce qui pourrait en faire des mesures idéales pour l'analyse de liaison de locus à caractère quantitatif ("Quantitative trait linkage analysis", qui vise à identifier les loci dont la variation allélique est associée à la variation d'un caractère d'intérêt). Les données quantitatives et potentiellement continues caractérisant l'endophénotype permettent une puissance statistique supérieure pour détecter ces associations, en comparaison de données cliniques qualitatives ou binaires issues des descriptions sémiologiques psychiatriques.
3. Si la physiologie de l'endophénotype est effectivement comprise, ou caractérisée par des examens d'imagerie ou des mesures biologiques chez l'humain ou l'animal, les gènes candidats pourront plus facilement être identifiés.
4. Les endophénotypes pourraient être utilisés dans les modèles animaux.

Les endophénotypes constituent une réponse aux difficultés rencontrées par les études sur les déterminants génétiques des pathologies psychiatriques. Ils pourraient également permettre d'améliorer la valeur prédictive dans l'évaluation de nouveaux médicaments (Koychev et al., 2012).

La probabilité qu'une molécule ayant franchi les étapes précliniques et entrant en phase I des essais cliniques devienne un jour un médicament commercialisé est de 9.6% en moyenne toutes spécialités confondues. Mais cette probabilité est différente suivant qu'un biomarqueur est utilisé dans les études ou non. Le taux de succès

est ainsi de 25.9% lorsqu'un biomarqueur est disponible contre 8.4% en l'absence de biomarqueur comme illustré Figure 3.6. Cette différence est majeure. L'analyse phase par phase montre que c'est dans les phases II et III de l'expérimentation clinique que la différence de succès suivant l'existence ou non d'un biomarqueur est la plus marquée. On passe ainsi de 28% à 46% pour le succès en phase II et de 55% à 76% pour le succès en phase III (Hay et al., 2014; Thomas et al., 2016).

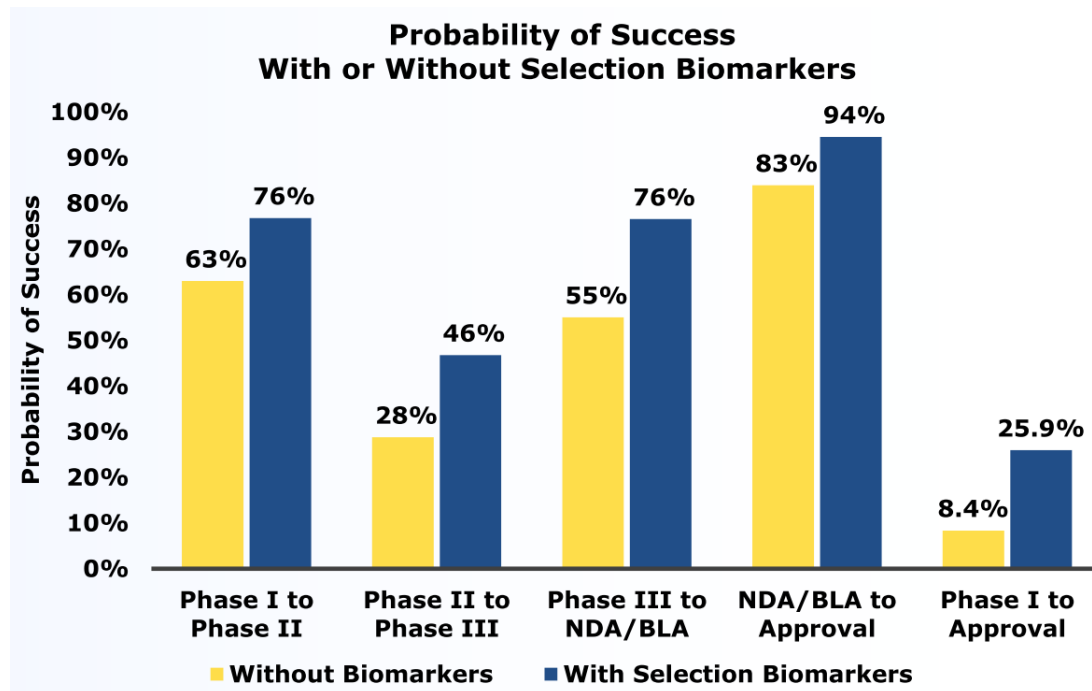


FIGURE 3.6 – Probabilité de succès aux différentes étapes des essais cliniques selon qu'un biomarqueur est utilisé ou non. NDA/BLA = New Drug Application / Biologic Licence Application. Approval = équivalent de l'Autorisation de Mise sur le Marché (Hay et al., 2014; Thomas et al., 2016)

#### 3.3.4 Les Research Domain Criteria (RDoC)

Avec les endophénotypes, deux possibilités semblent s'offrir à nous :

- enrichir la description actuelle des pathologies par des dimensions supplémentaires, dont les biomarqueurs, afin de stratifier plus finement les patients ;
- remplacer purement et simplement les classifications actuelles par une description à un niveau plus proche du fonctionnement cérébral.

Certains pensent que la première solution n'est pas assez radicale et que les classifications actuelles resteront un frein, même si elles sont affinées par des biomarqueurs. Ils proposent de se focaliser sur des symptômes spécifiques, ainsi que

sur des endophénotypes cognitifs et neurocognitifs, en soustrayant l'effet confondant des diagnostics catégoriels (Owen et al., 2011). Pour eux, les descriptions futures des troubles psychiatriques doivent adopter une approche dimensionnelle pour améliorer les classifications et l'innovation thérapeutique (Fernando et Robbins, 2011).

Le projet des "Research Domain Criteria" (RDoC) est une initiative du National Institute of Mental Health pour formaliser et promouvoir cette direction. Il encourage la recherche amenant à développer de nouvelles classifications des pathologies mentales sur la base de dimensions comportementales observables et de mesures neurobiologiques (Cuthbert et Insel, 2010 ; Insel et al., 2010). C'est un projet du domaine de la recherche pour le moment. L'objectif à long terme est cependant de remplacer l'approche psychiatrique catégorielle actuelle par une caractérisation dimensionnelle des troubles grâce à des indicateurs plus proches des mécanismes physiopathologiques à l'oeuvre. Ces indicateurs pourraient appartenir à différents niveaux d'analyse, génétiques, comportementaux, issus des neurosciences... (Kapur et al., 2012).

### 3.3.5 Variété des endophénotypes et biomarqueurs

Peu de biomarqueurs sont disponibles actuellement pour la psychiatrie en pratique clinique.

La commercialisation d'un test sanguin pour confirmer un diagnostic de schizophrénie débutante avait été tentée en 2010. Ce test sanguin se basait sur l'analyse de 51 données biologiques étudiées sur 250 patients présentant une schizophrénie et 230 sujets sains. Testé sur une cohorte indépendante de 577 patients avec schizophrénie et 229 sujets sains, ce test obtenait une sensibilité de 83% et une spécificité de 83% (Schwarz et al., 2010). Cependant, après une commercialisation initiale, le test a été finalement retiré du marché en 2013, en partie du fait du coût du test, en partie du fait d'une sensibilité et spécificité perfectibles (Scarr et al., 2015).

En revanche de nombreux biomarqueurs sont en phase d'étude ou de validation.

Des marqueurs d'imagerie ou d'électrophysiologie ont été testés, par exemple pour prédire la transition psychotique dans deux études rétrospectives en IRM fonctionnelle (Koutsouleris, Borgwardt, et al., 2012 ; Koutsouleris et al., 2009), ou une étude en spectro-imagerie s'intéressant à la libération de glutamate et glutamine en réponse à une perfusion de kétamine (Poels et al., 2014).

Des marqueurs sanguins sont à l'étude comme le dosage de la glycine. Un taux de plus de 300 microMol/l semble prédire une bonne réponse à l'addition de D-cyclosérine chez des patients résistants à une première ligne de traitement antidépresseur (Heresco-Levy et al., 2013).

De nombreux marqueurs génétiques sont à l'étude avec les difficultés décrites plus

haut, ainsi que des marqueurs issus de la transcriptomique et la protéomique (Glahn et al., 2014).

Des marqueurs peuvent également être de nature neurocognitive. Plusieurs études tentent par exemple de distinguer des sujets sains de sujets à risque de transition psychotique avec des marqueurs neurocognitifs avec un certain succès (Fusar-Poli et al., 2012 ; Koutsouleris, Davatzikos, et al., 2012). Ces marqueurs neurocognitifs ont pour certains une voie de choix :

“The most effective ‘stratifiers’ in psychiatry may well come from standardized cognitive and psychological measures”. (Kapur et al., 2012)

Examinons comment les neurosciences cognitives peuvent apporter leur contribution à l'élaboration d'endophénotypes.

## 3.4 L'apport des neurosciences cognitives

### 3.4.1 Définition et application à la psychiatrie

Les neurosciences cognitives se donnent pour tâche de comprendre l'architecture de la cognition et ses bases neurales, c'est donc une discipline qui est distincte mais recouvre en partie d'autres disciplines comme la psychologie physiologique, les neurosciences, la psychologie cognitive et la neuropsychologie (Michael Gazzaniga, 2014). Elles examinent les liens entre cognition et cerveau en utilisant une grande variété d'outils comme l'imagerie, l'électrophysiologie, l'étude des lésions cérébrales, chez l'homme comme chez l'animal (Michael Gazzaniga, 2014).

L'application des neurosciences cognitives à l'étude des pathologies psychiatriques est un champ très actif. Un terme spécifique a d'ailleurs été créé, c'est la neuropsychiatrie cognitive. Ce terme délimite une approche systématique et guidée par la théorie pour expliquer la psychopathologie clinique en termes de perturbations des mécanismes cognitifs normaux (Halligan et David, 2001).

Cette approche souhaite établir un pont entre la description clinique des pathologies psychiatriques et les dysfonctions cérébrales précises sous-tendant ces troubles sur la base de l'analyse du fonctionnement cognitif. Les troubles neuropsychiatriques y sont décrits en termes de perturbation (en défaut ou en excès) dans les systèmes de traitement de l'information (Halligan et David, 2001). Les bases cérébrales de ces perturbations sont ensuite recherchées. L'accent est mis sur les symptômes et non sur les diagnostics. Cette approche propose en cela la recherche des biomarqueurs de nature neurocognitive des pathologies psychiatriques.

### 3.4.2 Biomarqueurs neuro-cognitifs de pathologies psychiatriques

La recherche et la caractérisation de biomarqueurs cognitifs des pathologies psychiatriques fait l'objet d'une large littérature. Les pathologies du spectre schizophrénique y font l'objet d'une attention importante. Nous ne chercherons pas à faire une revue des différentes théorisations et des multiples candidats biomarqueurs cognitifs qui ont été proposés. Différentes revues de la littérature sont disponibles à ce sujet (voir par exemple Connors et Halligan (2015), Garety et Freeman (2013), Pickard (2015) ou Hallak et al. (2015) pour des revues récentes). Nous prendrons à titre illustratif quelques exemples de telles approches pour des symptômes spécifiques, comme les hallucinations ou le délire.

Des exemples de modèles cognitifs du délire incluent le raisonnement inférentiel, ou style de raisonnement en “jumping to conclusions” qui a fait l'objet d'un grand nombre de publications (Broome et al., 2007 ; Garety et Freeman, 2013, 1999 ; Hemsley, 1993), une incapacité à reconnaître des actions ou des pensées comme étant auto-générées (Frith, 1992), un déficit dans l'évaluation des croyances (Coltheart, 2007 ; Coltheart et al., 2011 ; Vinckier et al., 2016), une perturbation dans la mise à jour des inférences du fait d'erreurs de prédictions inappropriées (Corlett et al., 2010 ; Fletcher et Frith, 2009), ou une interaction entre délire et mémoire (Berna et al., 2014).

Des exemples de modèles cognitifs des hallucinations incluent un déséquilibre entre processus bottom-up et top-down (Aleman et Vercammen, 2012), des anomalies dans le traitement de la source du langage avec une incapacité à reconnaître le discours intérieur comme tel (Dollfus et al., 2012), un dysfonctionnement de la mémoire, les hallucinations auditives étant alors vues comme des fragments de mémoire rappelés de façon non intentionnelle (Rossell, 2012), ou encore une dysfonction des capacités de discrimination temporelle qui contribueraient à des distorsions perceptives sous-tendant les hallucinations (Waters, 2012).

Les modèles explicatifs sont donc nombreux dans ce champ. Cependant, tous n'ont pas la même capacité à produire des hypothèses expérimentales qui pourront être testées à la fois chez l'homme et chez l'animal. L'explication du délire par un déséquilibre entre les réseaux cérébraux du self intrinsèque et extrinsèque (Ebisch et Aleman, 2016) sera par exemple difficilement testable chez la souris. Il en va de même pour l'explication du délire par une mésattribution excessive à autrui d'événements négatifs dans le but de protéger son estime de soi (Bentall et al., 2001, 1994).

A l'inverse, certaines hypothèses explicatives peuvent être testées plus facilement par des dispositifs expérimentaux applicables chez l'homme comme chez l'animal. L'explication du délire comme résultat de perturbation dans la mise à jour des inférences du fait d'erreurs de prédiction inappropriées (Corlett et al., 2010 ; Fletcher et Frith, 2009) en fait partie. Cette hypothèse explicative repose en effet sur un

cadre descriptif des comportements qui a été utilisé chez l'homme comme chez l'animal, l'apprentissage par renforcement. Nous exposerons ce cadre au chapitre 4.

## 3.5 Couplage avec la psychopharmacologie

La psychopharmacologie est l'étude des effets des substances psychoactives sur le fonctionnement du système nerveux central à tous les niveaux d'analyse. Elle s'intéresse donc à la cognition, au comportement, aux états psychologiques, à la neurophysiologie, à la neurochimie, à l'expression génique et à la biologie moléculaire (Stolerman et Price, 2015). Elle étudie l'utilisation sociétale et médicale de ces substances, l'interaction des facteurs environnementaux et génétiques avec l'action des substances psychoactives, leur utilisation pour tester le fonctionnement du système nerveux central, ainsi que l'abus et la dépendance aux drogues (Stolerman et Price, 2015).

En quoi coupler psychopharmacologie et neurosciences cognitives pourrait-il être un apport pour la psychiatrie ?

### 3.5.1 Étude de l'effet de traitements

Malgré de nombreuses années de recherche le mode d'action des médicaments psychotropes, notamment antidépresseurs ou antipsychotiques, est encore largement incompris. Poursuivre les efforts pour en caractériser les effets n'est donc pas vain (Millan, 2008).

On sait que différents représentants de la classe des antipsychotiques ont des profils pharmacologiques très différents. La risperidone a une activité essentiellement anti D2 et anti 5HT2A, avec très peu d'activité sur les autres types de récepteurs dopaminergiques ou sur les autres récepteurs sérotoninergiques de type 2. L'olanzapine possède également cette activité anti D2 et anti 5HT2A, mais possède également une certaine affinité pour les récepteurs D1, D3, D4, ainsi que pour les récepteurs 5HT2B et 5HT2C. L'aripiprazole a quant à elle une activité agoniste partielle D2, une activité antagoniste 5HT2A, mais également une activité agoniste partielle 5HT1A, et une affinité pour plusieurs autres récepteurs dont les 5HT1D, 5HT2B, et D3 (Stahl, 2013).

Mais ces connaissances fondamentales ont eu un impact limité sur la pratique clinique. On reste loin d'une médecine personnalisée qui identifierait les dysfonctionnements cérébraux à l'œuvre chez le patient particulier à traiter, connaîtrait les effets relatifs des différents médicaments psychotropes sur ces circuits, et choisirait sur ces critères le traitement le plus adapté. Les cliniciens continuent de choisir la molécule à prescrire en fonction d'éléments essentiellement cliniques. L'intensité du délire fera préférer une molécule plus "incisive" comme la rispéridone, voire

l’halopéridol. Le risque de surpoids fera préférer l’aripiprazole à l’olanzapine. Dans certains cas, la connaissance fondamentale vient confirmer une pratique. Ainsi, la quasi absence d’activité anticholinergique de l’haloperidol est venu confirmer la pratique de son utilisation chez la personne âgée ou quand le risque de confusion mentale est élevé (Stahl, 2013). D’autres cliniciens s’en remettent à la médecine basée sur les preuves pour choisir le traitement le plus efficace globalement, ou celui ayant obtenu le meilleur taux d’observance à moyen terme (Cipriani et al., 2009 ; Lieberman et al., 2005).

Des progrès sont donc nécessaires en psychopharmacologie pour caractériser l’effet des traitements psychotropes au plan de leur effet sur le fonctionnement des circuits cérébraux et non seulement au plan de leur affinité avec tel ou tel récepteur. Les neurosciences cognitives se proposent de déterminer les substrats cérébraux du fonctionnement normal et pathologique.

Un couplage de ces deux approches pour caractériser l’effet des substances psychotropes sur les circuits cérébraux chez le sujet sain et en population clinique est donc indispensable. Ceci nécessite évidemment un choix judicieux dans le modèle cognitif choisi pour que ces deux approches se potentialisent. Nous défendons l’idée que l’étude de l’apprentissage par renforcement présenté au chapitre 4 est un choix judicieux. Caractériser les effets de traitements antidépresseurs chez le sujet sain est l’approche retenue pour notre premier article “Effects of acute and chronic antidepressant treatment on reinforcement learning and cognition : a 8 week randomized double blind placebo controlled study in healthy volunteers”. Caractériser l’effet d’un traitement antipsychotique dans une population clinique est l’approche retenue pour notre second article “Specific effect of a dopamine partial agonist on counterfactual learning : evidence from Gilles de la Tourette syndrome”.

### 3.5.2 Modélisation de pathologies psychiatriques

Une autre utilisation prometteuse de la psychopharmacologie pour la psychiatrie est la modélisation de pathologies psychiatriques chez le sujet sain.

Cette utilisation n’est pas nouvelle, et de nombreuses substances aux effets psychotomimétiques (qui miment par leur effet un ou des effets de la pathologie psychotique) ont été étudiées depuis plus de 50 ans (Carhart-Harris et al., 2013). Parmi ces substances psychotomimétiques, les plus étudiées ont été le cannabis ou son principe actif le THC (delta-9-tetrahydrocannabinol) (Kalant, 1971), les psychédéliques sérotoninergiques comme le LSD (Lysergic acid diethylamide) ou la psilocybine (Hoch et al., 1952), les amphétamines (Wallis et al., 1949) le PCP (phencyclidine) (Luby et al., 1959) ou la kétamine (Johnson, 1971).

Nous développerons ci-dessous l’exemple de la kétamine pour comprendre comment l’utilisation d’une substance psychoactive pour la modélisation d’une pathologie psychiatrique est possible et comment elle peut se coupler à une approche

neurocognitive.

La kétamine est utilisée cliniquement comme agent anesthésiant depuis 1970. Elle est fréquemment utilisée en chirurgie ambulatoire du fait de l'absence de dépression respiratoire ou cardiaque induite par son utilisation. Elle est également de plus en plus utilisée comme antalgique (Bell, 2009 ; Suzuki, 2009), comme traitement associé aux morphiniques d'une part (Carstensen et Møller, 2010), et comme potentiel traitement spécifique de l'hyperalgésie induite par les morphiniques par ailleurs (Gu et al., 2009). Son administration se fait classiquement en perfusion mais également désormais par voie orale (Blonk et al., 2010).

Plus récemment, les effets cognitifs de la kétamine à doses sub-anesthésiantes ont fait l'objet de nombreuses recherches qui ont permis de dire que la kétamine pouvait être un modèle pharmacologique utile de la schizophrénie (Corlett et al., 2006, 2007 ; Deakin et al., 2008 ; Krystal et al., 2005 ; Moore et al., 2013 ; Morgan et al., 2011 ; Schmidt et al., 2012). Plusieurs revues de la littérature sont disponibles à ce sujet (Javitt et al., 2012 ; Moghaddam et Krystal, 2012 ; Steeds et al., 2015).

La dose de kétamine utilisée en anesthésie est classiquement de 2 à 3 mg/kg en perfusion rapide initiale. La dose utilisée pour modéliser la pathologie psychotique se situe, elle, autour de 0.23 mg/kg en perfusion rapide initiale, puis perfusion continue lente de 0.009 mg/kg/min pendant trente minutes puis 0.004 mg/kg/min pendant une heure trente (D'Souza et al., 2012). Les doses utilisées pour l'étude de la kétamine comme modèle de psychose sont donc très inférieures (Figure 3.7) ce qui est un gage de sécurité d'utilisation.

A partir des études mentionnées ci-dessus, on note que :

- à des taux plasmatiques de 200 ng/mL, des phénomènes d'allure psychotique peuvent être observés : impression d'étrangeté, difficulté de coordination de la pensée, saillance anormale de certaines perceptions, particulièrement vives ;
- à des taux plasmatiques de 100 ng/mL, des modifications cérébrales, éventuellement associées à des modifications comportementales subtiles au cours d'une tâche cognitive ont été mises en évidence, mais sans franc symptôme d'allure psychotique, ou avec des modifications très subtiles.

L'étude de la sécurité d'utilisation de la kétamine à faible dose chez le sujet sain est rassurante. Une revue effectuée par Perry (Perry et al., 2007) portant sur 450 sujets ayant reçu des perfusions de kétamine ne retrouve pas d'effet indésirable grave. Des effets indésirables de nature psychiatriques ont été notés chez 2% des sujets avec des sensations désagréables rapportées ayant conduit certains sujets à demander l'interruption de la perfusion, ou l'absence de réponse aux questions de l'investigateur sans anomalie somatique associée ayant conduit l'investigateur à interrompre la perfusion. Ces réactions se sont en général amendées quelques minutes après l'arrêt de la perfusion, une personne a rapporté la persistance de cauchemars, insomnies et difficultés de concentration avec un retour complet à la

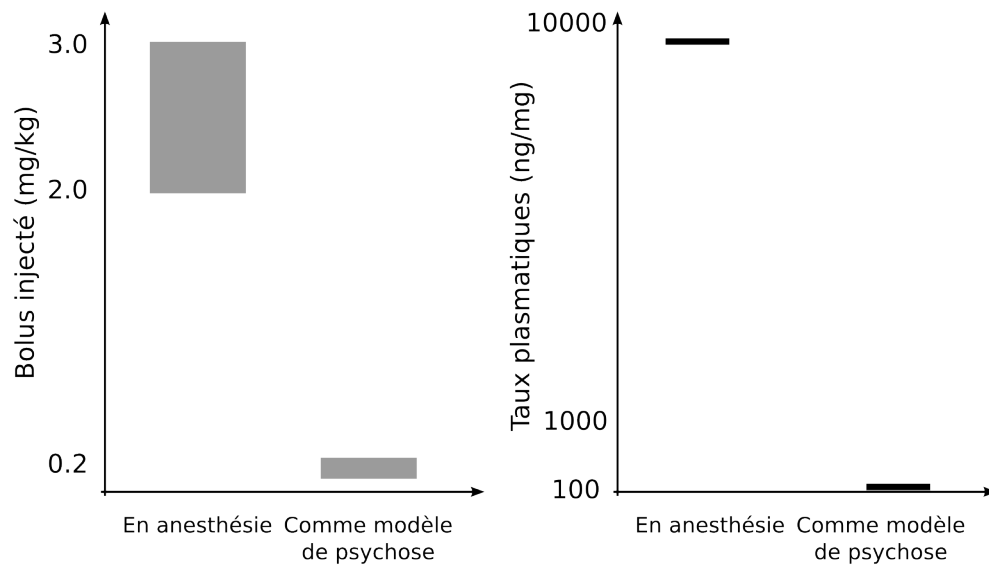


FIGURE 3.7 – Comparaison des doses de kétamine utilisées dans les études cognitives versus en anesthésiologie. A : Comparaison des doses injectées en perfusion rapide initiale en anesthésiologie versus dans les paradigmes de modèle pharmacologique de psychose. B : Comparaison des taux plasmatiques relevés en anesthésiologie (Induction entre 9000 et 25800 ng/ml, Réveil vers 1000 ng/ml) versus dans les paradigmes de modèle pharmacologique de psychose (100 à 200 ng/ml). Source (Domino et al., 1982)

normale après 4 jours (Perry et al., 2007). Les risques à distance d’une perfusion de kétamine unique évaluée sur 100 sujets une semaine après l’étude, 39 sujets après un mois, 50 sujets après trois mois et 34 sujets après six mois est également rassurant, avec une absence d’événement indésirable somatique ou psychiatrique et une absence de consommation récréative de kétamine. En revanche, certaines études indiquent de probables effets néfastes de l’utilisation chronique de la kétamine au plan cognitif (Morgan et al., 2012).

La kétamine à faible dose et en perfusion unique semble donc être un moyen efficace et sans danger connu d’induire des modifications cérébrales mimant des symptômes rencontrés dans la schizophrénie. Son utilisation chez le sujet sain peut être couplée avec l’approche par les neurosciences cognitives, pour tenter de caractériser les modifications des processus cognitifs et des circuits cérébraux touchés. Ce couplage a été fait par exemple avec une tâche d’apprentissage par renforcement chez le sujet sain sous kétamine, dans une expérience réalisée par l’équipe de recherche à laquelle j’étais rattaché au cours de ma thèse (Vinckier et al., 2016), ce qui m’a permis de me familiariser avec cette approche (Voir Article en Annexe 3). Ceci m’a conduit à poursuivre ces travaux en lançant une nouvelle étude visant à étudier les effets de la kétamine chez le sujet sain dans une tâche de prise de décision avec accumulation variable d’information, et mesure de la confiance par une loterie post-décisionnelle. Les corrélats cérébraux étaient mesurés par un enregistrement électroencéphalographique à haute densité de 64 canaux. Cette étude est en cours et 23 sujets ont été acquis, l’analyse des résultats comportementaux et d’électroencéphalographie sont en cours.

## 3.6 Résumé et conclusion

L’expérimentation chez l’homme visant à évaluer l’effet des traitements n’est donc pas exempte de biais et de limitations, particulièrement lorsque la recherche se fait en population clinique.

Une limitation majeure dans les progrès possibles est le regroupement syndromique opéré actuellement par les classifications nosographiques qui regroupent sous la même étiquette des patients pour lesquels les phénomènes pathologiques à l’œuvre sont probablement très différents, et à l’inverse, séparant dans des catégories de diagnostic différentes des patients pour lesquels des *dimensions* pathologiques sont communes.

Les endophénotypes sont une voie prometteuse, car elles devraient permettre une meilleure correspondance entre les différents niveaux de lecture des pathologies psychiatriques, entre l’observation comportementale, électrophysiologique ou d’imagerie, les dysfonctions cérébrales et biochimiques à l’œuvre et jusqu’aux anomalies génétiques sous-tendant ces anomalies *dimensionnelles*.

Les neurosciences cognitives peuvent s’inscrire dans cette approche et ont déjà

proposé divers candidats endophénotypes. Le couplage avec la psychopharmacologie pour étudier l'effet de traitements ou pour créer chez l'homme des modèles de certaines dimensions symptomatiques de pathologies psychiatriques est une approche prometteuse. elle tente de caractériser les éléments pathologiques et la réponse au traitement sur des axes qui ne sont plus ceux des descriptions syndromiques, mais qui sont issus de la recherche, en tentant de dégager les dysfonctions fondamentales, qui pourront ensuite être analysées en faisant le lien entre différents niveaux, comportementaux, cognitifs, biologiques, génétiques.

Tous les endophénotypes cognitifs proposés n'ont pas la même capacité à être opérationnalisés dans des paradigmes utilisables aussi bien chez l'homme que chez l'animal. L'analyse de l'apprentissage par renforcement possède probablement cette qualité, nous développons cette approche au prochain chapitre.

# Chapitre 4

## Focus sur l'apprentissage par renforcement

### 4.1 Historique et définition

#### 4.1.1 Historique

Une des premières expérimentations autour de l'apprentissage par renforcement a été effectuée à la fin du 19ème siècle par Edward Thorndike. Il avait construit un dispositif expérimental appelé “puzzle box” (boîte casse-tête) dont la porte ne pouvait s'ouvrir qu'après une suite précise d'actions (tirer sur une chaîne puis appuyer sur un levier par exemple). Il plaçait dans cette boîte un chat affamé et mettait de la nourriture à l'extérieur de la boîte et à la vue du chat. Un animal placé pour la première fois dans cette boîte exécute de nombreuses actions qui n'ont aucune conséquence sur l'ouverture de la porte, gratte la porte, tourne, fait sa toilette. Puis il peut arriver que par hasard, il tire la chaîne, puis un peu plus tard, également par hasard, appuie sur le levier, ouvrant ainsi la porte et obtenant la nourriture. Cet animal lorsqu'il redevient affamé et qu'il est replacé dans la boîte, exécutera moins d'actions qui n'ont pas eu de conséquence sur l'ouverture de la porte, et plus d'actions qui ont amené à l'ouverture de la porte. En renouvelant la procédure un grand nombre de fois, l'animal devient capable de sortir de la boîte en quelques secondes (Thorndike, 1911).

Dans cette expérience, sans instruction particulière et sans supervision, l'animal a appris progressivement, par ses actions et leur conséquence sur l'environnement, à choisir des actions qui améliorent son sort.

Thorndike se réfère au lien entre l'action et l'effet sous le terme “d'instrumentalisation” de l'action. Ces observations l'ont amené à formuler sa “loi de l'effet”, qui est la base de ce qui sera ensuite appelé le “conditionnement instrumental”.

Les travaux sur le conditionnement instrumental ont été poursuivis ensuite par Skinner, développant un environnement de tests encore largement utilisé aujourd'hui : la boîte de skinner. Il a exposé ses travaux dans "Le comportement des organismes", publié en 1938 (Skinner, 2006). Il a ensuite continué à consacrer une grande partie de sa vie à l'étude du conditionnement instrumental et à ses applications au comportement animal et humain. Le champ de l'apprentissage par renforcement s'est donc initialement largement développé sur la base de l'expérimentation animale.

### 4.1.2 Définition

L'apprentissage par renforcement décrit la capacité d'un organisme, humain ou animal, à améliorer progressivement ses choix par essai erreur dans son environnement, pour maximiser la fréquence de conséquences plaisantes (récompenses) ou minimiser la fréquence de conséquences adverses (punitions).

Le vocabulaire de l'apprentissage par renforcement est maintenant classiquement le suivant :

Le renforçateur (ou outcome O) est une récompense ou une punition, qui va avoir pour conséquence de modifier l'association entre l'état ou le stimulus (S) et la réponse (R). Lorsque le renforçateur est une récompense, ses propriétés sont classiquement de générer un comportement d'approche ou de consommation, d'accroître la probabilité et l'intensité des choix d'actions y menant, et de générer un sentiment subjectif de plaisir (Berridge et Robinson, 2003).

Ce qui est renforcé n'est pas décrit de façon unanime. Ainsi, pour Thorndike, ce que le chat apprenait était l'association entre la boîte et la séquence particulière d'actions menant au renforçateur, c'était donc le renforcement entre un stimulus et une réponse (S-R) (Thorndike, 1933). On peut alternativement dire que ce que le chat a appris, c'est le fait qu'une séquence particulière d'actions a amené à un certain résultat, donc un renforcement entre une réponse et un résultat (renforçateur) (R-O). Skinner a concilié ces deux façons de voir les choses dans une conceptualisation à trois termes : le renforçateur (O) renforce la réponse (R) en présence d'un stimulus discriminant (S), soit une association S-R-O (Skinner, 2006).

Les études récentes ont démontré que l'apprentissage était gouverné à la fois par l'association S-R et par l'association R-O (Redgrave et al., 2010). On peut même distinguer plusieurs phases dans l'apprentissage, une phase précoce nommée "goal-directed" où le comportement est essentiellement guidé par le renforçateur (O), et qui est donc une phase essentiellement R-O. Puis vient une phase tardive nommée "habitual" où le comportement devient exprimé automatiquement dans un certain contexte, et devient indépendant du renforçateur, c'est donc une phase essentiellement S-R. Ces différents mécanismes d'apprentissage sont illustrés Figure 4.1.

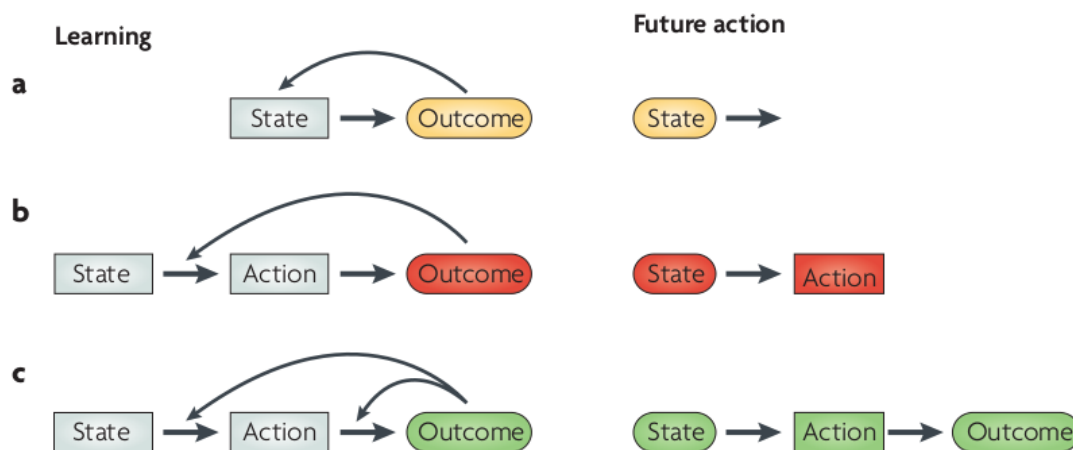


FIGURE 4.1 – Différents mécanismes d'apprentissage par renforcement et de contrôle de l'action. (a) Apprentissage pavlovien : l'association répétée d'un état ou d'un stimulus avec un renforçateur ayant une valence motivationnelle (résultat plaisant ou désagréable) amène une réponse conditionnée, typiquement innée, qui sera répétée lorsque l'état sera rencontré dans le futur. (b) Apprentissage instrumental habituel : si une action effectuée dans un certain état amène une récompense, l'action sera renforcée ; rencontrer à nouveau cet état dans le futur rendra plus probable l'exécution de cette action. (c) Apprentissage instrumental guidé vers un but : si une action effectuée dans un certain état amène une récompense, alors une représentation explicite de la séquence peut être formée ; cette représentation est disponible pour guider le comportement dans le futur. (Seymour et al., 2007)

La nécessité d'une erreur de prédiction peut être comprise intuitivement par le fait que si l'on sait tout prévoir dans notre environnement, il n'y a rien à apprendre, en revanche, si des événements arrivent comme des surprises, c'est que nous devons continuer à apprendre pour améliorer nos prédictions (Niv et Schoenbaum, 2008). La nécessité d'une erreur de prédiction pour l'existence d'un apprentissage a été démontrée expérimentalement dans le paradigme de "blocage de Kamin" (Kamin, 1969), où un stimulus est présenté avec un autre stimulus qui était déjà associé à une récompense : la récompense étant reçue comme prévu, il n'y a pas d'erreur de prédiction et pas d'apprentissage d'une association entre ce nouveau stimulus et la récompense (Waelti et al., 2001).

## 4.2 Formalisation et modélisation computationnelle

### 4.2.1 Formalisation

L'apprentissage par renforcement c'est donc apprendre quelles actions entreprendre pour maximiser un résultat (outcome) donné (Sutton et Barto, 1998). Il est possible de choisir une représentation numérique de l'outcome obtenu et ainsi de modéliser l'apprentissage par renforcement de façon computationnelle.

Il diffère de l'apprentissage supervisé où des exemples d'actions correctes et incorrectes sont donnés explicitement, et de l'apprentissage non supervisé strict où il n'y a pas d'erreur de prédiction (Dayan et Abbott, 2005).

L'apprentissage par renforcement est généralement formalisé comme un processus de Markov (Howard, 1960). A chaque pas de temps  $t$ , l'acteur se trouve dans un état  $s$  et peut choisir parmi plusieurs actions  $a$ . Choisir l'action  $a$  donne un outcome (récompense)  $r(s, a)$ , et amène dans l'état  $s'$  avec une certaine probabilité  $P(s, a, s')$ . Le fait que la probabilité d'arriver dans l'état  $s'$  ne dépende que de l'état  $s$  et de l'action  $a$ , indépendamment de l'histoire des états et actions antérieures constitue la propriété markovienne du processus.

Les algorithmes d'apprentissage par renforcement cherchent à résoudre deux problèmes : un problème de prédiction et un problème de contrôle. La résolution du problème de prédiction permet d'obtenir une fonction de valorisation indiquant pour chaque état  $s$  quelle valeur de renforcement  $r$  futur on peut attendre lorsque l'on effectue l'action  $a$  étant donné une "policy" ou conduite d'action choisie. La résolution du problème de contrôle permet d'obtenir une fonction de conduite qui maximise les renforcements obtenus de l'état initial à l'état final. Les deux problèmes sont interdépendants et nécessitent d'être résolus de façon conjointe.

Les variables d'un problème d'apprentissage par renforcement sont donc :

- un ensemble d'états  $s$  ;
- un ensemble d'actions  $a$  ;
- des règles de transition entre états  $P(s, a, s')$  ;
- des règles déterminant la délivrance de renforçateurs  $r(s, a)$ .

Les règles déterminant les probabilités de transition entre états  $P(s, a, s')$ , c'est-à-dire la probabilité d'entrer dans l'état  $s'$  sachant qu'on a effectué l'action  $a$  dans l'état  $s$ , et la délivrance des renforçateurs  $r(s, a)$ , c'est-à-dire la quantité de renforçateur reçu après avoir fait l'action  $a$  dans l'état  $s$ , représentent la façon dont l'environnement fonctionne.

Lorsque ce fonctionnement est connu à l'avance par le sujet, on n'est pas dans l'apprentissage par renforcement stricto sensu, et le champ qui étudie la résolution de ce problème est le "dynamic programming", le comportement optimal est alors déterminé par les équations de Bellman (Bellman, 1957).

Lorsque le fonctionnement de l'environnement n'est pas connu à l'avance, la résolution du problème de prédiction et du problème de contrôle se fait par apprentissage progressif. On est alors pleinement dans l'apprentissage par renforcement, la résolution du problème se faisant par essais erreurs, en fonction des résultats de l'interaction avec l'environnement. Plusieurs types d'algorithmes ont été formalisés pour converger vers une fonction de prédiction et une fonction de contrôle optimal.

Le champ de l'apprentissage par renforcement est très actif et un grand nombre d'algorithmes différents ont été développés pour modéliser le comportement de sujets dans ce type de problème. Parmi les plus utilisés, on peut citer l'algorithme "temporal difference" développé par Sutton et Barto (Sutton et Barto, 1998) qui consiste à résoudre la fonction de valorisation. Cet algorithme est une formalisation du modèle mathématique introduit par Rescorla-Wagner qui avait donné une place centrale à l'erreur de prédiction pour expliquer le conditionnement classique (Rescorla et al., 1972). Un autre algorithme très utilisé, le Q-learning, étend l'algorithme "temporal difference" et y ajoute la résolution de la fonction de conduite (Watkins et Dayan, 1992).

### 4.2.2 Algorithme "Temporal difference"

Etant donné une conduite  $\pi$ , et une séquence d'états suivis de renforçateurs  $s_1, r_1, s_2, r_2, \dots, s_t, r_t, s_{t+1}, r_{t+1}, \dots, s_T, r_T$ , le problème de prédiction est de trouver la fonction de valorisation des états  $V^\pi(s_t)$ .

Le renforçateur cumulé attendu dans le futur à partir de l'état  $s_t$  est

$$R_t = r_{t+1} + \gamma_1 r_{t+2} + \dots + \gamma_{T-t-1} r_T$$

où  $0 < \gamma < 1$  est un facteur d'actualisation traduisant le fait que des renforçateurs

qui seront reçus dans longtemps sont moins valorisés que des renforçateurs de même valeur qui seront reçus immédiatement.

Dans le cadre de l'apprentissage par renforcement, on suppose que la valeur accordée à un état  $V(s)$  est équivalente au renforçateur cumulé attendu dans le futur à partir de cet état.

$$V^\pi(s) = R_t(s_t = s)$$

Un estimateur  $V(s_t)$  de  $V^\pi(s_t)$  peut être estimé itérativement sur la base de ce qui se passe après la visite de l'état  $s_t$  :

$$V(s_t) \leftarrow V(s_t) + \alpha(R_t - V(s_t)) \quad (4.2.1)$$

où  $0 < \alpha < 1$  est un taux d'apprentissage qui représente dans quelle mesure la valeur de l'état  $V(s_t)$  est mise à jour en fonction des renforçateurs effectivement reçus.

On a ici une formulation qui est équivalent à la formalisation de Rescorla-Wagner, avec un terme  $R_t - V(s_t)$  qui représente une erreur de prédiction du renforçateur, c'est-à-dire la différence entre le renforçateur cumulé total et ce qui était prévu. Cette erreur de prédiction a une place centrale ici comme chez Rescorla-Wagner, car si la fonction de valorisation  $V(s_t)$  prédit parfaitement le renforçateur cumulé total, l'erreur de prédiction sera nulle et la valeur estimée de l'état  $V(s_t)$  ne sera pas modifiée, ce qui veut dire qu'on a convergé vers la valeur finale de  $V(s_t)$ .

Cette méthode demande cependant d'attendre la fin de la séquence pour pouvoir calculer le renforçateur cumulé global  $R_t$  pour ensuite mettre à jour les valeurs d'état  $V(s_t)$ .

Dans l'algorithme "temporal difference", on n'attend pas la fin de la séquence, mais dès le pas de temps suivant, on utilise le renforçateur reçu au temps  $t + 1$  et l'estimation de la valeur de l'état au temps  $t + 1$  comme approximation de la valeur des outcome cumulés à recevoir  $R_t$ . On peut ainsi dès le pas de temps  $t + 1$  mettre à jour l'estimation de la valeur de l'état  $s$  au temps  $t$  :

$$V(s_t) \leftarrow V(s_t) + \alpha[r_{t+1} + \gamma V(s_{t+1}) - V(s_t)] \quad (4.2.2)$$

En définissant l'erreur de prédiction comme :

$$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t) \quad (4.2.3)$$

On peut maintenant formuler plus simplement la méthode d'estimation de la fonction de prédiction optimale :

$$V(s_t) \leftarrow V(s_t) + \alpha \delta_t \quad (4.2.4)$$

Il a été démontré que pour une conduite  $\pi$  donnée, la méthode d'estimation séquentielle de la fonction de prédiction  $V(s_t)$  décrite ci-dessus converge vers la fonction optimale  $V^\pi$  en moyenne si le taux d'apprentissage  $\alpha$  est suffisamment petit, et avec une probabilité égale à 1 si le taux d'apprentissage  $\alpha$  remplit certaines conditions :  $\sum_{k=1}^{\infty} \alpha_k = \infty$  et  $\sum_{k=1}^{\infty} \alpha_k^2 < \infty$ .

### 4.2.3 Algorithme “Q-learning”

Dans l'algorithme Q-learning, la fonction de prédiction dépend non seulement des états  $s_t$  mais également des actions  $a_t$ . Ce sont des valeurs d'état-action plutôt que des valeurs d'état, qui représentent donc la valeur du choix d'une certaine action dans un certain contexte (Watkins et Dayan, 1992).

Une simple adaptation de l'algorithme temporel difference permet de calculer sur le même principe une estimation de la fonction de prédiction optimale :

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \delta_t \quad (4.2.5)$$

avec

$$\delta_t = r_{t+1} + \gamma \max_a [Q(s_{t+1}, a) - Q_t(s_t, a)] \quad (4.2.6)$$

Alors il peut être démontré que la fonction  $Q$  ainsi calculée converge vers la fonction  $Q^*$  optimale sous réserve que toutes les paires état-action continuent à être explorées.

Dans cet algorithme, on obtient donc une estimation des valeurs des actions, ce qui permet de définir une fonction de conduite (ou de sélection des actions). La méthode la plus simple est de choisir à chaque pas de temps  $t$  l'action qui a la valeur  $Q(s_t, a_t)$  la plus élevée. Cette méthode simple, appelée “greedy”, exploite la connaissance accumulée sur l'environnement, mais a l'inconvénient de ne plus explorer les actions qui se sont révélées moins avantageuses. Elle ne laisse donc pas de place à l'exploration, ce qui peut être désavantageux dans un environnement probabiliste qui change de façon dynamique.

Une modification simple à cette règle de conduite est de se comporter de façon “greedy” la plupart du temps, mais de sélectionner une action de façon aléatoire et indépendamment des valeurs relatives des actions de temps à autre (avec une probabilité  $\epsilon$ ). Cette règle de sélection d'actions appelée “ $\epsilon$ -greedy” permet de s'assurer que si le nombre d'essais est suffisant, toutes les actions continuent à

être explorées, ce qui améliore la convergence de la fonction de valorisation vers les valeurs optimales (Pearl, 1984). C'est aussi une façon de résoudre le problème de choisir l'exploration de l'environnement ou l'exploitation des connaissances accumulées. Dans ce cas cependant, l'exploration se fait vers n'importe quelle autre action, indépendamment de sa valeur. Une action alternative jugée très mauvaise serait donc choisie aussi fréquemment qu'une action alternative jugée seulement un peu moins bonne que l'action optimale.

Une modification qui permet que se poursuive une exploration, mais qui explore les actions alternatives en pondérant la probabilité de les explorer à leur valeur relative est appelée la règle de sélection d'action "softmax". La probabilité de choisir l'action  $a$  parmi l'ensemble possible des actions  $a^i, i = 1..n$  est donné par la formule :

$$P_t(a) = \frac{\exp(Q_t(a)/\beta)}{\sum_{i=1}^n \exp(Q_t(a^i)/\beta)}$$

où  $\beta > 0$  est un paramètre appelé la température. Une température basse aura pour conséquence que les différences de valeurs entre action a une grande importance sur la probabilité de sélection des actions, les actions ayant une valeur faible seront très peu souvent explorées. Une température très élevée, à l'inverse, aura tendance à lisser les différences de valeurs, les actions alternatives étant à l'extrême choisies de façon équiprobable.

## 4.3 Bases cérébrales de l'apprentissage par renforcement

### 4.3.1 Cerveau et récompense

L'étude du lien entre cerveau et récompense a débuté avec les expériences de Olds et Milner qui ont implanté dans le cerveau de rats des électrodes qui délivraient un courant électrique lorsque ces rats appuyaient sur un levier (Olds et Milner, 1954). Des électrodes étaient placées dans différentes régions du cerveau. Les auteurs ont observé un résultat remarquable : dans certains cas, les rats appuyaient volontairement et de façon répétée pour s'auto-administrer des stimulations électriques dans le cerveau. Cette activité était parfois tellement prenante pour les rats que certains ne se nourrissaient pas et ne buvaient plus. Les décharges neuronales déclenchées par les stimulations électriques dans certaines zones du cerveau poussaient donc les animaux à répéter l'action de presser le levier, ce qui est une manifestation classique de la réception d'une récompense comparable à ce qui était obtenu dans les expériences de Pavlov ou de Thorndike. Un corrélat cérébral de récompense, amenant un renforcement de l'action, avait ainsi été identifié (Olds et Milner, 1954).

Des études ultérieures ont permis de déterminer que de tels comportements d'auto-stimulation électrique intervenaient préférentiellement lorsque les électrodes étaient en contact avec des neurones dopaminergiques (Corbett et Wise, 1980).

Puis a été montré que ces neurones ne devaient pas nécessairement être activés artificiellement par des électrodes : ce sont bien ces neurones dopaminergiques qui sont actifs et déchargent lorsqu'un animal reçoit une récompense (Schultz, 1986). Le lien entre renforcement par la récompense et circuits dopaminergiques était ainsi démontré.

Les neurones dopaminergiques ont pour origine essentielle la substantia nigra pars compacta (SNpc) et l'aire tegmentale ventrale (VTA) (Gardner et Ashby, 2000). Les fibres dopaminergiques innervent principalement des structures sous-corticales en particulier le striatum, mais aussi des structures corticales et plus particulièrement le cortex préfrontal.

On distingue classiquement quatre voies dopaminergiques principales (Stahl, 2013) :

- la voie nigrostriée qui partant de la substantia nigra innerve le striatum ; cette voie fait partie du système extrapyramidal et contrôle le mouvement ;
- la voie mésolimbique qui partant du VTA innerve des structures sous-corticales (striatum ventral, amygdale) et corticales (hippocampe et des parties du cortex préfrontal) faisant partie du système limbique ;
- la voie mésocorticale qui partant du VTA innerve essentiellement des parties préfrontales ;
- la voie tubéro-infundibulaire qui part de l'hypothalamus et projette sur l'hypophyse, contrôlant la sécrétion de prolactine.

Une voie supplémentaire partant de différents sites comme la substance grise périaqueducule, l'hypothalamus et les noyaux latéraux parabrachiaux innervant le thalamus a récemment été mise en évidence chez l'animal, sa fonction est pour le moment mal connue.

#### 4.3.2 Cerveau et erreur de prédiction

Si le lien entre dopamine et récompense avait été fait depuis longtemps, ce sont les travaux de Schultz dans les années 90, qui vont permettre de réellement faire le lien avec les mécanismes d'apprentissage par renforcement.

Schultz montre que la réponse dopaminergique encode quelque chose de plus que la récompense. Si les neurones dopaminergiques déchargent lorsqu'une récompense non prévue est reçue, cette réponse dopaminergique disparaît lorsque la récompense est parfaitement prévisible. C'est alors au moment de la présentation du stimulus prédisant de façon fiable la récompense future qu'une décharge dopaminergique est alors observée. A l'inverse, si une récompense prévue n'est pas reçue, ou est moindre que prévue, alors l'activité des neurones dopaminergiques diminue (Figure 4.2). La

démonstration était faite : la réponse dopaminergique reflète une erreur de prédiction de la récompense (Schultz et al., 1997), brique essentielle de l'apprentissage par renforcement.

L'hypothèse d'un encodage dopaminergique de l'erreur de prédiction a été confirmée par les études ultérieures. Ces études ont vérifié que cette hypothèse était compatible avec des résultats jusque là mal expliqués, comme le blocage de Kamin. Dans ce dispositif, un stimulus est présenté avec un autre stimulus qui était déjà associé à une récompense. La récompense étant reçue comme prévu, il n'y a pas d'erreur de prédiction et pas d'apprentissage d'une association entre ce nouveau stimulus et la récompense (Waelti et al., 2001). Il a été montré plus tard que la réponse dopaminergique varie de façon monotone avec la probabilité (Fiorillo et al., 2003) et avec la magnitude de la récompense (Tobler et al., 2005).

La question de l'encodage non seulement des erreurs de prédiction ou des valeurs d'état, mais de la valeur des actions a été posée. La représentation cérébrale de la valeur des actions est en effet nécessaire, dans ce formalisme, pour expliquer le conditionnement instrumental. Des études ont retrouvé des corrélats de la valeur des actions dans l'activité des neurones du tronc (Morris et al., 2006), du striatum (Cromwell et Schultz, 2003 ; Lau et Glimcher, 2007 ; Samejima et al., 2005) et du cortex préfrontal (Kobayashi et al., 2002 ; Matsumoto et al., 2003, 2007).

Les travaux d'imagerie fonctionnelle ont permis de confirmer les bases cérébrales de l'apprentissage par renforcement chez l'homme également. Les travaux initiaux ont montré que des récompenses (du jus de fruit) induisaient une activation du striatum et du cortex préfrontal ventro-médian uniquement si elles étaient inattendues (donc lorsqu'une erreur de prédiction positive existe), pas quand elles étaient prévues (Berns et al., 2001). Après un entraînement intensif dans une tâche de conditionnement classique, l'indice annonceur d'une récompense engendrait une activation dans le striatum, l'amygdale et le cortex orbito-frontal, et plus la réception de la récompense prévue (O'Doherty et al., 2002). Des études ultérieures ont précisé plus avant les différences entre conditionnement classique du conditionnement instrumental, en montrant que le striatum dorsal, et peut-être plus particulièrement le putamen, avait un rôle plus spécifique pour la sélection d'action (Haruno et Kawato, 2006 ; Kahnt et al., 2009 ; Schönberg et al., 2007).

Une intervention pharmacologique avec un agoniste dopaminergique, la levodopa, un antagoniste dopaminergique, l'haloperidol et un placebo a permis de montrer que le striatum avait une activité qui corrélait avec l'erreur de prédiction, que l'intervention dopaminergique ces signaux, la levodopa les amplifiant et l'halopéridol les diminuant (Pessiglione et al., 2006), avec des conséquences cohérentes en termes de performance d'apprentissage. Cette manipulation pharmacologique apportait un argument supplémentaire fort pour une implication causale de la modulation dopaminergique dans l'apprentissage par renforcement (Pessiglione et al., 2006). Une étude d'imagerie ultérieure a montré que l'activité des neurones dopaminergique de l'aire tegmentale ventrale (VTA) était corrélée à l'erreur de prédiction (D'Ardenne

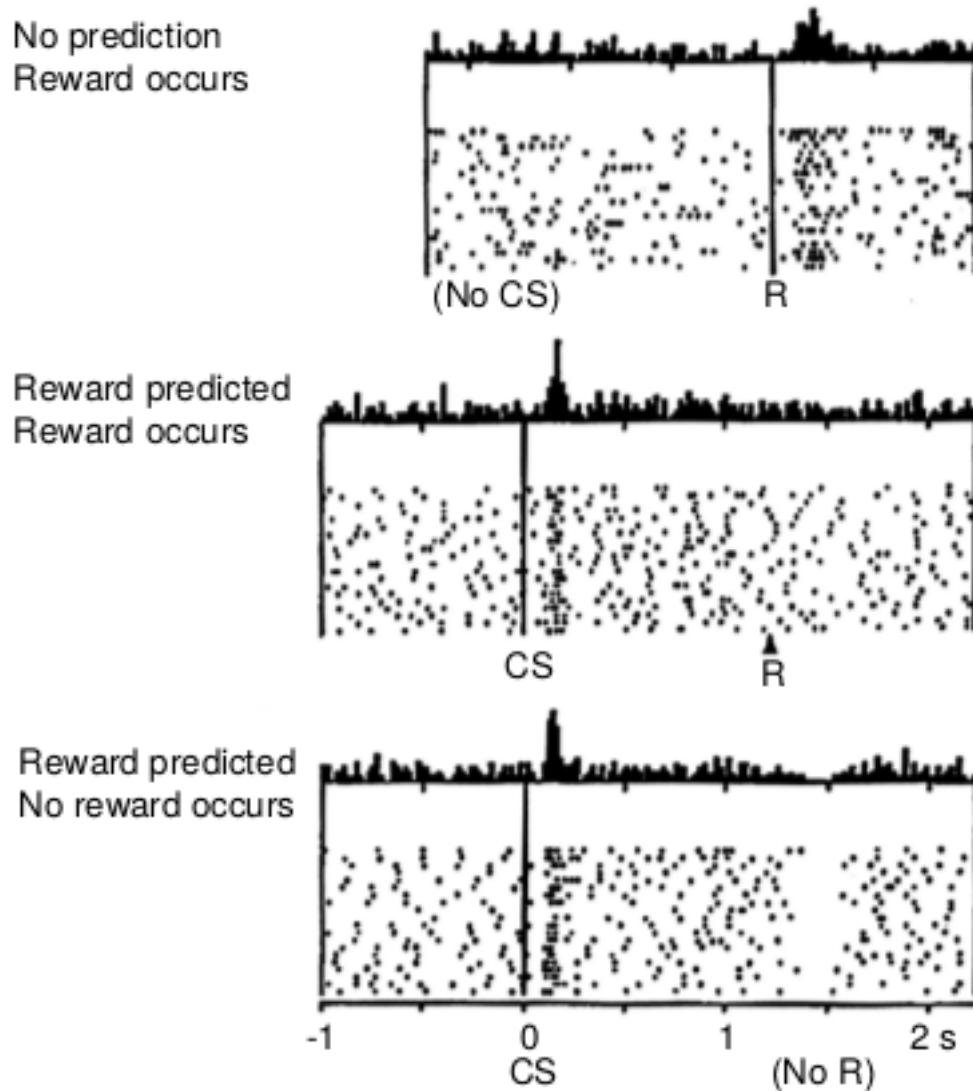


FIGURE 4.2 – La réponse dopaminergique encode l'erreur de prédiction de la récompense. Le neurone dopaminergique est activé lorsqu'une récompense non prévue est reçue (erreur de prédiction positive, haut), il ne montre pas de réponse lorsque la récompense est parfaitement prédite (pas d'erreur de prédiction, milieu), et diminue son activité lorsqu'une récompense attendue n'est pas reçue (erreur de prédiction négative, bas) (Schultz et al., 1997)

et al., 2008).

Au total, l'apprentissage par renforcement sur la base de récompenses a été bien documenté avec des études montrant le rôle clef de la dopamine, et l'implication du striatum et du cortex préfrontal ventro-médian, y compris chez l'homme avec des interventions pharmacologiques (Cools et al., 2009 ; Frank et al., 2004 ; Pessiglione et al., 2006).

### 4.3.3 L'apprentissage par punition

Si l'apprentissage par renforcement sur la base de récompenses a été bien documenté, l'apprentissage sur la base de punitions est moins connu (Dayan, 2012).

Un paradoxe existe d'ailleurs du point de vue de la formalisation de l'apprentissage par renforcement puisque dans l'algorithme, une fois une punition évitée, elle n'est plus renforcée, ce qui conduit le modèle à prédire que l'apprentissage sur la base de punitions sera moins bon que l'apprentissage à base de récompenses. Or les expérimentations chez l'homme indiquent que la vitesse d'apprentissage est comparable dans les deux conditions (Pessiglione et al., 2008, 2006).

Certains auteurs défendent une vision uniciste, les mêmes aires cérébrales seraient engagées dans l'apprentissage aversif, et la dopamine guiderait l'apprentissage dans les deux cas, avec pour le codage des erreurs de prédiction négative une baisse dans les décharges dopaminergiques (Schultz et al., 1997).

Le rôle de la dopamine dans l'apprentissage appétitif comme aversif est soutenue par des études pharmacologiques. Les patients présentant un syndrome de Parkinson ont une déplétion en dopamine des ganglions de la base, leur performances dans l'apprentissage par renforcement était spécifiquement détériorées pour les récompenses et non pour les punitions lorsque ces patients étaient sans traitement (Frank et al., 2004). Sous traitement par agoniste dopaminergique, cette asymétrie se renverse et les patients apprennent mieux des récompenses que des punitions (Frank et al., 2004). C'est un argument pour un rôle central de la dopamine dans l'apprentissage appétitif aussi bien qu'avérsif. Elle agirait à la fois en stimulant la voie directe ("Go") qui facilite la réponse comportementale et inhibant la voie indirecte ("No-Go") qui supprime la réponse comportementale. Les bouffées de libération de dopamine augmenteraient la plasticité de la voie directe et abaisseraient celle de la voie indirecte. Ce serait le support de l'apprentissage par renforcement basé sur la recherche des récompenses. Une diminution de la libération de dopamine auraient l'effet inverse, ce serait le support de l'apprentissage par renforcement basé sur l'évitement des punitions (Frank et al., 2004). Ces résultats ont été répliqués et étendus à des tâches différentes chez le patient parkinsonien, ou à des pathologies différentes comme le syndrome Gilles de la Tourette, étudiant l'effet d'un traitement antipsychotique (Palmiter et al., 2009).

Cette capacité de la dopamine à encoder aussi bien les erreurs de prédiction sur les

récompenses que les punitions a cependant été critiquée. Bayer et Glimcher ont montré que si le taux de décharge des neurones dopaminergiques après réception de la récompense est linéairement corrélée à l'erreur de prédiction lorsque le renforçateur est positif, ce n'était pas le cas quand le renforçateur était négatif (Bayer et Glimcher, 2005). Les auteurs émettent l'idée que la dopamine sous-tendrait les signaux d'erreur de prédiction uniquement pour les récompenses, et qu'il existerait un système complémentaire pour les punitions.

Daw et collaborateurs a présenté un modèle dans lequel le système sérotoninergique aurait ce rôle mutuellement opposé au système dopaminergique. Dans leur modèle, les décharges phasiques de sérotonine encodent les erreurs de prédiction concernant les punitions, ce qui sous-tendrait l'apprentissage par renforcement basé sur l'évitement des punitions (Daw et al., 2002). Des études d'intervention pharmacologique sur cette hypothèse ont apporté des résultats contrastés. Cools et collaborateurs ont étudié l'effet de la déplétion en tryptophane, un précurseur de la sérotonine sur le sujet sain. Ils montrent que cette intervention augmente la prédiction des punitions, mais pas des récompenses (Cools et al., 2008), ce qui est en faveur d'un rôle spécifique de la sérotonine pour l'apprentissage aversif. Une autre recherche a étudié l'effet d'un traitement par inhibiteur de recapture de la sérotonine chez le sujet sain. Les auteurs retrouvent une diminution de la réponse cérébrale sous traitement sérotoninergique, pour les stimuli appétitifs comme pour les stimuli aversifs (McCabe et al., 2010). De même une étude a recherché l'effet d'un traitement par inhibiteur de recapture de la sérotonine chez des patients présentant un trouble obsessionnel compulsif. Sans traitement, ces derniers présentent un déficit dans l'apprentissage par récompenses comme par punitions. Le traitement par inhibiteur de recapture de la sérotonine corrige ces déficits dans les deux conditions (Stefano Palminteri, Clair, et al., 2012). La manipulation pharmacologique de la sérotonine n'induisent pas d'asymétrie entre récompense et punition dans ces deux dernières expériences, ce qui va à l'encontre d'un rôle spécifique de la sérotonine pour l'apprentissage par punition (Stefano Palminteri, Clair, et al., 2012).

Le même débat existe concernant les circuits cérébraux sous-tendant l'apprentissage par renforcement basé sur les punitions. Sont-ils les mêmes que ceux qui sous-tendent l'apprentissage par la récompense, ou un système distinct et complémentaire existe-t-il ?

Plusieurs candidats ont été proposés, les plus mentionnés étant l'insula antérieure, et le cortex préfrontal dorso-médian (pour des revues, voir (Bartra et al., 2013; Garrison et al., 2013; Liu et al., 2011)). Il a ainsi été montré que l'insula s'activait non seulement à réception de punitions, mais également à la présentation des stimuli prédisant ces punitions, et ce à la fois pour de stimuli physiquement aversifs comme des chocs électriques ou des goûts désagréables et pour des stimuli plus abstraits comme des pertes financières (Kim et al., 2011; Nitschke et al., 2006). Une étude chez des patients présentant des tumeurs cérébrales donnait des arguments pour l'implication causale de l'insula antérieure de façon sélective dans l'apprentissage aversif (Stefano Palminteri, Justo, et al., 2012). Mais dans cette

même étude, l'atteinte du striatum affectait également l'apprentissage aversif, les auteurs proposant un rôle complémentaire de ces deux structures, l'insula sous-tendrait l'apprentissage des valeurs négatives, et le striatum serait impliqué dans la sélection d'action sur la base de ces valeurs (Stefano Palminteri, Justo, et al., 2012). D'autres études retrouvent une activation de l'insula avec des renforçateurs positifs, et d'autres régions comme le cortex orbito-frontal a été rapporté comme encodant des renforçateurs négatifs (Gottfried et Dolan, 2004 ; Kirsch et al., 2003).

Au total, la question de savoir si ce sont les mêmes régions (le striatum ventral et le cortex préfrontal ventro-médian) qui sous-tendent l'apprentissage de valeurs positives et négatives, ou si l'apprentissage de valeurs négatives est sous-tendu par un système opposé (l'insula et le cortex préfrontal dorso-médian) n'est pas tranchée.

Une théorisation plus récente du rôle de la dopamine et de la sérotonine dans l'apprentissage par renforcement fait la distinction entre valence et action, et les orthogonalise (Guitart-Masip et al., 2014). Les auteurs partent de la constatation du fait que dans un apprentissage instrumental, on peut demander à un sujet d'exécuter une action ou au contraire lui demander de retenir une action, dans le but d'obtenir une récompense. Dans ce cadre instrumental, et contrairement à l'apprentissage Pavlovien, il n'y a donc pas de lien à priori entre le type d'action à effectuer (approche / évitement) et la valence du renforçateur reçu (récompense / punition). Les auteurs assignent à la dopamine une responsabilité dans les contextes d'approche et de récompense, et à la sérotonine une responsabilité dans les contextes d'évitement et de punition (Figure 4.3) (Guitart-Masip et al., 2014). Cette distinction pourrait expliquer la variété des résultats obtenus concernant le rôle de la dopamine et de la sérotonine, selon le type d'action demandé au sujet dans un même contexte (récompense ou punition).

#### 4.3.4 L'apprentissage contrefactuel

Si l'intérêt pour un sujet, humain ou animal, d'avoir une représentation cérébrale de la valeur du choix effectué est clair, il pourrait y avoir également un intérêt à maintenir une représentation cérébrale des conséquences des choix qui n'ont pas été faits (Boorman et al., 2011).

L'observation que les sujets humains ne se comportent pas comme le voudraient les théories économiques classiques a été le point de départ. Pour rendre compte de cette observation, une des propositions était que les individus pouvaient ressentir du regret et de la réjouissance en conséquence de leurs choix, et que lorsqu'un choix doit être fait en situation d'incertitude, ces individus essaient d'anticiper et de prendre en compte ces éléments (Loomes et Sugden, 1982).

L'étude expérimentale de ce rôle du regret dans la prise de décision s'est faite en utilisant des paradigmes d'apprentissage (Coricelli et al., 2005). Le regret, versant

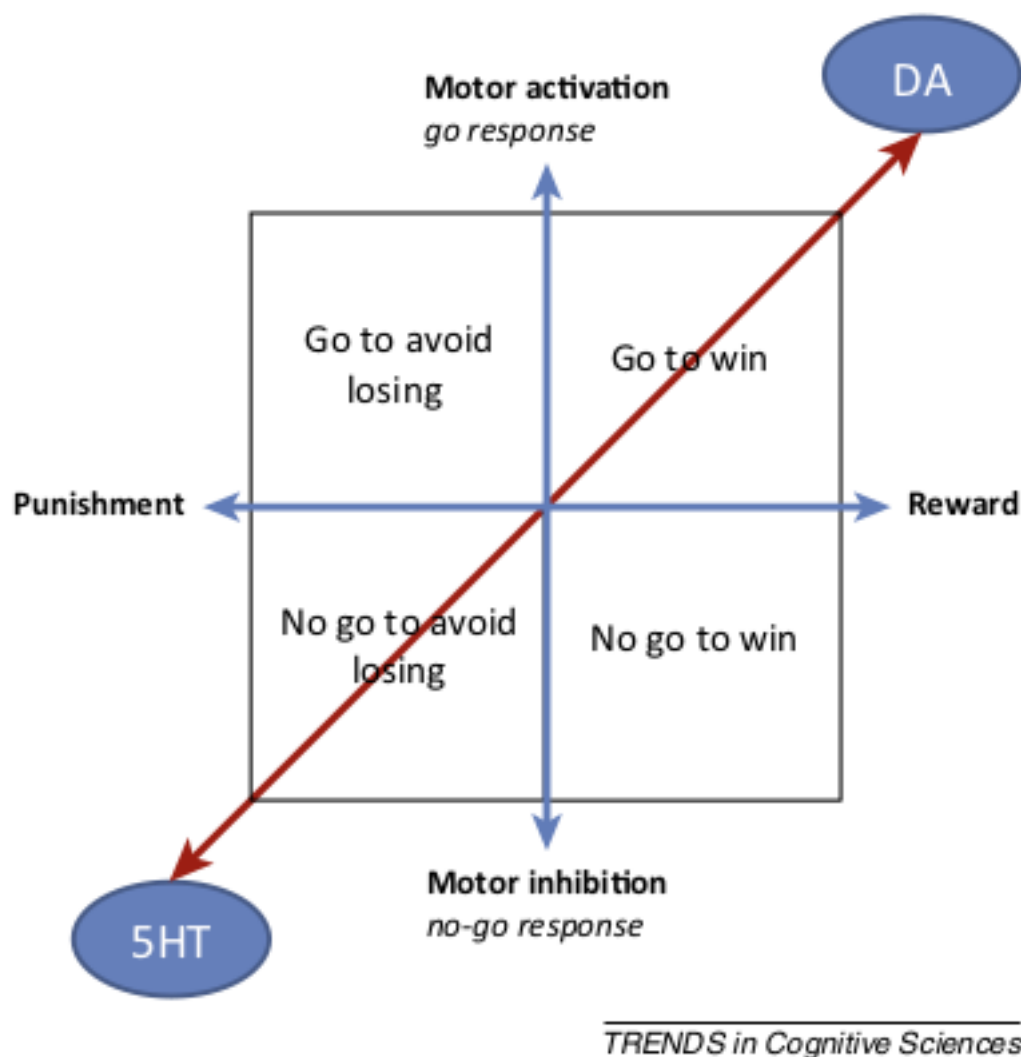


FIGURE 4.3 – Deux axes du contrôle du comportement, valence (récompense versus punition) et action (activation motrice ou approche versus inhibition motrice). Ces deux systèmes sont indépendants dans un contexte d'apprentissage instrumental (en bleu), le sujet pouvant apprendre à faire une action aussi bien qu'à retenir une action pour obtenir une récompense. A l'inverse, dans un contexte d'apprentissage Pavlovien, ces deux systèmes sont liés (en rouge), car la récompense est associée à une activation comportementale, l'approche, alors que la punition est associée à l'inhibition de l'action. Selon ce modèle, au niveau neuronal, la Dopamine serait associée aux comportements d'approche et à la prédiction de récompense, tandis que la sérotonine serait associée à l'inhibition comportementale dans des contextes aversifs (Guitart-Masip et al., 2014).

émotionnel, a été distingué de la pensée contrefactuelle, le versant cognitif (Coricelli et Rustichini, 2010). Le regret est en effet conceptualisé par les neurosciences cognitives de la prise de décision comme une émotion négative résultant de l'écart entre le résultat d'une action choisie et d'une autre action non choisie (Baskin-Sommers et al., 2016). C'est ce qu'il ressent lorsqu'il apprend que le résultat de son choix est moins bon que ce qu'il aurait pu obtenir s'il avait fait un choix différent (Zeelenberg et Pieters, 2007).

Les sujets humains sont en général averses au regret. Face à un choix comportant plusieurs options, ils estiment et prennent en compte le regret qui pourrait résulter de leur choix (Coricelli et al., 2005). Pour ce faire, ils doivent être capables de comparer les valeurs relatives des actions qui ont été choisies et des actions qui n'ont pas été choisies. Cette activité de simulation mentale des valeurs des actions non expérimentées est baptisée pensée contrefactuelle, ou inférence contrefactuelle. Elle serait fondamentale pour expliquer l'influence du regret dans la prise de décision observée par les économistes (Coricelli et Rustichini, 2010).

Les bases neurales de la pensée contrefactuelle ont commencé à être explorées.

Certains auteurs retrouvent l'implication d'un réseau incluant le cortex préfrontal, en particulier le cortex préfrontal dorso-médian, postero-médian et fronto-polaire latéral (Boorman et al., 2011). Ce réseau sous-tendrait le codage des valeurs contrefactuelles ainsi que les erreurs de prédiction contrefactuelles (Boorman et al., 2011). Des études d'électroencéphalographie apportent également des arguments pour un processus spécifique et initialement séparé du traitement des valeurs factuelles, avant que les deux processus ne convergent (Fischer et Ullsperger, 2013). Le traitement des inférences contrefactuelles est parfois attribué au cortex fronto-polaire (Koechlin, 2014), ou au cortex orbito-frontal (Camille et al., 2004 ; Sommer et al., 2009).

Cependant pour d'autres auteurs, le traitement contrefactuel serait sous-tendu par les mêmes systèmes neuronaux que le traitement des choix factuels, en particulier par le striatum (Li et Daw, 2011 ; Lohrenz et al., 2007).

Malgré la jeunesse du domaine de recherche, des propositions de vision complète des différentes étapes dans le traitement contrefactuel émergent, avec l'activation des simulations mentales, l'inférence contrefactuelle, et l'apprentissage avec adaptation ; elles tentent de lier les différents types de traitement à l'activation de différentes aires cérébrales (Van Hoeck et al., 2015) (Figure 4.4).

Au plan pharmacologique, peu d'interventions ont été tentées pour étudier l'impact sur le raisonnement contrefactuel. Avec Tobia et collaborateurs des sujets sains ont effectué une tâche d'investissement séquentiel mettant en jeu le raisonnement contrefactuel avec intervention pharmacologique. Une déplétion en tryptophane était réalisée sur un groupe de sujets, une déplétion en tyrosine et phénylalanine sur un autre groupe, ces interventions diminuaient respectivement les taux de sérotonine et de dopamine. Une intervention placebo était pratiquée sur un troisième groupe. Les

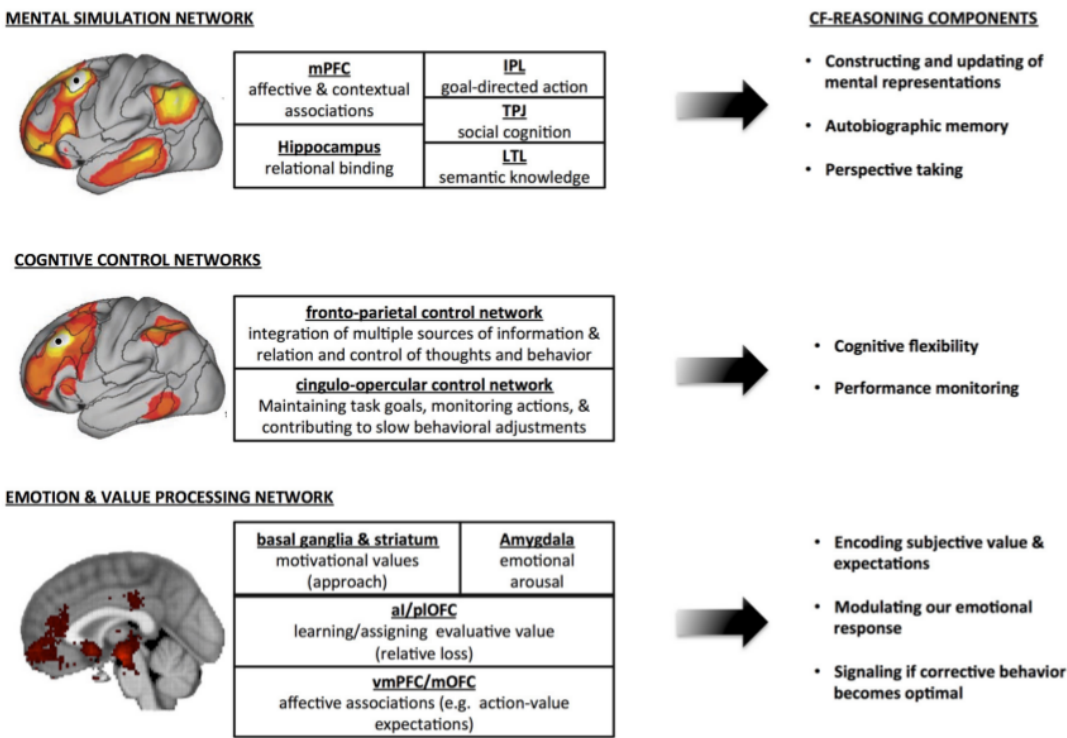


FIGURE 4.4 – Proposition de vision globale des différentes aires cérébrales participant au raisonnement contrefactuel et des différentes fonctions cognitives s’y rapportant. CF contrefactuel, mPFC cortex préfrontal médian, IPL lobe inféro-pariétal, TPJ jonction temporo-pariétale, LTL lobe temporal latéral, aI insula antérieure, pOFC cortex orbito-frontal postéro-latéral, vmPFC cortex préfrontal ventro-médian, dmPFC cortex préfrontal dorso-médian (Van Hoeck et al., 2015)

résultats comportementaux n'étaient pas affectés par cette intervention. L'analyse en imagerie a cependant montré des modifications spécifiques à chacune des deux manipulations. Pour les auteurs, l'apprentissage contrefactuel est sous-tendu par des aires corticales et sous-corticales qui impliquent une interaction entre dopamine et sérotonine (Tobia et al., 2014).

### 4.3.5 L'apprentissage model-based et model-free

Avant même l'intérêt pour l'apprentissage contrefactuel, des observations avaient déjà indiqué que l'apprentissage observé chez les animaux ne pouvait se résumer à l'acquisition de valeurs au gré de renforçateurs reçus. Ce sont par exemple les observations d'apprentissage latent, mis en évidence lorsque des rats explorent un labyrinthe en l'absence de tout renforcement (Tolman et Honzik, 1930), ou les procédures de type effet d'un incitateur sans objet (Krieckhaus et Wolf, 1968).

Dans les expériences de Tolman, plusieurs groupes de rats évoluaient dans un labyrinthe complexe. Un groupe de rats recevaient systématiquement une récompense en cas de succès sous forme de nourriture. Un autre groupe ne recevait aucun renforçateur initialement, ceux-ci étaient introduits à partir du 11ème jour seulement. Le temps de parcours et le taux d'erreurs des rats initialement non renforcés a alors diminué de façon rapide et importante à l'introduction du renforçateur (Figure 4.5). Les performances de ces rats ont rattrapé puis dépassé celles des rats qui recevaient un renforcement depuis le début (Tolman et Honzik, 1930). L'hypothèse formulée est que ces animaux ont constitué, durant la période où aucun renforcement n'était reçu, une carte cognitive, un modèle de leur environnement. Ce modèle interne leur a ensuite permis, une fois motivés par une récompense, de prendre plus facilement la meilleure décision à chaque intersection pour obtenir cette récompense. L'observation des hésitations motrices et des attitudes de regard en avant et en arrière des rongeurs aux endroits où un choix est possible, est venu renforcer cette hypothèse. Ce comportement est appelé "vicarious trial and error (VTE)" et évocateur de simulation mentale (Muenzinger, 1938).

Ces observations ont initié l'hypothèse qu'un comportement pouvait être contrôlé par deux systèmes vus initialement comme mutuellement exclusifs, le système model-based, et le système model-free.

Le système model-based exploite un modèle interne chez le sujet, une connaissance du fonctionnement de l'environnement (et de la tâche dans un contexte expérimental) qui permet à l'individu de simuler mentalement les conséquences probables d'une action (Sutton et Barto, 1998). Il apprend en particulier les probabilités de transition entre états  $P(s, a, s')$  qui définit quel état  $s'$  suivra l'état  $s$  si le sujet choisit de faire l'action  $a$ . Ce modèle interne permet au sujet d'évaluer le bénéfice des actions futures, éventuellement en prenant en compte plusieurs pas de temps. Elle permet donc au sujet de s'adapter de façon flexible à des modifications du

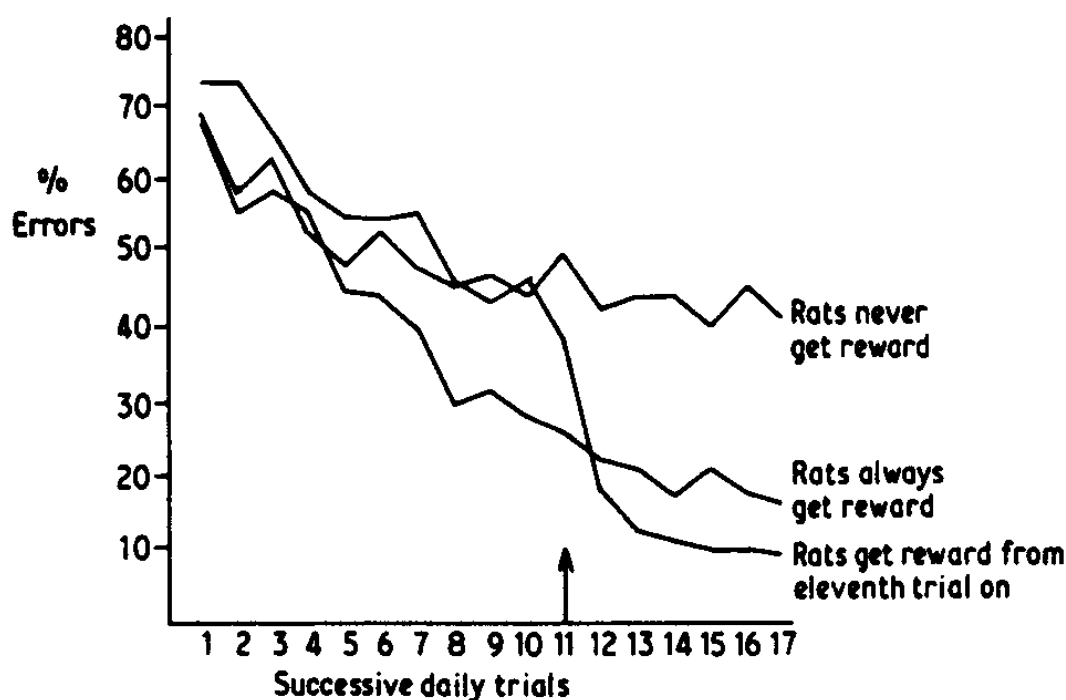


FIGURE 4.5 – Démonstration du phénomène d'apprentissage latent dans un labyrinthe complexe. Les rats recevant une récompense ont une diminution progressive du taux d'erreur bien supérieure à ce qui est observé chez des animaux ne recevant jamais de récompense. Cependant, l'introduction de récompenses à partir du 11ème jour, dans un groupe de rats n'ayant jusque là jamais eu de renforçateur et ayant un taux d'erreur élevé, amène une amélioration des performances quasi immédiate. Ce phénomène indique qu'un apprentissage avait en fait eu lieu durant la période où ces rats ne recevaient pourtant aucune récompense, bien que cela n'ait pas été visible dans leur comportement avant l'introduction du renforçateur. (Tolman et Honzik, 1930)

fonctionnement de l'environnement ou à une modification des outcome des actions (Lee et al., 2012).

Le système model-free à l'inverse ne prend des décision que sur la base des récompenses reçues dans le passé. Aucune représentation du fonctionnement de l'environnement n'est nécessaire. Ce type de contrôle du comportement a l'avantage d'être moins coûteux en termes de charge cognitive. Il a l'inconvénient d'être moins flexible dans son adaptation à un environnement qui se modifie de façon dynamique.

Daw et collaborateurs ont proposé une correspondance formelle entre ce découpage et la distinction entre comportement goal-directed et comportement habituel chez l'animal. Lorsque le comportement est goal-directed, l'animal est capable de modifier rapidement son comportement en réponse à des modifications dans les récompenses, tandis que lorsque le comportement est habituel, l'animal persévère dans ses actions, indépendamment des récompenses reçues. Cette définition est opérationnelle et rétrospective, car elle ne peut être faite qu'après une intervention d'extinction ou de dévaluation du renforçateur (Dickinson, 1985). Daw et collaborateurs ont ainsi proposé que le comportement goal-directed serait le résultat d'un apprentissage model-based et que le comportement habituel serait le résultat d'un apprentissage model-free (Daw et al., 2005).

Un consensus semble exister pour dire que le striatum dorso-létéral est nécessaire pour le comportement habituel, et le striatum dorso-médian est nécessaire pour le comportement goal-directed, sur la base d'études de lésion, ou d'inactivation pharmacologique chez l'animal (Balleine, 2005 ; Yin et Knowlton, 2006).

Les études récentes ont montré que les choix model-based incluent une activité prospective, faisant intervenir des régions préfrontales, notamment le cortex préfrontal dorso-médian et le cortex fronto-polaire (Doll et al., 2015), ou le cortex préfrontal ventro-médian (O'Doherty et al., 2015). Ces structures sous-tendraient en particulier les erreurs de prédiction en lien avec l'acquisition du modèle. Cependant la façon dont le modèle lui-même est représenté et stocké n'est pas encore élucidée (O'Doherty et al., 2015).

Cette distinction entre contrôle model-based et model-free du comportement est particulièrement intéressante pour la psychiatrie. En effet, une perturbation dans l'équilibre entre les systèmes de décision goal-directed et habituels pourraient être à l'origine de divers troubles dont certains sont du ressort de la psychiatrie, comme les troubles obsessionnels compulsifs (Gillan et al., 2011), les addictions (Belin et al., 2013), ou la schizophrénie (Dolan et Dayan, 2013).

## 4.4 Résumé et conclusion

L'apprentissage par renforcement décrit la façon dont un organisme améliore progressivement ses choix par essai-erreur dans l'interaction avec son environnement,

dans le but de maximiser la fréquence de conséquences plaisantes (récompenses) ou minimiser la fréquence de conséquences adverses (punitions). Ce mécanisme fondamental a fait l'objet d'une conceptualisation classique distinguant des stimuli (S) des actions (A) et des renforçateurs (O). Deux classes principales d'apprentissage ont été décrites : le conditionnement Pavlovien qui associe un stimulus (ou un état) à un renforçateur (O) et le conditionnement instrumental qui associe un stimulus (ou un état) à une action (A) par l'effet d'un renforçateur (O) déclenché par l'action.

Ces différents types d'apprentissage font l'objet d'une formalisation mathématique et algorithmique. Elle permet d'analyser le comportement des sujets au cours de tâches d'apprentissage par renforcement de façon fine. Les substrats cérébraux de l'apprentissage par renforcement ont été étudiés et ont attribué un rôle particulièrement important à la dopamine, qui encode, notamment dans le striatum, les erreurs de prédictions (différence entre le renforçateur attendu et le renforçateur effectivement reçu), qui sont un élément central et indispensable pour ce type d'apprentissage.

Des enrichissements sont venus compléter les modèles initiaux. L'apprentissage par renforcement par les punitions recherche si des circuits cérébraux spécifiques, notamment l'insula antérieure peut être impliquée dans ces situations, et si la sérotonine encode les erreurs de prédiction de punitions de façon symétrique et complémentaire à la dopamine. L'apprentissage contrefactuel analyse de quelle façon des sujets peuvent apprendre non seulement du résultat des actions choisies, mais également du résultat (hypothétique) des actions non choisies. L'apprentissage model-based rend compte de la façon dont un sujet construit un modèle interne du fonctionnement de son environnement et l'exploite pour choisir des actions de façon plus efficace que ce qu'il réussirait à faire sur la base des seuls renforçateurs reçus. Les bases cérébrales de ces deux derniers types d'apprentissage sont en cours d'exploration mais semblent impliquer de façon particulière les aires préfrontales.

L'apprentissage par renforcement permet donc d'analyser de façon fine le comportement d'un sujet au cours de son interaction avec l'environnement. Il mobilise des aires cérébrales qui ont un rôle probablement crucial dans bon nombre de pathologies psychiatriques, comme le striatum et le cortex préfrontal. Il est sous-tendu par le système dopaminergique avec un rôle possible du système sérotoninergique. L'étude de l'apprentissage par renforcement semble donc être une approche intéressante pour caractériser les modifications du fonctionnement cérébral du fait de pathologies psychiatriques ou pour caractériser l'effet des traitements psychotropes en termes comportementaux ou neurocognitifs et cérébraux. Nous illustrons cette possibilité dans les deux études rapportées ci-après.



Deuxième partie

Partie expérimentale



# Liste de publications et articles en cours

## Travaux directement liés à la thèse de sciences

Specific effect of a dopamine partial agonist on counterfactual learning : evidence from Gilles de la Tourette syndrome

**Salvador A**, Delorme C, Hartmann A, Coricelli G, Gaillard R, Worbe Y, Palminteri S.

En révision à *Scientific Reports*. [IF : 5.2].

Effects of acute and chronic antidepressant treatment on reinforcement learning and cognition : a 8 week randomized double blind placebo controlled study in healthy volunteers

**Salvador A**, Vinckier F, Pessiglione M, Deakin JFW, Harmer CJ, Montejo AL, Jabourian AP, Gruget C, Gabriel C, Goodwin GM, Gaillard R.

En cours de soumission à *NeuroPsychoPharmacology*. [IF : 6.4].

## Travaux connexe à la thèse de sciences

Enhanced habits formation in Gilles de la Tourette syndrome.

Delorme C, **Salvador A**, Valabrègue R, Roze E, Palminteri S, Vidailhet M, de Wit S, Robbins T, Hartmann A, Worbe Y.

*Brain* 2016 Feb ;139(Pt 2) :605-15. [IF : 10.1].

Illusion of agency in patients with Gilles de la Tourette Syndrome.

Delorme C, **Salvador A**, Voon V, Roze E, Vidailhet M, Hartmann A, Worbe Y. *Cortex*. 2016 Apr ;77 :132-40. [IF : 4.3]

Confidence and psychosis : a neuro-computational account of contingency learning disruption by NMDA blockade.

Vinckier F, Gaillard R, Palminteri S, Rigoux L, **Salvador A**, Fornito A, Adapa R, Krebs MO, Pessiglione M, Fletcher PC.

*Molecular Psychiatry*. 2015 Jun 9. [IF : 9.0].

Habitual and goal-directed behavioural control in Gilles de la Tourette syndrome. Worbe Y, **Salvador A**, Delorme C, Palminteri S, Roze E, Voon V, Vidailhet V, Daw N, Hartmann A.

En cours soumission à *Translational Psychiatry* [IF : 5.5]

## Autres travaux de recherche

Mast cells' involvement in inflammation pathways linked to depression : evidence in mastocytosis.

Georgin-Lavialle S, Moura DS, **Salvador A**, Chauvet-Gelinier JC, Launay JM, Damaj G, Côté F, Soucié E, Chandesris MO, Barète S, Grandpeix-Guyodo C, Bachmeyer C, Alyanakian MA, Aouba A, Lortholary O, Dubreuil P, Teyssier JR, Trojak B, Haffen E, Vandel P, Bonin B ; French Mast Cell Study Group, Hermine O, Gaillard R.

*Molecular Psychiatry*. 2016 Jan 26. [IF : 9.0].

Neuroimaging evidence of brain abnormalities in mastocytosis.

**Salvador A\***, Boddaert N\*, Chandesris MO, Lemaître H, Grévent D, Gauthier C, Naggara O, Georgin-Lavialle S, Moura DS, Munsch F, Jaafari N, Zilbovicius M, Lortholary O, Gaillard R, Hermine O. (\* = co-premiers auteurs)

En révision à **Translational Psychiatry** [IF : 5.5].

Unconscious memory suppression.

**Salvador A**, Vinckier F, Smadja S, Cohen L, Naccache L, Dehaene S, Gaillard R.

En review à *Cognition*. [IF : 3.4].

Everolimus induced mood changes in breast cancer patients : a case-control study.

O Mir, **A Salvador**, S Dauchy, S Ropert, S Delaloge, C Lemogne, R Gaillard.

En review à *European Journal of Cancer*. [IF : 6.2].

## Chapitre 5

### Article 1 : Specific effect of a dopamine partial agonist on counterfactual learning : evidence from Gilles de la Tourette syndrome

Le syndrome Gilles de la Tourette est caractérisé par des symptômes hyperkinétiques moteurs et verbaux. Ce syndrome est défini dans le DSM-V comme l'association de plusieurs tics moteurs et d'au moins un tic vocal à au moins un moment de la pathologie, mais pas nécessairement simultanément. La fréquence de ces manifestations peut varier, mais le tableau doit durer plus de un an depuis l'apparition des premiers tics. Le début des troubles doit être intervenu avant l'âge de 18 ans, et le tableau ne doit pas être dû à une substance (comme la cocaïne) ni à une pathologie médicale (comme la maladie de Huntington, ou une encéphalite post-virale) (American Psychiatric Association, 2013).

C'est un syndrome qui n'est pas rare puisque sa prévalence se situerait entre un et dix pour mille enfants (Bloch et Leckman, 2009 ; Scahill et al., 2005). L'étiologie en est mal connue, il existe en effet des preuves d'une transmission héréditaire, mais le substrat génétique n'est pas connu, et les facteurs environnementaux ont probablement un rôle important, mais ne sont pas non plus connus. Des hypothèses ont été avancées concernant le rôle d'une possible réaction auto-immune suivant une infection à streptocoques, mais n'est pas confirmée (Leckman, 2002).

La physiopathologie de ce trouble n'est pas encore bien caractérisée. Les symptômes ont été attribués à des particularités anatomiques et fonctionnelles des ganglions de la base (Albin et Mink, 2006 ; Graybiel et Canales, 2001). Une hyper-dopaminergie à la fois au niveau des ganglions de la base et au niveau du cortex frontal a été retrouvé par des études d'imagerie de type tomographie par émission de positons (bien

qu'elles ne soient pas unanimes) (Gilbert et al., 2006 ; Malison et al., 1995 ; Segura et Strafella, 2013 ; Wong et al., 2008), mais également par les études post-mortem (Minzer et al., 2004 ; Yoon et al., 2007).

Les symptômes sont notablement améliorés par l'administration d'antipsychotiques, qui sont le traitement de référence pour la prise en charge de ce trouble (Leckman, 2002 ; Singer, 2005). Leur mode d'action dans cette pathologie n'est elle non plus pas clairement établi.

Nous rapportons dans cette étude l'utilisation d'une tâche d'apprentissage par renforcement dans le but de caractériser les particularités du fonctionnement cérébral chez les patients présentant un syndrome Gilles de la Tourette comparés à des sujets sains et de caractériser l'effet d'un traitement antipsychotique fréquemment utilisé dans la prise en charge de ce syndrome, l'aripiprazole.

**Title:**

**Specific effect of a dopamine partial agonist on counterfactual learning: evidence from Gilles de la Tourette syndrome**

**Authors:**

Alexandre Salvador<sup>1,2</sup>, Yulia Worbe<sup>3</sup>, Cécile Delorme<sup>3</sup>, Giorgio Coricelli<sup>4</sup>, Raphaël Gaillard<sup>2</sup>, Trevor W. Robbins<sup>5</sup>, Andreas Hartmann<sup>3</sup>, Stefano Palminteri<sup>1,6</sup>

**Affiliations:**

<sup>1</sup>Laboratoire de Neurosciences Cognitives, Institut National de la Santé et de la Recherche Médicale, Paris, FR

<sup>2</sup>Centre de Psychiatrie et Neuroscience, Hôpital Sainte Anne, Paris, FR

<sup>3</sup>National Reference Centre for Gilles de la Tourette Syndrome, Department of Neurology, Institut du Cerveau et de la Moelle Epinière, Hôpital de la Pitié-Salpêtrière, Paris, FR

<sup>4</sup>Department of Economics, University of Southern California, USA

<sup>5</sup>Department of Psychology and Behavioural and Clinical Neuroscience Institute, University of Cambridge, UK

<sup>6</sup>Institut d'Étude de la Cognition, Département d'Études Cognitives, École Normale Supérieure, Paris, FR

**Corresponding author:**

Stefano Palminteri, PhD ([stefano.palminteri@ens.fr](mailto:stefano.palminteri@ens.fr))

**Funding and Disclosure**

AS was supported by a Fondation pour la Recherche Médicale grant (grant number FDM20120624489). SP was supported by a Marie Skłodowska-Curie Individual European Fellowship (PIEF-GA-2012 Grant 328822) and is supported by an ATIP-Avenir starting

grant (R16069JS). The study was supported by a grant from the Association Française du Syndrome de Gilles de la Tourette to YW. The funding agencies did not influence the content of the manuscript. The authors declare no competing financial interest. The Institut d'Etude de la Cognition is supported financially by the LabEx IEC (ANR-10-LABX-0087 IEC) and the IDEX PSL\* (ANR-10-IDEX-0001-02 PSL\*).

### **Author contributions**

AS performed the experiments, analysed the data, interpreted the results and wrote the manuscript. YW designed the study, performed the experiment, interpreted the results and edited the manuscript. CD performed the experiments and edited the manuscript. GC, RG, TWR interpreted the results and edited the manuscript. AH designed the study and interpreted the results. SP designed the study, analysed the data, interpreted the results and wrote the manuscript.

## **Abstract**

The dopamine partial agonist aripiprazole is increasingly used to treat pathologies for which other antipsychotics are indicated because it displays fewer side effects, such as sedation and depression-like symptoms, than other dopamine receptor antagonists. Previously, we showed that aripiprazole may protect motivational function by preserving reinforcement-related signals used to sustain reward-maximization behaviour in a simple action-outcome learning task. However, the effect of aripiprazole on more cognitive facets of human reinforcement learning, such as learning from the hypothetical outcomes of alternative courses of action (i.e., counterfactual learning), is unknown.

To test the influence of aripiprazole on counterfactual learning, we administered a reinforcement-learning task that involves both direct learning from obtained outcomes and indirect learning from forgone outcomes to two groups of Gilles de la Tourette (GTS) patients, one consisting of patients who were completely unmedicated and the other consisting of patients who were receiving aripiprazole monotherapy, and to healthy subjects. We replicated a previous finding showing that aripiprazole does not affect direct learning from obtained outcomes in GTS. We also found that whereas learning performance improved in the presence of counterfactual feedback in both healthy controls and unmedicated GTS patients, this was not the case in aripiprazole-medicated GTS patients.

Our results suggest that whereas aripiprazole preserves direct learning of action-outcome associations, it may impair more complex inferential processes, such as counterfactual learning, from forgone outcomes.

## Introduction

Aripiprazole is a recently introduced antipsychotic medication with dopamine receptor partial agonist mechanisms<sup>1</sup>. In schizophrenia patients, aripiprazole exhibited efficacy comparable to that of typical and atypical antipsychotics in the treatment of positive and negative symptoms as well as in the prevention of relapse<sup>2,3</sup>. Its efficacy has also been demonstrated in other neurological disorders for which antipsychotics are indicated, such as Gilles de la Tourette syndrome (GTS)<sup>4,5</sup>. Its tolerability is often considered superior to that of typical antipsychotics, and it is associated with fewer adverse side effects, such as extrapyramidal and metabolic syndromes<sup>6,7</sup>. In GTS, aripiprazole is effective for suppressing tics while displaying a less severe side effect profile than dopamine receptor antagonists with regard to motivational deficits such as sedation and depressive reactions<sup>8</sup>. Due to this advantageous cost-benefit trade-off, aripiprazole has become a widely prescribed treatment for schizophrenia and GTS.

Pharmacological studies in humans suggest that dopamine receptor antagonist-induced sedation and depressive states may be the consequence of blunting of reward-related signals<sup>9,10</sup>. For instance, reward-seeking behaviour and reward prediction errors encoded in the ventral striatum were reduced by haloperidol administration in healthy volunteers<sup>11</sup>. In contrast, both functions were preserved in GTS patients medicated with aripiprazole<sup>12</sup>. These results are consistent with the idea that, as opposed to dopamine receptor antagonists, aripiprazole preserves motivational functions via preserving reward-related signals.

Despite the observations described above, in humans, reinforcement learning is seldom solely based on learning from obtained outcomes (i.e., factual outcomes)<sup>13</sup>. In fact, human reinforcement learning often makes use of more abstract inferential processes, including learning from counterfactual outcomes<sup>14,15</sup>. In other words, counterfactual learning refers to the ability to learn not only from direct experience but also from hypothetical outcomes (the outcomes of the option that was not chosen). Previous neuroimaging studies have demonstrated that subjects take into account counterfactual feedback when it is available and have suggested that this form of learning could be underpinned by a dorsal prefrontal system<sup>16–18</sup>, despite the fact that some other studies have suggested that factual and counterfactual outcomes may be processed by the same neural system involving the ventral striatum<sup>19,20</sup>. On the pharmacological level, a recent study that investigated counterfactual learning under a regimen of dietary depletion of serotonin and

*Specific effect of dopamine partial agonist on counterfactual learning*

dopamine/noradrenalin precursors indicated that brain tracking of fictive prediction error might be sensitive to both dopamine and serotonin manipulation<sup>21</sup>.

Here, to characterize the effect of aripiprazole on counterfactual learning, we administered a recently developed behavioural task that contrasts learning from obtained and hypothetical outcomes to a group of medication-free GTS patients, a group of GTS patients with aripiprazole as their only medication, and matched healthy controls<sup>22</sup>. GTS patients represented an ideal model for this investigation because in this population we have already shown the beneficial effects of aripiprazole compared to a dopamine receptor antagonist in factual learning, that is, in learning from obtained outcomes (as opposed to counterfactual learning, which is learning from forgone outcomes)<sup>12</sup>. In addition, unlike many schizophrenic patients, GTS patients have relatively preserved cognitive efficiency<sup>23,24</sup>. Finally, by investigating the effect of aripiprazole in a population that is commonly treated with this medication, our findings have direct clinical relevance.

## Results

### Demographic and psychometric results

Forty consecutive patients consulting at the GTS Reference Center in Paris were included in the study. Patients taking any medication other than aripiprazole (i.e., pimozide, risperidone, haloperidol, bromazepam or escitalopram) were not included in the final analysis. One patient did not complete the task and was also excluded from the analysis. This led to a final sample of 31 GTS patients (17 without any medication and 14 with aripiprazole monotherapy). In addition, 20 healthy controls obtained through advertising were included. Two significant differences were found when comparing the 3 groups: the number of years of study ( $F(2,48)=3.2$ ,  $p=0.050$ ) (although the results of the National Adult Reading Test (NART;  $F(2,46)=1.2$ ,  $p=0.31$ ), which measures verbal intelligence, did not differ among the groups) and the Beck Depression Inventory score (BDI,  $F(2,48)=6.72$ ,  $p=0.0027$ ). In addition, two scales showed marginal differences among the three groups: the Barratt Impulsiveness Scale (BIS-11;  $F(2,48)=3.00$ ,  $p=0.059$ ) and the Behavioural Inhibition / Behavioural Activation Scale (BIS/BAS;  $F(2,48)=3.04$ ,  $p=0.057$ ). However, there was no significant difference between the aripiprazole and unmedicated GTS groups on any demographic or psychometric scale or in disease severity index or measures of intelligence (all  $p>0.2$ , two-sample t-test).

### Behavioural results

Subjects performed a probabilistic instrumental learning task adapted from previous imaging studies<sup>22,25</sup>. During the instrumental probabilistic learning task, the subjects were required to learn which stimuli had the greatest likelihood of resulting in an advantageous outcome through trial and error. The dependent variable was the subject's accuracy; an "accurate" (or "correct") choice was choice of the stimulus that is on average more rewarding or less punishing. Both feedback valence (reward vs. punishment) and feedback information (partial feedback vs. complete feedback) were manipulated using a within-subjects factorial design (Figure 1A).

A preliminary analysis of variance (ANOVA) evaluating accuracy as a function of trial, feedback information, feedback valence and group (aripiprazole-treated GTS patients, unmedicated GTS patients, and controls) showed that there was a significant main effect of group ( $F(2,4864)=14.5$ ,  $p<0.001$ ), a significant main effect of trial number ( $F(1,4864)=46.4$ ,  $p<0.001$ ), and a significant main effect of feedback information ( $F(1,4864)=23.7$ ,  $p<0.001$ ), with an interaction effect of group with feedback information

*Specific effect of dopamine partial agonist on counterfactual learning*

( $F(2,4864)=15.35$ ,  $p<0.001$ ) but no main effect of feedback valence ( $F(1,4864)=0.8$ ,  $p>0.3$ ). Feedback valence was therefore excluded from the subsequent ANOVA analysis. The final ANOVA, which evaluated accuracy as a function of trial, feedback information, and group, showed that there was a significant main effect of trial number on correct choice rate ( $F(1,2432)=41.1$ ,  $p<0.001$ ), indicating that accuracy improved with learning (Figure 2A). In addition, there was a significant main effect of feedback information ( $F(1,2432)=21.0$ ,  $p<0.001$ ) and of group ( $F(2,2432)=12.9$ ,  $p<0.001$ ), with a significant interaction between group and feedback information ( $F(2,2432)=13.6$ ,  $p<0.001$ ), indicating that counterfactual feedback did not positively modulate the accuracy similarly in the three groups. The interaction effect of group and feedback information remained significant when years of study, BDI score, and BIS.11 or BIS.BAS impulsivity scores were included either individually or all together ( $F(2,2260)=22.8$ ,  $p<0.001$ ).

We then defined the « counterfactual gain » as the difference in accuracy between the complete feedback and the partial feedback conditions. A confirmatory ANOVA evaluating counterfactual gain as a function of group and trial showed a significant main effect of group ( $F(2,1216)=17.9$ ,  $p<0.001$ ), indicating that the groups benefited differently from counterfactual information. There was no main effect of trial ( $F(1,1216)=1.00$ ,  $p=0.3$ ), indicating that the benefit from counterfactual information was present for the entire duration of the experiment. The effect of group remained significant when years of study, BDI score, and BIS.11 or BIS.BAS impulsivity scores were included either individually or all together ( $F(2, 1130)=25.1$ ,  $p<0.001$ ).

To assess the directionality of the critical ANOVA effects in a concise manner, we plotted and analysed the accuracy and the counterfactual gain averaged across trials. Accuracy in the complete feedback condition was significantly higher than accuracy in the partial feedback condition in both the control and unmedicated GTS groups ( $t(19)=-2.6$ ,  $p<0.05$ , and  $t(16)=-2.3$ ,  $p<0.05$ , respectively) but not in the aripiprazole GTS group ( $t(13)=0.151$ ,  $p=0.882$ ) (Figure 2B). The average counterfactual gain was significantly higher in the control group and in the unmedicated group than in the aripiprazole group (8.3% controls versus 0.3% aripiprazole,  $t(31.5)=2.2$ ,  $p<0.05$ ; t-test 10.5% unmedicated versus 0.3% aripiprazole,  $t(23.3)=2.1$ ,  $p<0.05$ ). There was no difference in counterfactual gain between the control and unmedicated groups ( $t(29.3)=-0.4$ ,  $p=0.701$ ) (Figure 2C).

Interestingly, in the aripiprazole GTS group, we found a significant correlation between treatment dose and accuracy in the complete feedback condition (Spearman's correlation

*Specific effect of dopamine partial agonist on counterfactual learning*

test  $\rho=-0.66$ ,  $p<0.01$ ), indicating that the higher the dose of atypical antipsychotic medication received, the lower the accuracy in the complete feedback condition. This correlation remained significant when the only patient with a 30-mg dose of aripiprazole was excluded ( $\rho=-0.57$ ,  $p<0.05$ ). In contrast, treatment dose did not correlate with accuracy in the partial feedback condition (Spearman's correlation test  $\rho=-0.47$ ,  $p>0.05$ ).

To assess evidence in favour of the null hypothesis (i.e., the absence of counterfactual learning in the aripiprazole group), we performed a Bayesian equivalent of a t-test (see the Materials and Methods section). This analysis showed that the probability that the average counterfactual gain falls within a region of practical equivalence (i.e., ROPE: [-1.5%, 1.5%] interval) is small (4%) in both the unmedicated GTS group and in the control group, whereas this probability is 5.75 times larger (23%) in the aripiprazole group (Figure 4). This Bayesian approach provides evidence that the absence of counterfactual learning in the aripiprazole group does not arise from a lack of power in detecting it.

Finally, we analysed the reaction times using an ANOVA with trial, group and feedback information as factors (see Table 2). The ANOVA revealed a significant effect of trial, indicating that reaction time decreased across trials ( $F(1,2432)=17.8$ ,  $p<0.001$ ). Feedback information and group did not affect reaction time ( $F(1,2432)=0.08$ ,  $p=0.772$  and  $F(2,2432)=2.74$ ,  $p=0.064$ ), nor did any of the interaction factors, indicating that reaction time was not modulated by the pathology or by the treatment.

## Discussion

Using a probabilistic learning paradigm with factual (where subjects were informed only about the result of the chosen outcome) and counterfactual (where subjects were also informed about the result of the unchosen outcome) feedback information conditions in healthy controls and aripiprazole-medicated and unmedicated GTS patients, we showed a specific effect of aripiprazole on counterfactual learning in the absence of any effect on factual learning.

All subjects performed significantly above chance level in the partial feedback condition, and accuracy did not differ across the groups. As previously shown, control subjects had higher accuracy in the complete feedback condition than in the partial feedback condition<sup>22,26</sup>. However, in GTS patients, counterfactual learning was affected by medication status: unmedicated patients behaved similarly to healthy controls, whereas aripiprazole-medicated patients did not improve their performance in the presence of counterfactual feedback. Counterfactual learning, as implemented in our task, requires at least two component processes: a “peripheral” process, which corresponds to the fact that the complete feedback condition mobilizes additional attentional resources; and a “central” process, which corresponds to the integration of hypothetical prospects (“counterfactuals”) in reinforcement learning<sup>27</sup>.

Was the effect of aripiprazole on counterfactual learning the hallmark of a specific impairment in counterfactual reasoning or a consequence of reduced attention? Evidence against a nonspecific “peripheral” effect is the fact that baseline performance (as measured in the partial feedback condition) did not differ in the unmedicated and the aripiprazole GTS groups. If anything, average accuracy in the partial feedback condition was numerically higher in the aripiprazole GTS group than in the unmedicated GTS group (67.5% vs. 63.0%). Further evidence against the idea of a general attentional decrease is our finding that reaction times did not differ among the groups, indicating a similar level of engagement in the task for the subjects in all groups.

The preserved ability to learn from factual outcomes during aripiprazole treatment replicates previous work by Worbe and colleagues showing that aripiprazole did not affect probabilistic reward learning in GTS patients<sup>12</sup>. This characteristic of aripiprazole strikingly contrasts with the robust observation that dopamine antagonists blunt reward-related learning and performance<sup>9–11</sup>. Reduced counterfactual learning has been reported in nicotine addiction<sup>28</sup>. It may therefore be argued that this learning deficit may contribute to

*Specific effect of dopamine partial agonist on counterfactual learning*

the increased likelihood of behavioural addictions, such as pathological gambling, in aripiprazole-medicated patients<sup>29</sup> as well as the observation that aripiprazole increases cigarette smoking<sup>30</sup>.

The reduction in counterfactual learning in the aripiprazole group may come as a surprise because early studies proposed that, given its particular mechanism of action as a partial D2 agonist, aripiprazole would preserve or even improve high-level cognitive function through improved frontal dopamine release<sup>31,32</sup>. However, later studies mitigated this expectation by showing no effect or even a deterioration in executive functions following aripiprazole administration<sup>33–39</sup>. Consistent with those findings, imaging studies failed to demonstrate an increase in frontal function and even showed reduced frontal metabolism associated with reduced performance in executive function tasks<sup>40–42</sup>. These contrasting findings on prefrontal dopamine might be linked to aripiprazole dose; low-dose systemic aripiprazole (0.5 mg/kg) was shown to increase extracellular dopamine levels in the cortex, whereas high-dose aripiprazole (10 mg/kg) reduced dopamine levels in mice<sup>43,44</sup>. This might explain why we find a detrimental effect on counterfactual learning that increases significantly with aripiprazole dose in our study.

Aripiprazole has binding affinity for both dopamine and serotonin receptors<sup>1</sup>. Its affinity for 5-HT1A receptors in particular may provide an explanation for the apparent discrepancy between preserved factual learning and reduced counterfactual learning. Dopamine and serotonin have been associated with both automatic/reflexive and deliberative/reflective reinforcement processing, which in our task could be incorporated into factual and counterfactual learning<sup>45–48</sup>. In particular, it has been shown that dietary depletion of tryptophan reduces the propensity to implement cognitive strategies in a reinforcement-learning task<sup>46,49</sup>, and aripiprazole has been shown to reduce serotonin output in the prefrontal cortex and the dorsal raphe nucleus<sup>50</sup>. In the light of these findings, it may be argued that aripiprazole's detrimental effect on counterfactual learning is mediated by an alteration in serotonergic neurotransmission. Aripiprazole would therefore have a specific pattern of effects on reinforcement learning that includes both preserved direct/factual learning due to its dopamine partial agonist property and reduced indirect/counterfactual learning due to its prefrontal serotonin disruption effect. Future studies in which specific dopaminergic and serotonergic pharmacological agents are used are required to determine whether the effect of aripiprazole on counterfactual learning is mediated by its dopaminergic affinity or its serotonergic affinity or by a combination of the two.

*Specific effect of dopamine partial agonist on counterfactual learning*

Is the detrimental effect of aripiprazole on counterfactual learning specific to GTS patients or generalizable? The answer to this question is closely related to the mechanism through which aripiprazole blunts counterfactual learning. GTS has been hypothesized to be dependent on a constitutive hyperactivity of dopaminergic neurotransmission<sup>51,52</sup>. Positron emission tomography and post-mortem studies found evidence of increased dopamine concentration in both the striatum and the prefrontal cortex of GTS patients<sup>53–57</sup>. Thus, if the negative effect of aripiprazole on counterfactual learning is mediated by the dopamine system, the alterations of this neuromodulator that occur in GTS may preclude the generalizability of our result to other pathologies and to the general population. On the other hand, there is no current clear hypothesis and no replicated data implicating the serotonergic system in GTS, with recent negative results<sup>58,59</sup>. Thus, if the negative effect of aripiprazole on counterfactual learning is mediated by the serotonin system, it is less likely that it is specific to GTS patients.

Previous studies suggest that hyperactive reinforcement learning may contribute significantly to GTS symptoms<sup>60,61</sup>. Assuming that positive and negative prediction errors are represented by phasic fluctuations in dopaminergic transmission, we previously found that reinforcement learning in unmedicated GTS subjects was consistent with a functional hyperdopaminergia, both in terms of enhanced reward<sup>62</sup> and reduced punishment learning<sup>63</sup>. These findings were obtained when reward-outcome contingencies were deterministic and unconscious as measured by post-test assessments of cue visibility and stimulus-reward associations and using a task that has been shown to rely specifically on the ventral striatum<sup>64</sup>. In contrast, our current results were obtained using a probabilistic instrumental learning task that employs perfectly visible (therefore conscious) cues and recruits both the cortex (namely, the medial prefrontal cortices and the insula) and the ventral striatum<sup>22</sup>. In the present study, we found that unmedicated GTS patients performed equally well in the reward-seeking condition and in the punishment-avoidance condition. Taken together, our current and previous findings might suggest that in GTS excessive positive reinforcement occurs for implicit/unconscious (as opposed to explicit/conscious) learning processes; this possibility could hypothetically be supported by the dependence of these two types of processes on different neural substrates (subcortical vs. prefrontal)<sup>65</sup>.

A potential limitation of our study is that it was observational, meaning that we included patients who were already either unmedicated or undergoing treatment with aripiprazole.

*Specific effect of dopamine partial agonist on counterfactual learning*

The absence of group randomization raises the question that the effects attributed to aripiprazole may arise from pre-existing phenotypic between-group differences that determined the patients' medication status. This is unlikely for several reasons. First, very few patients are either never or always medicated; thus, medication status fluctuates and does not seem to be a patient-specific characteristic<sup>8</sup>. Second, to control as much as possible for this potential confounding factor, we measured several clinical, psychological and cognitive parameters and compared them between groups. We ensured through a psychiatric interview that the included GTS subjects did not have comorbid psychiatric diagnoses. We further measured a number of dimensions that could account for learning differences in our task, including general intelligence, disease severity, and cognitive and clinical dimensions, such as impulsivity, anxiety and depression, through rating scales. None of these dimensions differed significantly between aripiprazole-treated and unmedicated GTS patients; it is therefore unlikely that they could account for the reduction in counterfactual learning found in the aripiprazole group. Two of these dimensions, depression (BDI) and level of study, differed significantly between the control group and the two GTS groups, and impulsivity differed marginally. This, of course, does not affect our main claim regarding a difference between unmedicated and aripiprazole GTS patients. However, as a formal control, we included these dimensions as cofactors in the ANOVA analysis and found that the reported effects remained significant. To summarize, while the absence of group randomization is a methodological limitation that requires us to be cautious in concluding that the difference in counterfactual learning across groups is a specific result of treatment, we controlled for a number of potential confounding factors to make this conclusion as robust as possible.

Another potential limitation of our study is its relatively small sample size. While the number of participants analysed in each group is consistent with previous publications targeting reinforcement learning in GTS patients<sup>12,62,63</sup>, one could wonder whether the absence of difference in accuracy between the complete feedback and partial feedback conditions in the aripiprazole group might be due to a lack of power. However, the Bayesian analysis provided evidence that the absence of a difference in accuracy between the partial and complete feedback conditions specifically in the aripiprazole GTS group did not arise as the result of a lack of power.

To conclude, we demonstrated a negative effect of aripiprazole on counterfactual learning in a population of aripiprazole-medicated GTS patients. At present, it is unclear whether

*Specific effect of dopamine partial agonist on counterfactual learning*

this finding can be generalized to other clinical populations and, most notably, to schizophrenic patients, for whom treatment with aripiprazole is becoming increasingly popular. It is also unclear at this stage whether the observed effect is mediated by the dopaminergic system or the serotonergic system. Further research is needed to address these questions. Nonetheless, our findings shed new light on the characterization of the cognitive profile of aripiprazole and represent a first step in understanding the neuropharmacology of counterfactual learning.

## **Materials and methods**

### **Participants**

The ethics committee of the Pitié-Salpêtrière Hospital approved the study (N° Promoteur: C11-34; N° CPP: 97\*19; N° ANSM: B121209-31). All participants gave written informed consent prior to participation in the study, and the study was conducted in accordance with the Declaration of Helsinki (1964).

Forty consecutive patients of the GTS Reference Center in Pitié-Salpêtrière Paris were screened for inclusion and offered participation in this study. They were examined by a multidisciplinary team experienced in GTS (AH and YW). Tic severity was rated on the Yale Global Tic Severity Scale<sup>66</sup>. Inclusion criteria for the patients were as follows: age above 18 and confirmed diagnosis of GTS fulfilling the DSM-5 criteria. Aripiprazole-medicated GTS patients were on stable antipsychotic treatment for at least 4 weeks. Exclusion criteria included the co-occurrence of Axis I psychiatric disorders, including current major depressive disorder, current anxiety disorder, current or past psychotic disorder, substance abuse aside from nicotine, attention deficit disorder and obsessive-compulsive disorders. These pathologies were screened using both subsections of the Mini International Neuropsychiatry Inventory (MINI) and a clinical evaluation performed by a trained neurologist who investigated the presence of diagnostic criteria for OCD or ADHD as detailed in the DSM-IV TR. Exclusion criteria also included any neurological or movement disorder other than tics as evaluated by a qualified neurologist before inclusion, as well as the use of any medication (including antidepressants and benzodiazepines) in the unmedicated GTS group and the presence of any co-medication in the aripiprazole GTS group.

Healthy volunteers were recruited by hospital-based advertisements. The inclusion criterion for healthy volunteers was age over 18 years. The exclusion criteria were the same as for TS patients plus a personal history of tics and any concomitant treatment except contraceptive pills for women.

### **Behavioural task**

The subjects performed a probabilistic instrumental learning task adapted from previous imaging studies<sup>22,25</sup>. The subjects were first provided with written instructions, and these were reformulated orally if needed. Each participant first performed a short training session (16 trials) aimed at familiarizing the participant with the task's timing and responses. The participant then performed three learning sessions.

### *Specific effect of dopamine partial agonist on counterfactual learning*

The options were abstract symbols taken from the Agathodaimon font. Each session contained eight novel options grouped into four novel fixed pairs. The pairs of options were fixed so that a given option was always presented with the same other option. Thus, within each session, pairs of options represented stable choice contexts. Within sessions, each pair of options was presented 24 times for a total of 96 trials. The four option pairs corresponded to the four conditions (reward/partial, reward/complete, punishment/partial and punishment/complete), and the four conditions were associated with different pairs of outcomes (reward contexts: winning 0.5 € versus nothing; punishment contexts: losing 0.5 € versus nothing) and the provision of a different amount of information at feedback (partial or complete). In the partial feedback contexts, only the outcome regarding the chosen option was provided, whereas in the complete feedback contexts the outcomes of both the chosen and the unchosen option were provided (Figure 1A). Within each pair, the two options were associated with the two possible outcomes with reciprocal probabilities (0.75/0.25 and 0.25/0.75). The subjects were informed that the aim of the task was to maximize their payoff and that only factual (and not counterfactual) outcomes counted. Between cues, the outcome probability was independent on a trial-by-trial basis, although it was anti-correlated on average. Thus, in a complete feedback trial, the subjects could observe the same outcome from both cues on 37.5% of trials and different outcomes from each cue on 62.5% of trials.

Pairs of options were presented in a pseudo-randomized and unpredictable manner to each subject (intermixed design). During each trial, one option was randomly presented on the left and one on the right side of a central fixation cross (Figure 1B). The side on which a given option was presented was also pseudo-randomized such that a given option was presented an equal number of times on the left and on the right of the central cross.

### **Statistical analysis**

From the learning task, we extracted accuracy. An “accurate” (or “correct”) choice occurs when the subject selects the most rewarding symbol in the reward-seeking context (i.e., the symbols that yielded +0.5 € in 75% of cases and 0 € in 25% of cases) or selects the less punishing symbol in the punishment-avoidance context (i.e., the symbols that yielded 0 € in 75% of cases and -0.5 € in 25%). The central hypothesis regarding the effect of aripiprazole was assessed using a repeated measures ANOVA with group (« Controls », « Unmedicated », and « Aripiprazole ») as the between-subjects factor and trial (1:24), feedback valence (reward vs. punishment avoidance) and feedback information (partial

*Specific effect of dopamine partial agonist on counterfactual learning*

feedback vs. complete feedback) as the within-subject factors. A specific effect of aripiprazole on counterfactual learning would result in a significant group x feedback information interaction. As a control, we performed an ANCOVA analysis that included additional psychometric measures (education, BDI, BIS.11 and BIS.BAS) as covariates. As an additional control, we performed the same ANOVA on reaction times. We also performed a confirmatory ANOVA on the “counterfactual gain”, which was defined as the difference in the accuracy achieved in the complete and partial feedback contexts (a positive “counterfactual gain” indicates a beneficial effect of counterfactual feedback on performance).

To assess the directionality of the effect, we reported the average values of all dependent variables as well as the results of Student’s t-tests (two sample t-tests were used when comparing groups, and one sample t-test was used when comparing group-values to zero). A Welch-Satterthwaite approximation was used when the variance was not equal across the groups<sup>67</sup>. Correlations, in particular the correlation between treatment dose and accuracy, were computed using Spearman's test. The significance level was set at 5%.

To assess evidence in favour of the null hypothesis (i.e., no difference in accuracy in the complete and partial feedback conditions), we performed a Bayesian equivalent of a t-test using Monte Carlo Markov chain analysis to compute the posterior probability of the difference in accuracy between the two feedback information conditions for each group. This analysis was conducted using the Bayesian Estimation Supersedes the t-Test (BEST) package version 0.4.0 in R statistical software with the BESTmcmc and plot.BEST functions. The region of practical equivalence (ROPE) was set at 3%, meaning that we considered a difference included in the [-1.5%, 1.5%] interval to be too small to be considered relevant. Evidence for the null hypothesis was the probability that the actual difference between feedback information conditions lies in the ROPE for each group. This probability was assessed through Monte Carlo Markov chain analysis; 100 000 simulations were run for each test.

All statistical analyses were performed using R statistical software.

## Bibliography

1. Potkin, S. G. et al. Aripiprazole, an antipsychotic with a novel mechanism of action, and risperidone vs placebo in patients with schizophrenia and schizoaffective disorder. *Arch Gen Psychiatry* 60, 681–690 (2003).
2. Pigott, T. A. et al. Aripiprazole for the prevention of relapse in stabilized patients with chronic schizophrenia: a placebo-controlled 26-week study. *J Clin Psychiatry* 64, 1048–1056 (2003).
3. Kane, J. M. et al. Efficacy and safety of aripiprazole and haloperidol versus placebo in patients with schizophrenia and schizoaffective disorder. *J Clin Psychiatry* 63, 763–771 (2002).
4. Kawohl, W., Schneider, F., Vernaleken, I. & Neuner, I. Aripiprazole in the pharmacotherapy of Gilles de la Tourette syndrome in adult patients. *The world journal of biological psychiatry* 10, 827–831 (2009).
5. Neuner, I., Nordt, C., Schneider, F. & Kawohl, W. Effectiveness of aripiprazole in the treatment of adult Tourette patients up to 56 months. *Hum Psychopharmacol* 27, 364–369 (2012).
6. Citrome, L., Kalsekar, I., Baker, R. A. & Hebden, T. A review of real-world data on the effects of aripiprazole on weight and metabolic outcomes in adults. *Curr Med Res Opin* 30, 1629–1641 (2014).
7. Stroup, T. S. et al. A randomized trial examining the effectiveness of switching from olanzapine, quetiapine, or risperidone to aripiprazole to reduce metabolic risk: comparison of antipsychotics for metabolic problems (CAMP). *Am J Psychiatry* 168, 947–956 (2011).
8. Hartmann, A. & Worbe, Y. Pharmacological treatment of Gilles de la Tourette syndrome. *Neurosci Biobehav Rev* 37, 1157–1161 (2013).
9. Frank, M. J., Seeberger, L. C. & O'reilly, R. C. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306, 1940–1943 (2004).
10. Eisenegger, C. et al. Role of dopamine D2 receptors in human reinforcement learning. *Neuropsychopharmacology* 39, 2366–2375 (2014).
11. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045 (2006).
12. Worbe, Y. et al. Reinforcement learning and Gilles de la Tourette syndrome: dissociation of clinical phenotypes and pharmacological treatments. *Arch Gen Psychiatry* 68, 1257–1266 (2011).
13. O'Doherty, J. P., Lee, S. W. & McNamee, D. The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences* 1, 94–100 (2015).
14. Coricelli, G. & Rustichini, A. Counterfactual thinking and emotions: regret and envy learning. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 365, 241–247 (2010).
15. Coricelli, G., Dolan, R. J. & Sirigu, A. Brain, emotion and decision making: the paradigmatic example of regret. *Trends in cognitive sciences* 11, 258–265 (2007).
16. Boorman, E. D., Behrens, T. E. & Rushworth, M. F. Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS biology* 9, e1001093 (2011).
17. Fischer, A. G. & Ullsperger, M. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron* 79, 1243–1255 (2013).
18. Koechlin, E. An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 369, (2014).

*Specific effect of dopamine partial agonist on counterfactual learning*

19. Kishida, K. T. et al. Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proceedings of the National Academy of Sciences of the United States of America* 113, 200–205 (2016).
20. Li, J. & Daw, N. D. Signals in human striatum are appropriate for policy update rather than value prediction. *The Journal of neuroscience* 31, 5504–5511 (2011).
21. Tobia, M. J. et al. Neural systems for choice and valuation with counterfactual learning signals. *NeuroImage* 89, 57–69 (2014).
22. Palminteri, S., Khamassi, M., Joffily, M. & Coricelli, G. Contextual modulation of value signals in reward and punishment learning. *Nat Commun* 6, 8096 (2015).
23. Leckman, J. F. Tourette's syndrome. *Lancet* 360, 1577–1586 (2002).
24. Singer, H. S. Tourette's syndrome: from behaviour to biology. *Lancet Neurol* 4, 149–159 (2005).
25. Palminteri, S., Boraud, T., Lafargue, G., Dubois, B. & Pessiglione, M. Brain hemispheres selectively track the expected value of contralateral options. *J Neurosci* 29, 13465–13472 (2009).
26. Palminteri, S., Kilford, E. J., Coricelli, G. & Blakemore, S.-J. The Computational Development of Reinforcement Learning during Adolescence. *PLoS computational biology* 12, e1004953 (2016).
27. Baird, A. A. & Fugelsang, J. A. The emergence of consequential thought: evidence from neuroscience. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 359, 1797–1804 (2004).
28. Chiu, P. H., Lohrenz, T. M. & Montague, P. R. Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nature neuroscience* 11, 514–520 (2008).
29. Moore, T. J., Glenmullen, J. & Mattison, D. R. Reports of pathological gambling, hypersexuality, and compulsive shopping associated with dopamine receptor agonist drugs. *JAMA internal medicine* 174, 1930–1933 (2014).
30. Lofwall, M. R., Nuzzo, P. A., Campbell, C. & Walsh, S. L. Aripiprazole effects on self-administration and pharmacodynamics of intravenous cocaine and cigarette smoking in humans. *Experimental and clinical psychopharmacology* 22, 238–247 (2014).
31. Jordan, S. et al. The antipsychotic aripiprazole is a potent, partial agonist at the human 5-HT<sub>1A</sub> receptor. *European journal of pharmacology* 441, 137–140 (2002).
32. Tamminga, C. A. & Carlsson, A. Partial dopamine agonists and dopaminergic stabilizers, in the treatment of psychosis. *Current drug targets CNS and neurological disorders* 1, 141–147 (2002).
33. Yasui-Furukori, N., Kaneda, A., Sugawara, N., Tomita, T. & Kaneko, S. Effect of adjunctive treatment with aripiprazole to atypical antipsychotics on cognitive function in schizophrenia patients. *J Psychopharmacol* 26, 806–812 (2012).
34. Schlagenhauf, F. et al. Switching schizophrenia patients from typical neuroleptics to aripiprazole: effects on working memory dependent functional activation. *Schizophr Res* 118, 189–200 (2010).
35. Suzuki, H., Gen, K. & Inoue, Y. An unblinded comparison of the clinical and cognitive effects of switching from first-generation antipsychotics to aripiprazole, perospirone or olanzapine in patients with chronic schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 35, 161–168 (2011).
36. Wang, C., Zheng, D., Xu, J., Lam, W. & Yew, D. T. Brain damages in ketamine addicts as revealed by magnetic resonance imaging. *Front Neuroanat* 7, 23 (2013).

*Specific effect of dopamine partial agonist on counterfactual learning*

37. Lee, B. J. et al. Effect of aripiprazole on cognitive function and hyperprolactinemia in patients with schizophrenia treated with risperidone. *Clin Psychopharmacol Neurosci* 11, 60–66 (2013).
38. Maat, A. et al. Open, randomized trial of the effects of aripiprazole versus risperidone on social cognition in schizophrenia. *Eur Neuropsychopharmacol* 24, 575–584 (2014).
39. Yeh, C.-B. et al. Neurocognitive effects of aripiprazole in adolescents and young adults with schizophrenia. *Nord J Psychiatry* 68, 219–224 (2014).
40. Kim, E. et al. The relationship between antipsychotic D2 occupancy and change in frontal metabolism and working memory : A dual [(11)C]raclopride and [(18)F]FDG imaging study with aripiprazole. *Psychopharmacology* 227, 221–229 (2013).
41. Handley, R. et al. Acute effects of single-dose aripiprazole and haloperidol on resting cerebral blood flow (rCBF) in the human brain. *Hum Brain Mapp* 34, 272–282 (2013).
42. Bolstad, I. et al. No difference in frontal cortical activity during an executive functioning task after acute doses of aripiprazole and haloperidol. *Frontiers in human neuroscience* 9, 296 (2015).
43. Zocchi A, Fabbri D, Heidbreder CA (2005). Aripiprazole increases dopamine but not noradrenaline and serotonin levels in the mouse prefrontal cortex. *Neuroscience letters* **387**: 157–161.
44. Tanahashi S, Yamamura S, Nakagawa M, Motomura E, Okada M (2012). Dopamine D2 and serotonin 5-HT1A receptors mediate the actions of aripiprazole in mesocortical and mesoaccumbens transmission. *Neuropharmacology* **62**: 765–774.
45. Wunderlich, K., Smittenaar, P. & Dolan, R. J. Dopamine enhances model-based over model-free choice behavior. *Neuron* 75, 418–424 (2012).
46. Worbe, Y. et al. Valence-dependent influence of serotonin depletion on model-based choice strategy. *Molecular psychiatry* 21, 624–629 (2016).
47. Doll, B. B., Bath, K. G., Daw, N. D. & Frank, M. J. Variability in Dopamine Genes Dissociates Model-Based and Model-Free Reinforcement Learning. *The Journal of neuroscience* 36, 1211–1222 (2016).
48. Sharp, M. E., Foerde, K., Daw, N. D. & Shohamy, D. Dopamine selectively remediates 'model-based' reward learning: a computational approach. *Brain* 139, 355–364 (2016).
49. Tanaka, S. C. et al. Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. *PloS one* 2, e1333 (2007).
50. Bortolozzi A, Díaz-Mataix L, Toth M, Celada P, Artigas F (2007). In vivo actions of aripiprazole on serotonergic and dopaminergic systems in rodent brain. *Psychopharmacology* **191**: 745–758.
51. Graybiel, A. M. & Canales, J. J. The neurobiology of repetitive behaviors: clues to the neurobiology of Tourette syndrome. *Adv Neurol* 85, 123–131 (2001).
52. Albin, R. L. & Mink, J. W. Recent advances in Tourette syndrome research. *Trends Neurosci* 29, 175–182 (2006).
53. Malison, R. T. et al. [123I]beta-CIT SPECT imaging of striatal dopamine transporter binding in Tourette's disorder. *Am J Psychiatry* 152, 1359–1361 (1995).
54. Gilbert, D. L. et al. Altered mesolimbocortical and thalamic dopamine in Tourette syndrome. *Neurology* 67, 1695–1697 (2006).

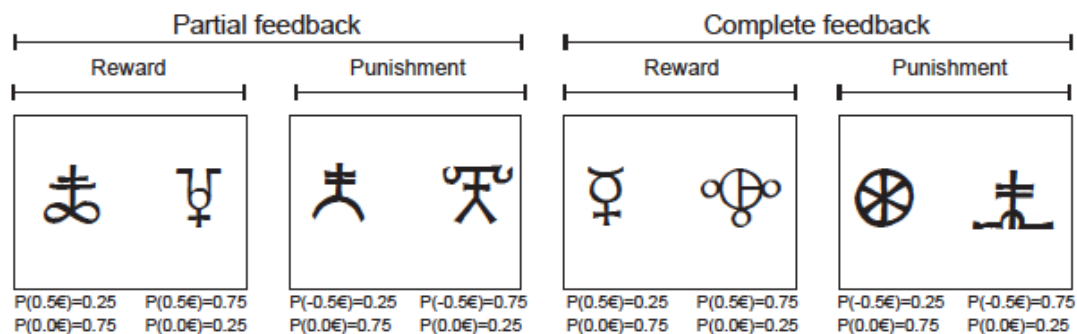
*Specific effect of dopamine partial agonist on counterfactual learning*

55. Wong, D. F. et al. Mechanisms of dopaminergic and serotonergic neurotransmission in Tourette syndrome: clues from an in vivo neurochemistry study with PET. *Neuropsychopharmacology* 33, 1239–1251 (2008).
56. Minzer, K., Lee, O., Hong, J. J. & Singer, H. S. Increased prefrontal D2 protein in Tourette syndrome: a postmortem analysis of frontal cortex and striatum. *J Neurol Sci* 219, 55–61 (2004).
57. Yoon, D. Y., Gause, C. D., Leckman, J. F. & Singer, H. S. Frontal dopaminergic abnormality in Tourette syndrome: a postmortem analysis. *J Neurol Sci* 255, 50–56 (2007).
58. Xu, L. et al. Lack of genetic association of 5-HT2A 102 T/C and -1438A/G polymorphisms with Tourette syndrome in a family-based association study in a Chinese Han population. *Asia-Pacific psychiatry* 8, 87–91 (2016).
59. Yi, M., Zhang, Y., Wang, Y., Su, N. & Liu, S. Association between the polymorphism of C861G (rs6296) in the serotonin 1B receptor gene and Tourette syndrome in Han Chinese people. *Asia-Pacific psychiatry* (2015). doi:10.1111/appy.12196
60. Palminteri, S. & Pessiglione, M. Reinforcement learning and Tourette syndrome. *International review of neurobiology* 112, 131–153 (2013).
61. Maia, T. V. & Frank, M. J. From reinforcement learning models to psychiatric and neurological disorders. *Nature neuroscience* 14, 154–162 (2011).
62. Palminteri, S. et al. Dopamine-dependent reinforcement of motor skill learning: evidence from Gilles de la Tourette syndrome. *Brain* 134, 2287–2301 (2011).
63. Palminteri, S. et al. Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proc Natl Acad Sci U S A* 106, 19179–19184 (2009).
64. Pessiglione, M. et al. Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59, 561–567 (2008).
65. Delorme, C. et al. Enhanced habit formation in Gilles de la Tourette syndrome. *Brain* 139, 605–615 (2016).
66. Leckman, J. F. et al. The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. *J Am Acad Child Adolesc Psychiatry* 28, 566–573 (1989).
67. Welch, B. On the comparison of several mean values: an alternative approach. *Biometrika* 38, 330–336 (1951).

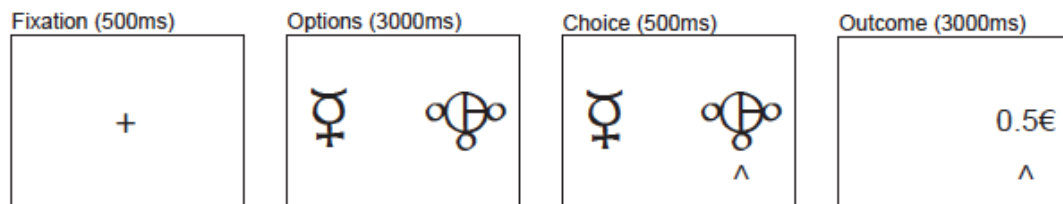
*Specific effect of dopamine partial agonist on counterfactual learning*

## Figures

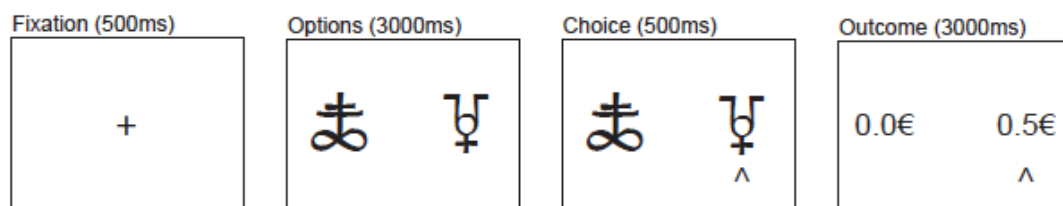
### (A) Task contingencies



### (B) Partial feedback condition

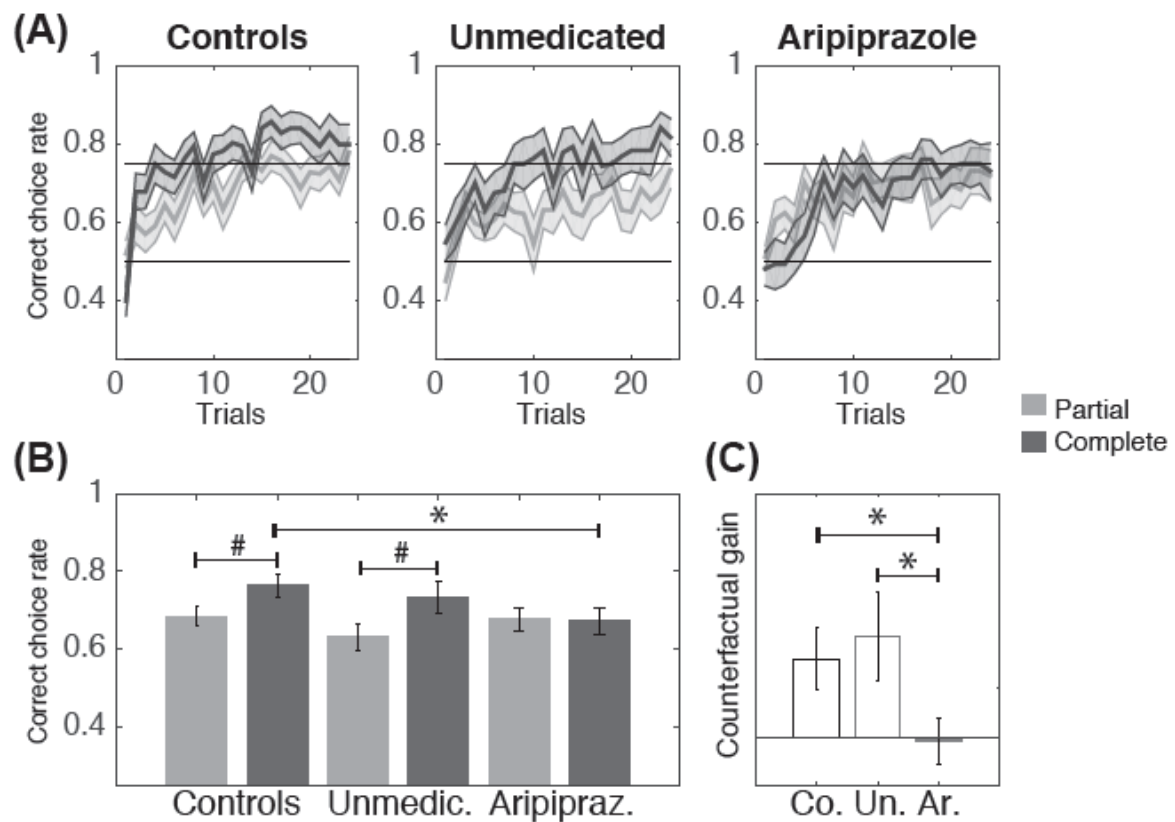


### Complete feedback condition



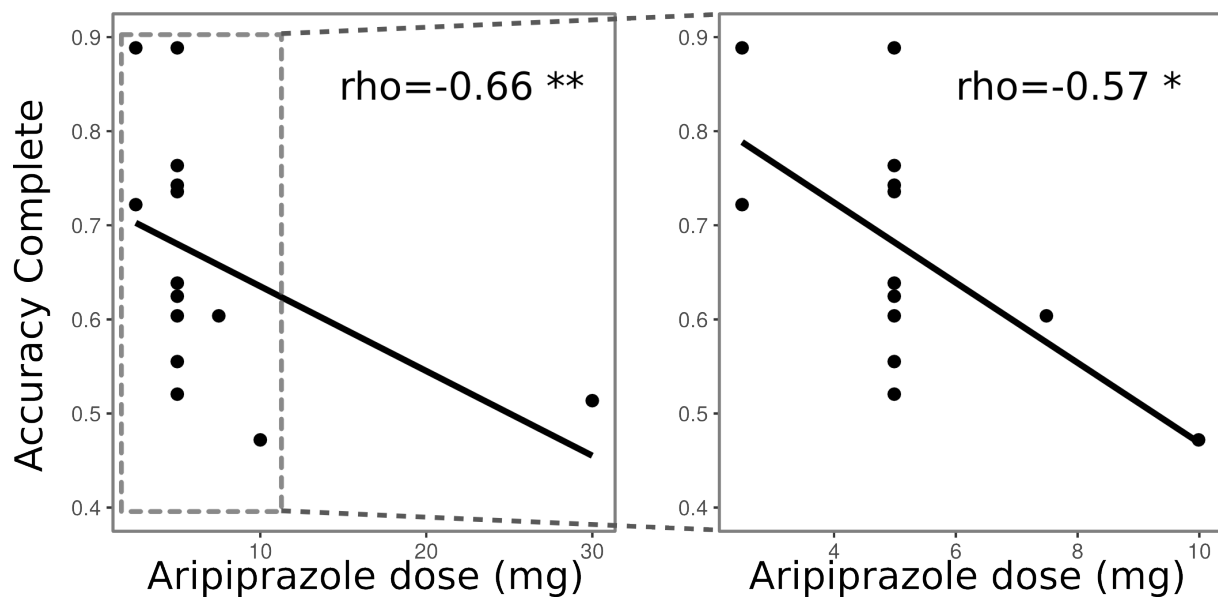
**Figure 1: Task. (A) Task contingencies and factorial design. The attribution of symbols to each condition was randomized across subjects. (B) Typical trials.** A novel trial started after a fixation screen (500 ms). Subjects were required to select between the two options by pressing one of the corresponding two buttons with their left or right index fingers to select the leftmost or the rightmost option, respectively, within a 3000-ms time window. After the choice window, a red pointer appeared below the selected option for 500 ms. At the end of the trial, the options disappeared and the selected option was replaced by the outcome (“+0.5 €”, “0.0 €” or “-0.5 €”) for 3000 ms. In the complete information context, the outcome corresponding to the unchosen option (counterfactual) was also displayed.

*Specific effect of dopamine partial agonist on counterfactual learning*

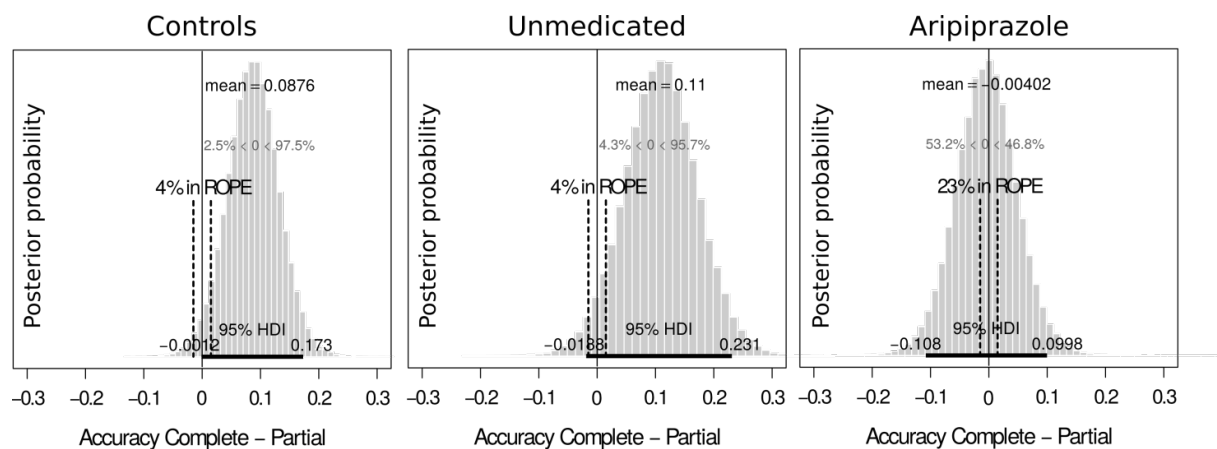


**Figure 2:** (A) Learning curves as a function of feedback information and group. The bold lines represent trial-by-trial average accuracy over the three sessions. The shaded areas between the thin lines represent the standard error of the mean with subject as the random factor. Light and dark grey represent partial and complete feedback conditions, respectively. (B) Average performance averaged across the 24 trials. (C) Average « counterfactual gain » (complete minus partial accuracy) (Co. = controls; Un. = unmedicated; Ar. = aripiprazole). # $p < 0.05$ , one-sample t-test; \*  $p < 0.05$ , two-sample t-test. The error bars represent the standard error of the mean with subject as the random factor.

*Specific effect of dopamine partial agonist on counterfactual learning*



**Figure 3:** Correlation between aripiprazole dose and accuracy in the complete feedback context in all subjects (left) and excluding the subject with the highest dose (right).



**Figure 4:** Results of the Bayesian analysis for evidence in favour of the null hypothesis for the difference between accuracy in the complete and partial feedback conditions (i.e., « counterfactual gain ») for each group. ROPE is the region of practical equivalence, which was set at the level of 3%  $[-1.5\%, 1.5\%]$ . “4% in ROPE” means that there is a 4% chance that the true average counterfactual gain lies within the region of the practical equivalence interval.

*Specific effect of dopamine partial agonist on counterfactual learning*

**Tables**

**Table 1: Demographic characteristics and psychometric scales for participants by treatment group.**

The “p2” p-values correspond to the comparison between the unmedicated and aripiprazole groups (two-sample t-test). The “p3” p-values correspond to the comparison among all three groups (one-way analysis of variance).

|                                | Controls          | Unmedicated       | Aripiprazole       | P2    | P3    |
|--------------------------------|-------------------|-------------------|--------------------|-------|-------|
| <b>N</b>                       | 20                | 17                | 14                 |       |       |
| <b>Age (m sd)</b>              | 32.35 (13.36)     | 30.76 (10.33)     | 32.07 (11.52)      | 0.742 | 0.916 |
| <b>Sex = Female (%)</b>        | 9 (45.0 %)        | 4 (23.5 %)        | 5 (35.7 %)         | 0.729 | 0.395 |
| <b>Study (m sd)</b>            | 14.40 (1.70)      | 13.24 (1.95)      | 12.79 (2.26)       | 0.557 | 0.050 |
| <b>NART (m sd)</b>             | 308.50<br>(94.70) | 262.81<br>(93.72) | 267.00<br>(106.20) | 0.911 | 0.310 |
| <b>YGTSS (m sd)</b>            | -                 | 37.87 (15.95)     | 31.14 (13.29)      | 0.230 | -     |
| <b>Disease Duration (m sd)</b> | -                 | 23.88 (10.95)     | 23.29 (11.81)      | 0.888 | -     |
| <b>BDI (m sd)</b>              | 1.75 (2.38)       | 6.59 (5.97)       | 7.07 (5.81)        | 0.822 | 0.003 |
| <b>STAI A (m sd)</b>           | 47.60 (3.86)      | 47.76 (2.97)      | 47.08 (4.25)       | 0.606 | 0.874 |
| <b>STAI B (m sd)</b>           | 47.71 (4.21)      | 50.29 (6.94)      | 51.09 (6.36)       | 0.762 | 0.272 |
| <b>BIS BAS (m sd)</b>          | 44.90 (5.18)      | 41.41 (5.76)      | 40.36 (6.55)       | 0.637 | 0.059 |
| <b>BIS 11 (m sd)</b>           | 60.40 (8.71)      | 65.94 (11.64)     | 69.29 (11.90)      | 0.437 | 0.057 |
| <b>Years on Aripip. (m sd)</b> | -                 | -                 | 5.75 (3.33)        | -     | -     |
| <b>Aripip. Dose, mg (m sd)</b> | -                 | -                 | 6.96 (6.88)        | -     | -     |

**Table 2: Average behavioural data (mean±s.e.m.).**

| Condition         | Accuracy (correct choice rate) |           |           |           | Reaction time (seconds) |           |           |           |
|-------------------|--------------------------------|-----------|-----------|-----------|-------------------------|-----------|-----------|-----------|
|                   | Rew./Par.                      | Rew./Com. | Pun./Par. | Pun./Com. | Rew./Par.               | Rew./Com. | Pun./Par. | Pun./Com. |
| <b>Controls</b>   | 0.68±0.04                      | 0.76±0.04 | 0.68±0.03 | 0.77±0.04 | 1.16±0.04               | 1.21±0.06 | 1.37±0.06 | 1.27±0.06 |
| <b>Unmedic.</b>   | 0.61±0.05                      | 0.74±0.05 | 0.65±0.04 | 0.72±0.04 | 1.14±0.06               | 1.16±0.06 | 1.29±0.06 | 1.28±0.07 |
| <b>Aripipraz.</b> | 0.66±0.05                      | 0.64±0.04 | 0.69±0.03 | 0.70±0.05 | 1.15±0.06               | 1.16±0.06 | 1.28±0.06 | 1.27±0.07 |

## Chapitre 6

### Article 2 : Effects of acute and chronic antidepressant treatment on reinforcement learning and cognition : a 8 week randomized double blind placebo controlled study in healthy volunteers

La dépression est une pathologie fréquente et invalidante, pour laquelle la dimension cognitive et motivationnelle est importante, associée à la fois à la rémission fonctionnelle et au sentiment subjectif de bien-être (Fervaha et al., 2016).

Les anomalies dans l'apprentissage par renforcement du fait de la dépression pourraient fournir des candidats endophénotypes pour cette pathologie (Chen et al., 2015). Les modifications dans l'apprentissage par renforcement pourraient également permettre de caractériser les effets des traitements antidépresseurs comme les inhibiteurs de recapture de la sérotonine (IRS). Un débat existe à ce sujet entre les tenants d'un rôle bénéfique des antidépresseurs comme les IRS sur la sensibilité à la récompense (Stoy et al., 2012; Wichers et al., 2009), et les tenants d'un rôle délétère de ces traitements, qui causeraient apathie et détachement émotionnel (Cooper et al., 2014; Price et al., 2009).

Pour mieux comprendre l'effet des traitements antidépresseurs, celui-ci peut être caractérisée chez le sujet sain (Antypa et al., 2014). Cependant, les études publiées jusqu'ici ont essentiellement étudié les effets des traitements en administration aiguë (Chamberlain et al., 2006), ou pendant une durée relativement courte, jusqu'à trois semaines maximum (Julian Macoveanu et al., 2014). Or l'effet des traitements antidépresseurs en utilisation clinique est retardé, avec un délai de deux à quatre semaines, et les recommandations internationales sont d'évaluer l'effet des traite-

ments après 6 à 12 semaines (National Collaborating Centre for Mental Health (UK), 2010).

Dans cette étude nous étudions l'effet des antidépresseurs chez le sujet sain à la fois en administration à court terme (3 jours) et en administration chronique (8 semaines) chez le sujet sain, en caractérisant son action dans deux tâches d'apprentissage par renforcement. L'une est une tâche probabiliste de sélection de stimulus simple. La seconde est une tâche d'apprentissage probabiliste plus complexe avec renversements occasionnels, qui permet d'étudier la propension des sujets à choisir des actions risquées, et leur propension à persévérer dans l'erreur. Nous utilisons ce dispositif pour comparer l'effet de deux traitements antidépresseurs de classe différente, l'escitalopram, un IRS typique, et l'agomélatine, un antidépresseur de mécanisme d'action différent avec une action monoaminergique combinée à une action mélatoninergique (Stahl, 2013).

# Effects of acute and chronic antidepressant treatment on reinforcement learning and cognition: a 8-week randomized double blind placebo controlled study in healthy volunteers.

Salvador A 1, Vinckier F 5, Pessiglione M 5, Deakin JFW 3, Harmer CJ 2, Montejo AL 4, Jabourian AP 6, Gruget C 7, Gabriel C 7, Goodwin GM 2, Gaillard R 1.

1. Centre Hospitalier Sainte Anne, INSERM U894, Université Paris Descartes, Sorbonne Paris Cité, Paris, France;
2. University Department of Psychiatry, Warneford Hospital, Oxford, UK;
3. Neuroscience and Psychiatry Unit, The University of Manchester, Manchester, UK;
4. Hospital Universitario de Salamanca, University of Salamanca, Salamanca, Spain;
5. Motivation Brain & Behavior lab (MBB), Brain & Spine Institute (ICM), Paris, France - CNRS UMR 7225 - UPMC-P6 UMR S1127;
6. Rue Marbeuf, Paris, France;
7. Institut de Recherches Internationales Servier, Suresnes, France.

## Abstract

The difference in mechanisms of action of different antidepressants is not fully understood. Studying their effects on reinforcement learning in healthy subjects is a way to characterise these mechanisms. Given the slow onset of action of antidepressants, these effects need to be studied after both short term and long term treatment, but previous studies did not evaluate these effects beyond 3 weeks.

The present double-blind, placebo-controlled study evaluates the effects of two classes of antidepressants (a typical SSRI, escitalopram, and an antidepressant with a different mechanism of action, agomelatine) on reinforcement learning and cognition in healthy subjects, both after short-term (3 days) and long-term (8 weeks) administration.

A total of 132 volunteers were randomly assigned to escitalopram (20mg), agomelatine (25 or 50mg) or placebo. Subjects were evaluated through a simple probabilistic stimulus selection task, and a complex probabilistic learning task encompassing measures of risk taking and perseveration.

First we demonstrated a beneficial effect of both agomelatine and escitalopram compared to placebo after chronic administration (8 weeks) in terms of accuracy in Task 1. Second, we found with agomelatine an early (3 days) beneficial effect in both Task 1 and Task 2, specifically on accuracy on the positive cue, and in terms of reaction time in Task 2. Third, the improved speed and accuracy with agomelatine were not accompanied by hallmarks of decreased flexibility or maladaptive risk taking.

Fourth, we obtained no beneficial effect of an increased dose of agomelatine (50mg compared to 25mg) on either Task 1 or 2.

These results challenge previous reports of impaired probabilistic learning as a result of acute antidepressant treatment, and illustrate the importance of not relying only on acute effects to extrapolate on the actual long term action of antidepressants, as it is already known in clinical practice.

## 1 Introduction

Depression is considered primarily a mood disorder, with at least one of the symptoms being either (1) depressed mood or (2) loss of interest or pleasure<sup>1</sup>. Cognitive symptoms have been recently shown to be frequent<sup>2</sup>, to be a major source of complaint<sup>3</sup> and to have important functional consequences<sup>4,5</sup>.

In line with motivation/pleasure impairment in depression, dysfunctions in reinforcement learning are thought to be promising candidates to serve as intermediate phenotypes<sup>6</sup> and might bridge a gap between mood and cognition<sup>7,8</sup>. Most studies have hitherto concentrated on the peculiarities of reinforcement learning in depressed patients (for reviews see 9 and 10). Another approach is to study how medication used to treat depression affects reinforcement learning mechanisms. Selective serotonin reuptake inhibitors (SSRI) like escitalopram, are the most prescribed antidepressants<sup>11</sup>, but some results suggest they might be detrimental on the motivational dimension, with increased apathy, loss of interest and emotional detachment<sup>12–30</sup>. On the contrary, other studies have shown that SSRI could increase reward sensitivity<sup>31–40</sup>, a positive effect that is commonly observed with antidepressants targeting catecholamines<sup>16,24,28,41–44</sup>. Finally, other antidepressants like agomelatine, have a different mechanism of action, combining monoaminergic and melatonergic properties<sup>45</sup>, with a more selective effect than SSRIs on emotional processing<sup>46</sup> and an earlier effect on pleasure and interest<sup>47</sup>. These clinical profiles might be related to different properties in terms of reinforcement learning. Testing antidepressants with different mechanisms of action might therefore be of interest in order to disentangle common and specific antidepressant effects.

One issue with clinical studies is that antidepressant effects may be confounded by an interaction with the pathological state: for instance an SSRI treatment was reported to decrease apathy in late-life depression<sup>40</sup> and to increase it in Parkinson's disease<sup>30</sup>. Studying the effect of antidepressants on healthy subjects might be more suited<sup>48</sup> and should help to characterize the differences in therapeutic action in terms of reinforcement learning.

In this view, Chamberlain and colleagues demonstrated that acute inhibition of central serotonin reuptake impaired probabilistic learning in healthy humans<sup>49</sup>. But only few studies on the cognitive impact of a longer term administration of antidepressants in healthy humans have been published to date, with 1 week<sup>50–54</sup>, 2 weeks<sup>22</sup> or 3 weeks treatment<sup>21,32,41</sup>.

This is an important shortcoming as the onset of antidepressant action appears with

a delay of 2-4 weeks and can only be fully evaluated after 6-12 weeks<sup>55</sup>. Short term and long term antidepressant treatments might therefore have different effects on reinforcement learning. For instance, Bari and colleagues found in a reversal learning task in rats that a single high dose of citalopram resulted in a decrease in lose-shift behaviour while repeated administration resulted in increased frequency of win-stay behaviour<sup>56</sup>.

Taking into account these different lines of results, our goal was to assess the comparative effect of two classes of antidepressants (a typical SSRI, escitalopram, and an antidepressant with a different mechanism of action, agomelatine) on human reinforcement learning both after sub-acute (3 days) and chronic administration (up to 8 weeks), using two tasks, a simple probabilistic learning task measuring classic reinforcement learning ability, and a more complex probabilistic learning task encompassing measures of the subject's propensity to take risk and to persevere in error, both adapted from previous paradigms (from 57 and 58 respectively).

## 2 Methods

### 2.1 Participants

Participants were recruited at two distinct sites in the UK (Oxford and Manchester) through public advertisement. The study was approved by the local Ethics Committee and registered in the International Standard Randomised Controlled Trial Number Registry under reference ISRCTN75872983. Inclusion criteria were healthy male or female volunteers, between 18 and 45 years, with a body mass index between 18 and 35 kg.m<sup>-2</sup>, non or moderate smokers (under 10 cigarettes a day), fluent in English. Participants gave their written informed consent prior to inclusion. Clinical examination including the structured clinical interview for DSM-IV-TR axis 1 disorder (SCID), physical examination (heart rate, blood pressure), and laboratory examination (haematology, blood and urine biochemistry) were performed to exclude any medical condition, psychiatric disease, use of psychotropic medication, drug consumption, pregnancy (blood and urine pregnancy test). All women had to use effective contraception. No treatment was allowed during the study, with the exception of oral contraception and paracetamol if required.

This study was exploratory and sample size was estimated based on comparable studies<sup>46,49</sup>. Exclusions and early drop outs (before week 5) were compensated to include 32 participants completing intervention (16 males and 16 females) in each treatment group, totalling 128 subjects (See CONSORT Flow Diagram in SOM Fig.S1). All 128 subjects were included in the analysis, totalling 379 visits, with 1 missing visit at day 3, no missing visit at day 14, and 4 missing visits at week 8. In addition, 3 visits were excluded for the analysis of Task 1, where subjects received a total negative monetary outcome, a sign that the task was misunderstood. Number of excluded visits was similar in each group with one visit in the placebo group, one in the escitalopram 20mg and one in the agomelatine 25mg group.

In addition, participants were evaluated for anxiety and depression levels using the Hospital Anxiety and Depression Scale<sup>59</sup>.

## 2.2 Study design

This study is a randomized, double blind, placebo controlled, parallel groups study. Treatment group was assigned at inclusion using a balanced (non adaptive) randomisation with stratification on gender and center and an allocation ratio of 1 to one of four treatment groups: placebo, agomelatine 25 mg, agomelatine 50 mg, escitalopram 20mg (10mg during week 1 and 9, 20mg during weeks 2 to 8).

Treatment or placebo was taken orally around 8 p.m. daily, and consisted of a single red capsule taken orally with a glass of water. Treatment was undistinguishable across groups and no case of unblinding was reported.

Testing was carried out on three occasions, after 3 days, 2 weeks and 8 weeks of treatment (Fig. 1.a). The study also included other evaluations, including questionnaires, emotion recognition tasks, and effort tasks, which will be presented elsewhere.

## 2.3 Tasks and data acquisition

At each visit, subjects were tested on two distinct tasks, a simple probabilistic task and a complex probabilistic task (Fig. 1.b and 1.c)

### 2.3.1 Simple probabilistic task

After a short practice, subjects had to perform three sessions of the same instrumental learning task, each proposing two new pairs of abstract stimuli randomly taken among letters from the Agathodaimon font. Each stimulus was associated with a given probability of gain or loss: one pair was associated with gains, (£1 or £0) and the other pair was associated with loss (-£1 or £0), with probabilities of 0.8 and 0.2 reciprocally. On each trial, one pair was randomly presented and the two stimuli were displayed on screen, to the right and left of a fixation cross, their relative position being counterbalanced across trials. Subjects were required to choose the right or left stimuli by pressing the 's' and 'l' keys respectively. To win money subjects had to learn, by trial and error, the stimulus-outcome associations.

After fixation delay and cue display, a 4 seconds interval was available for subjects to respond. The response was taken at the end of the period and categorized as 'risky' or 'less risky'. Monetary outcome was then displayed for 2 seconds. Participants were explicitly told that they would not receive the amounts earned during the task; they were paid a fix amount that compensated for their time and expenses associated with taking part in the study.

### 2.3.2 Complex probabilistic task

In this task, participants had to choose on each trial between a more and a less risky option, indicating their choice by pressing a key or not. The risky choice lead to the gain or the loss of £1, whereas the less risky option lead to the gain or loss of 10 pence. There were two different contextual cues. One was associated with a 80% chance of winning £1 (and a corresponding 20% chance of losing £1) following a risky choice and with 80% chance of winning 10 pence (and a corresponding 20% chance of losing 10 pence) following a less risky choice. The other cue had opposite contingencies, the risky choice gave an 80% chance of losing £1, while the less risky choice gave an 80% chance of losing 10 pence. The two cues described above are referred to as the “positive” and “negative” cue respectively. An unannounced contingency reversal occurred after the subject responded accurately 13 to 17 times in a row, or at least every 60 trials. A reversal means that the positive cue unexpectedly became negative and vice-versa.

After fixation delay and cue display, a 4 seconds interval was available for subjects to respond. The response was taken at the end of the period and categorized as ‘risky’ or ‘less risky’. Monetary outcome was then displayed for 2 seconds. Participants were explicitly told that they would not receive the amounts earned during the task; they were paid a fix amount that compensated for their time and expenses associated with taking part in the study.

Before performing the task, participants were familiarized with the task structure and with the notion that cue-outcome relationships were not necessarily constant. However, they were not warned that contingencies could be reversed.

## 2.4 Statistical analysis

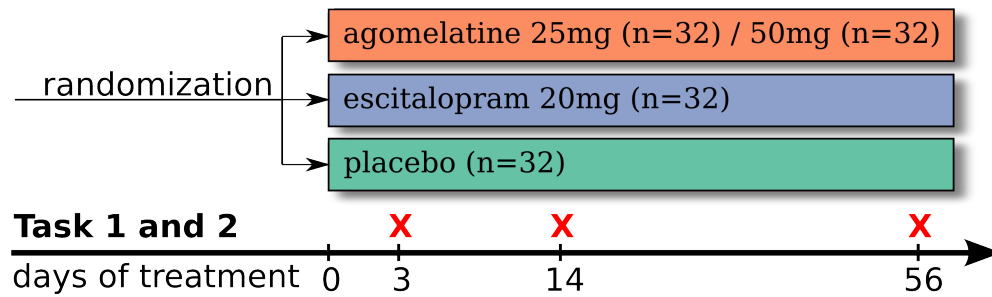
A comparison of demographic characteristics by group was performed using a MANOVA for quantitative variables (Age, anxiety, depression) and a chi-squared test for qualitative data (Sex).

The effect of the different treatment groups was analysed in two steps. First an assessment of the effect of treatment was performed comparing placebo, agomelatine 25mg and escitalopram 20mg. Second, a dose-effect analysis was performed, comparing the effect of placebo, agomelatine 25mg and agomelatine 50mg.

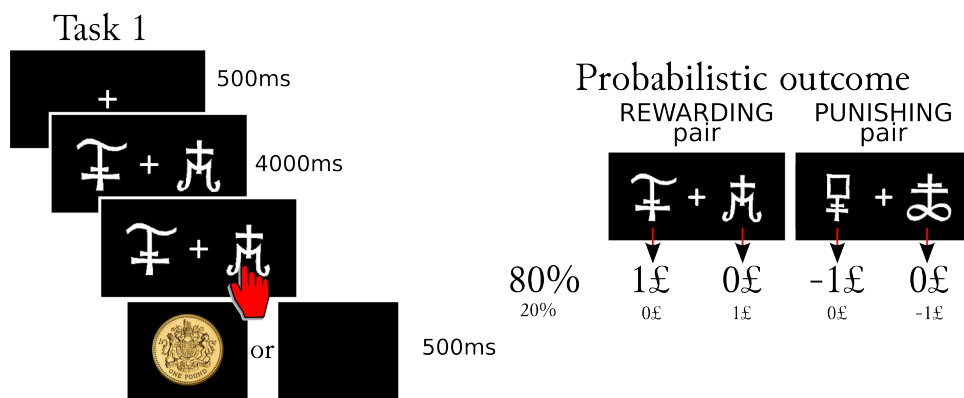
In task 1 and task 2, accuracy and reaction time were analysed using a repeated measure ANOVA with visit and valence as within subject variables, treatment as a between subject variable, and subject as random factor. Post hoc tests were performed using uncorrected pairwise t-tests. Significance level was set at 0.05.

In task 2, the additional variables of interest were the percentage of risky actions, the number of reversals achieved, and the number of perseverative errors, defined as the number of times a subject made the wrong choice given that at the preceding trial, he already made the wrong choice and feedback indicated him that the choice was wrong. Percentage of risky actions was analysed using repeated measure ANOVA with visit

a



b



c

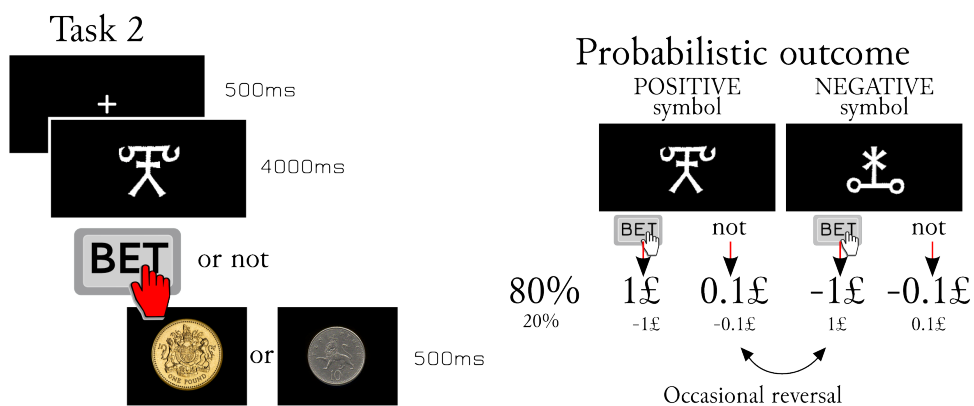


Figure 1: Study design and Tasks description. (a). Design of this double blind placebo controlled parallel groups study. At each of three visits, subjects performed two distinct tasks (b and c). (b). A simple probabilistic task in which subjects had to choose between two visual stimuli, with probabilistic gains or losses. (c). A complex probabilistic task in which subjects had to bet or not on a positive and negative symbol, with probabilistic outcome and occasional reversals.

and valence as within subject variables, treatment as between subject variable, and subject as random factor. The influence of valence was specifically scrutinized, as risky action with the positive symbol is adaptive while risky action with the negative symbol is detrimental. Number of reversals and perseverative errors were analysed using the same ANOVA, but without valence (both variables are global measures that cannot be split on a trial by trial basis). Post hoc tests used uncorrected pairwise t-tests.

Dose effect was analysed using the same ANOVA analysis as described above, with placebo, agomelatine 25mg and agomelatine 50mg as treatment groups.

While the main text concentrates on subacute versus chronic effects for clarity reasons, repeated measure ANOVAs were properly performed on all available subjects including all three visits (subacute, subchronic and chronic).

### 3 Results

#### 3.1 Demographics

Included participants are summarized in Table 1. Groups did not differ significantly in either demographic characteristics or in anxiety or depression levels (all  $p > 0.05$ ) albeit there was a small, marginally significant difference in Age ( $p=0.09$ ), with post-hoc tests showing no significant difference between groups (pairwise t-tests, all  $p > 0.16$ ).

Table 1: Demographic and psychometric characteristic of included subjects

Table 1 : Demographic and psychometric characteristic of included subjects

|                         | Placebo     | Escitalopram20 | Agomelatine25 | Agomelatine50 | p    |
|-------------------------|-------------|----------------|---------------|---------------|------|
| N                       | 32          | 32             | 32            | 32            |      |
| Sex (F/M)               | (16/16)     | 16/16)         | (16/16)       | (16/16)       | 1    |
| Age mean (sd)           | 23.0 (4.1)  | 24.2 (4.1)     | 24.2 (4.9)    | 21.8 (3.8)    | 0.09 |
| HADdepression mean (sd) | 1.19 (1.35) | 0.84 (1.65)    | 0.81 (1.15)   | 0.91 (1.20)   | 0.67 |
| HADanxiety mean (sd)    | 3.53 (2.55) | 3.78 (2.76)    | 3.72 (2.53)   | 3.50 (2.33)   | 0.96 |

#### 3.2 Effects of valence and visit on accuracy

Learning in both tasks was affected by valence [ $F(1,86)=7.291$ ,  $p=0.00834$  in task 1 and  $F(1,88)=11.674$ ,  $p=0.00096$  in task 2] with significantly higher performance with positive cues compared to negative cues [87.7% vs. 85.6%,  $t=-2.42$ ,  $p=0.0175$  in task 1 and 75.4% vs. 69.8%,  $t=-3.2$ ,  $p=0.0018$  in task 2].

Learning was also affected by visit [ $F(2,172)=5.630$ ,  $p=4.28 \text{ e-}3$  in task 1 and  $F(2,176)=13.422$ ,  $p=3.76 \text{ e-}6$  in task 2] with significantly higher accuracy in the chronic compared to sub-acute phase [88.1% vs. 84.8%,  $t=-2.79$   $p=0.0059$  in task 1

and 73.6% vs. 70.3%,  $t=-2.53$ ,  $p=0.0121$  in task 2]. These effects of valence and visit are visible in the average learning curves of task 1 in Figure 2.a.

### 3.3 Effect of treatment on accuracy and reaction time

In task 1, the effect of treatment on accuracy interacted with visit [ $F(4,172)=2.709$ ,  $p=0.0319$ ], and marginally with visit and valence [ $F(4,172)=2.236$ ,  $p=0.0672$ ]. In task 2, the effect of treatment on accuracy interacted with visit and valence [ $F(4,176)=2.953$ ,  $p=0.0215$ ].

This triple interaction results in a complex pattern that can be visualised in Figure 2.a. We focus below on the contrast between effects after chronic versus subacute treatment. Additional analyses including effects after subchronic treatment are included in supplementary online material.

#### 3.3.1 Both escitalopram and agomelatine improved performance on the positive stimuli against placebo after chronic treatment (8 weeks) in task 1

After chronic treatment (8 weeks), accuracy in task 1 with the positive pair was significantly higher with both antidepressants compared to placebo [pairwise t-tests against placebo: agomelatine  $p=0.0415$ , escitalopram  $p=0.0146$ ], with no difference between the two antidepressant groups [ $p=0.67$ ]. In contrast, with the negative pair, there was no such significant difference between treatment groups [all  $p>0.35$ ] (Figure 2.b and 2.c, right column).

In task 2, no significant difference in accuracy was found across treatment groups after chronic treatment either with the positive or negative cue [all  $p>0.05$ ] (Figure 2.c).

#### 3.3.2 Agomelatine yielded better performance on the positive stimuli against escitalopram after sub-acute treatment (3 days) in both task 1 and 2

After sub-acute treatment (3 days), accuracy in task 1 with the positive pair was higher with agomelatine compared to escitalopram [pairwise t-tests,  $p=0.00420$ ], and marginally so compared to placebo [ $p=0.0929$ ], while no difference between escitalopram and placebo existed [ $p=0.91$ ] (Figure 2.b left column). And no differences existed with the negative cues [all  $p>0.10$ ].

In Task 2 with the positive symbol, accuracy was higher with agomelatine compared to both escitalopram and placebo [pairwise t-tests,  $p=0.0199$  and  $p=0.0163$  respectively], while no difference between escitalopram and placebo existed [ $p=0.47$ ]. Again no difference existed with the negative cues across treatment groups [all  $p>0.10$ ] (Figure 2.c left column).

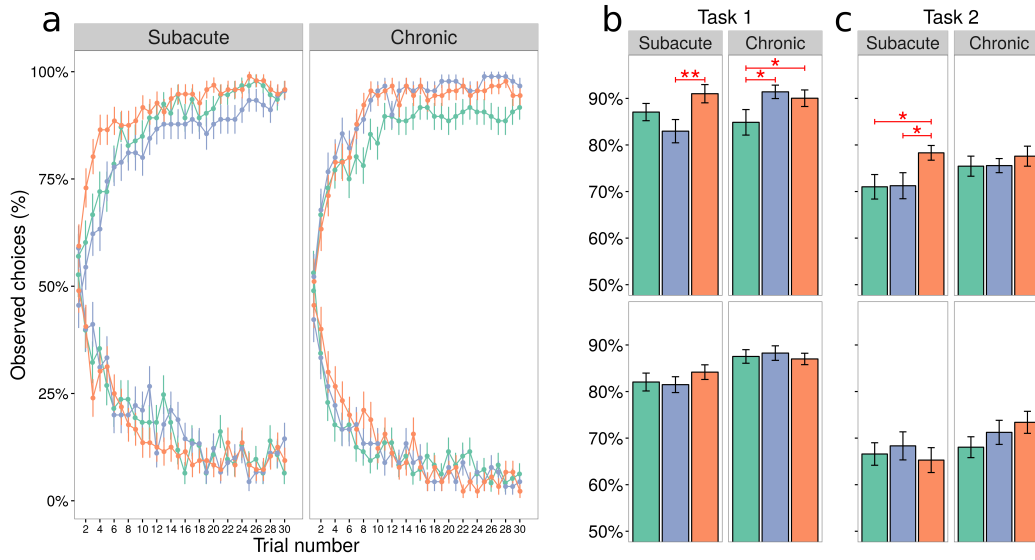


Figure 2: Accuracy in task 1 and 2, green=placebo, purple=escitalopram, orange=agomelatine. Learning curves by treatment group in task 1 after subacute (3 days) and chronic (8 weeks) treatment (a). Average accuracy by treatment group as a function of visit and valence for task 1 (b) and task 2 (c). Error bars represent standard error of the mean. Significance levels are : \*  $p<0.05$ , \*\*  $p<0.01$ .

### 3.3.3 Reaction time was shorter with agomelatine than with placebo in task 2

Reaction time was affected by valence and visit in both tasks [ $F(1,86)=246.7$ ,  $p<2e-16$  and  $F(2, 172)=36.4$ ,  $p=6.4e-14$  for valence and visit respectively in task 1;  $F(1, 88)=56.4$ ,  $p=4.6e-11$  and  $F(2, 176)=4.44$ ,  $p=0.013$  for valence and visit respectively in task 2], with shorter reaction times with the positive cues and shorter reaction times at the third visit (chronic) in all three treatment groups.

No effect of treatment on reaction time was found in task 1 [ $F(2,86)=0.984$ ,  $p=0.378$ ]. However, treatment had a main effect on reaction time in task 2 [ $F(2,88)=4.724$ ,  $p=0.0113$ ], with shorter reaction time in the agomelatine group compared to placebo [pairwise t-test,  $p=0.00262$ ] and marginally so compared to escitalopram [ $p=0.05208$ ], and no significant difference between escitalopram and placebo [ $p=0.28$ ].

No effect of treatment on reaction time was found in task 1 [ $F(2,86)=0.984$ ,  $p=0.378$ ].

### 3.3.4 Percentage of risky actions increased in the subacute phase with agomelatine

Regarding percentage of risky actions, there was an interaction between treatment and visit [ $F(4,176)=2.611$ ,  $p=0.0372$ ], with subjects in the agomelatine group taking marginally more risky actions in the subacute phase than subjects in the escitalopram or placebo groups ( $p=0.07$  and  $p=0.1$  respectively). Examining separately risky

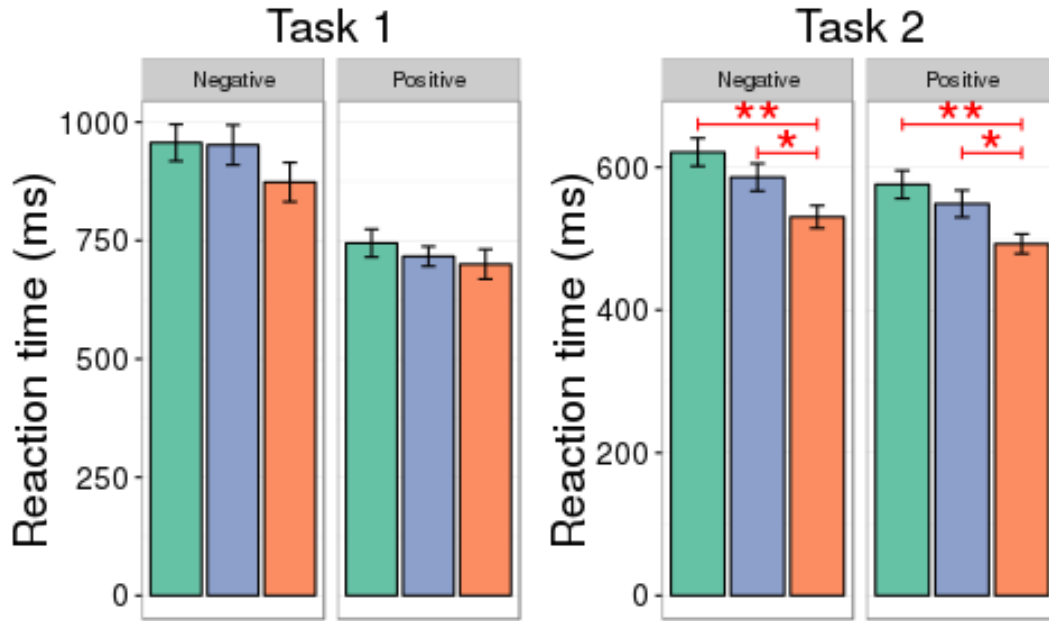


Figure 3: Reaction time by treatment group in task 1 (a) and 2 (b), green=placebo, purple=escitalopram, orange=agomelatine. Significance levels are : \*  $p < 0.05$ , \*\*  $p < 0.01$ .

actions by cue valence shows that percentage of risky action on the positive cue was higher with agomelatine compared to both escitalopram and placebo [pairwise t-test,  $p = 0.033$  and  $p = 0.021$  respectively], while percentage of risky actions on the negative cue did not differ across treatment groups [all  $p > 0.45$ ] (Fig. 4.a).

### 3.3.5 Number of reversals achieved and perseverative errors were not adversely affected by active treatment

The number of reversals achieved was not affected by treatment [ $F(2,90) = 1.042$ ,  $p = 0.357$ ], or by an interaction of treatment and visit [ $F(4,181) = 0.408$ ,  $p = 0.803$ ] (Fig. 4.b).

Regarding perseverative errors, there was an interaction between treatment and visit [ $F(4,87) = 5.921$ ,  $p = 0.00029$ ], with no difference across treatment groups in the sub-acute phase, but with fewer perseverative errors with agomelatine compared to placebo ( $p = 0.0148$ ) in the chronic phase, and no difference the other treatment groups (Fig. 4.c).

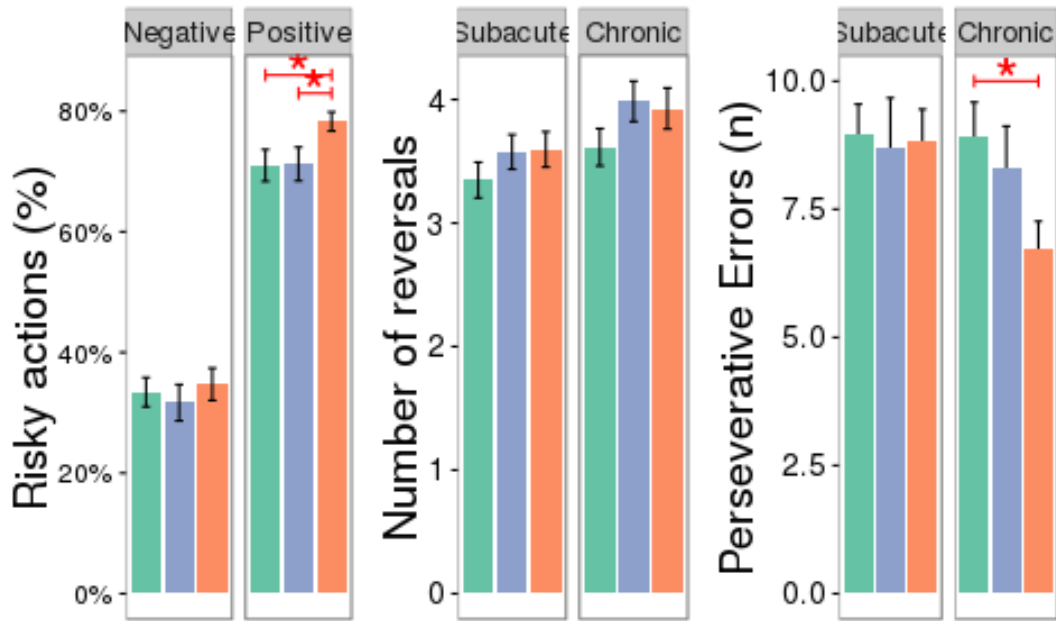


Figure 4: Task 2 : Number of reversals (a) and perseverative errors (b) for each treatment group in the subacute and chronic phases, percentage risky actions (c) for each treatment group in the subacute phase for the positive and negative symbols. Green=placebo, purple=escitalopram, orange=agomelatine.

### 3.3.6 No improvement in performance was obtained with higher dose of agomelatine

The effect of treatment dose on average accuracy interacted with visit and valence in task 1 [ $F(4,180)=2.674$ ,  $p=0.0335$ ], and in task 2 [ $F(4,182)=3.254$ ,  $p=0.0132$ ]. Post-hoc tests showed that there was no additional benefit of 50mg of agomelatine compared to 25mg of agomelatine :

After chronic treatment (8 weeks), average accuracy in Task 1 with the positive pair was significantly higher with 50mg of agomelatine compared to placebo [t-test,  $p=0.0304$ ] as is the case with 25mg of agomelatine, but there was no difference between 50mg and 25mg of agomelatine [t-test,  $p=0.505$ ]. In Task 2, accuracy with the positive symbols with agomelatine 50mg was not different from placebo [t-test,  $p=0.19$ ] and even marginally lower compared to agomelatine 25mg [t-test,  $p=0.0638$ ]. With the negative pair, there was no difference between 50mg and 25mg of agomelatine [all  $p>0.10$ ].

After sub-acute treatment (3 days), the early beneficial effect found with 25mg of agomelatine was not present with 50mg of agomelatine. In Task 1, average accuracy on the positive pair with agomelatine 50mg was not different from that with placebo [t-test,  $p=0.690$ ] and marginally lower than that with 25mg of agomelatine [t-test,  $p=0.0788$ ]. Likewise in Task 2, average accuracy on the positive symbol with 50mg of agomelatine was not different from that with placebo [t-test,  $p=0.671$ ] and marginally

lower than that with 25mg of agomelatine [t-test,  $p=0.0627$ ]. With the negative pair or symbol, there was no difference in average accuracy with 50mg compared to 25mg of agomelatine [t-test, all  $p>0.25$ ].

## 4 Discussion

This study of the sub-acute and chronic effects of two types of antidepressants escitalopram and agomelatine compared to placebo yielded four main results. First we demonstrated a beneficial effect of both agomelatine and escitalopram compared to placebo after chronic administration (8 weeks) in terms of accuracy in Task 1. Second, we found with agomelatine an early (3 days) beneficial effect in both Task 1 and Task 2, specifically on accuracy on the positive cue, and in terms of reaction time in Task 2. Third, these improved accuracy and speed were not accompanied by hallmarks of decreased flexibility or maladaptive risk taking, with even a decrease of perseverative errors. Fourth, we obtained no beneficial effect of an increased dose of agomelatine (50mg compared to 25mg) on either Task 1 or 2.

The first effect, a beneficial cognitive effect on positive cues for both escitalopram and agomelatine after chronic administration may come as a surprise with regards to previously published papers. Indeed, some clinicians reported patients experiencing blunted affects under SSRI (see for instance 12; 13), some even arguing that the persistence of cognitive symptoms might not be a lack of efficiency of treatments, but rather a side effect of specific classes of antidepressants<sup>60,61</sup>. And some experimental work seemed to illustrate this clinical observation, McCabe and colleagues showing after 2 weeks of treatment a decreased neural response to the sight and taste of chocolate in the ventral striatum and orbitofrontal cortex compared to placebo<sup>22</sup>. Moreover, Chamberlain and colleagues demonstrated that acute inhibition of central serotonin reuptake impaired probabilistic learning in healthy humans<sup>49</sup>. This effect is however consistent with animal studies that report a beneficial effect of long-lasting citalopram treatment on rat efficiency in a reversal learning task, whereas serotonin depletion had the opposite effect<sup>56</sup>. Thus, our results emphasize the danger of relying only on the short-term effects of medication to infer the effects of treatments in the long run. A beneficial effect on positive cues with long term antidepressant treatment is in line with the idea that antidepressant reverse a negative affective bias in depression by increasing the relative processing of positive versus negative affective information<sup>62,63</sup>. In this cognitive neuropsychological view of antidepressant action, antidepressants should lead to a better exploitation of positive events and to the learning of new emotional associations through this positive bias, both in the daily life and in the course of psychotherapy.

The second effect, an earlier beneficial effect of agomelatine specifically on the positive cue compared to escitalopram in both tasks may be related to the peculiar pharmacological profile of agomelatine compared to escitalopram. This early effect of agomelatine is consistent with a dopaminergic effect. The asymmetry in response to reward versus punishment is usually interpreted as the hallmark of a dopaminergic effect<sup>57,64-67</sup>. Likewise, the increase in percentage risky actions in task 2 only on the

positive stimulus is in line with the literature showing that boosting dopamine levels increased the number of risky options chosen in trials involving potential gains but not trials involving potential losses<sup>68,69</sup>. The overall improvement in reaction time regardless of stimulus valence in task 1 is also coherent with the cognitive improvement resulting from moderate dopamine increase, and decreased perseverative errors is in line with the remediation of cognitive inflexibility observed in Parkinson's disease treated with L-Dopa<sup>70</sup>, an effect we observed in the chronic condition. These results are coherent with the mechanisms of action of agomelatine, a 5-HT<sub>2C</sub> antagonist. 5-HT<sub>2C</sub> antagonists can indeed reverse decreased motivation following knockout or blockage of serotonin transporter<sup>71</sup>. Indeed electrophysiological studies in rats found an increase in VTA dopaminergic neurons activity after both acute and chronic administration of agomelatine<sup>72</sup>. Specifically, agomelatine was found to increase extracellular level of dopamine specifically in the frontal cortex, but not in the nucleus accumbens or striatum<sup>73</sup>. Crucially, the dopaminergic like effect observed here with agomelatine is not to blindly increase risk taking, but to improve subjects' willingness to bet on the positive cue. This effect, which contributes to the overall increase in accuracy, may be related to a prefrontal dopaminergic effect.

In contrast, escitalopram does not increase dopaminergic transmission and could even inhibit dopamine release through the 5HT<sub>2C</sub> receptor<sup>74,75</sup>. Furthermore, acute administration of citalopram impairs probabilistic learning in healthy humans<sup>49</sup>. In our experiment, accuracy on the positive cue with escitalopram was below that obtained with agomelatine, which might be attributed to this difference in pharmacological profile on dopamine. There was no significant detrimental effect of escitalopram compared to placebo, but the learning curves in task 1 seem to indicate that the significant difference between agomelatine and escitalopram is attributable to both a positive effect of agomelatine against placebo and a detrimental effect of escitalopram against placebo: although the first comparison was only marginal and the second failed to reach significance, the sum of both effects was significant. Our behavioural measures might have lacked of sensitivity. Indeed McCabe and colleagues similarly found no behavioural effect of escitalopram while there was a clear and significant brain effect<sup>22</sup>.

The early beneficial effect of agomelatine in healthy humans is in line with clinical reports of an early motivational effect of agomelatine in depressed patients compared to other antidepressants. Martinotti and colleagues for instance compared the effects of agomelatine and venlafaxine on depression in general and anhedonia in particular, and found similar antidepressant efficacy, but a higher efficacy on anhedonia<sup>47</sup>.

Limitations of this study include the fact that subjects did not perform the tasks before treatment onset. As a result, we cannot fully guarantee that groups performed equally in tasks before treatment. However, the fact that the study was randomized and controlled against placebo offsets this limitation. Another limitation is the complex pharmacokinetic profile of the drugs we investigated. This effect is not linear with time, as demonstrated by the effects obtained after sub-chronic treatment (2 weeks), presented in Supplementary Online Material. Further studies could help to more precisely describe this pharmacokinetic profile and to relate it to neurobiological changes across time.

In conclusion, this paper illustrates the importance of not relying only on acute effects to extrapolate on the actual action of antidepressants, as it is already known in clinical practice. The comparison of acute and chronic antidepressant administration therefore seems of great theoretical and practical importance. Cognition and motivation are now recognized as important clinical dimensions in depression, associated to both functional impairment and subjective well-being<sup>76</sup>. Targeting these dimensions through reinforcement learning paradigms to screen and select drugs that are active more specifically on those dimensions should be encouraged.

## 5 Funding and disclosure

AS, FV and RG have consulted for Servier and received compensation. RG has received compensation as a member of the scientific advisory board of Janssen, Lundbeck, Roche, Takeda. He has served as consultant and/or speaker for Astra Zeneca, Pierre Fabre, Lilly, Otsuka, SANOFI, Servier and received compensation, and he has received research support from Servier. MP declares no potential conflict of interest.

## References

1. American Psychiatric Association, A. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*. (American Psychiatric Publishing, 2013).
2. Conradi, H. J., Ormel, J. & de Jonge, P. Presence of individual (residual) symptoms during depressive episodes and periods of remission: a 3-year prospective study. *Psychol Med* **41**, 1165–1174 (2011).
3. Iverson, G. L. & Lam, R. W. Rapid screening for perceived cognitive impairment in major depressive disorder. *Ann Clin Psychiatry* **25**, 135–140 (2013).
4. Lam, R. W. *et al.* Which depressive symptoms and medication side effects are perceived by patients as interfering most with occupational functioning? *Depress Res Treat* **2012**, 630206 (2012).
5. Rock, P. L., Roiser, J. P., Riedel, W. J. & Blackwell, A. D. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol Med* **44**, 2029–2040 (2014).
6. Chen, C., Takahashi, T., Nakagawa, S., Inoue, T. & Kusumi, I. Reinforcement learning in depression: A review of computational research. *Neuroscience and biobehavioral reviews* **55**, 247–267 (2015).
7. Adams, R. A., Huys, Q. J. M. & Roiser, J. P. Computational Psychiatry: towards a mathematically informed understanding of mental illness. *Journal of neurology, neurosurgery, and psychiatry* **87**, 53–63 (2016).
8. Heinz, A., Schlagenhauf, F., Beck, A. & Wackerhagen, C. Dimensional psychiatry:

- mental disorders as dysfunctions of basic learning mechanisms. *Journal of neural transmission* **123**, 809–821 (2016).
9. Eshel, N. & Roiser, J. P. Reward and punishment processing in depression. *Biol Psychiatry* **68**, 118–124 (2010).
  10. Huys, Q. J., Pizzagalli, D. A., Bogdan, R. & Dayan, P. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of mood and anxiety disorders* **3**, 12 (2013).
  11. Bauer, M. *et al.* Prescribing patterns of antidepressants in Europe: results from the Factors Influencing Depression Endpoints Research (FINDER) study. *European psychiatry* **23**, 66–73 (2008).
  12. Opbroek, A. *et al.* Emotional blunting associated with SSRI-induced sexual dysfunction. Do SSRIs inhibit emotional responses? *Int J Neuropsychopharmacol* **5**, 147–151 (2002).
  13. Price, J., Cole, V. & Goodwin, G. M. Emotional side-effects of selective serotonin reuptake inhibitors: qualitative study. *Br J Psychiatry* **195**, 211–217 (2009).
  14. Barnhart, W. J., Makela, E. H. & Latocha, M. J. SSRI-induced apathy syndrome: a clinical review. *Journal of psychiatric practice* **10**, 196–199 (2004).
  15. Besnier, N. *et al.* Effects of paroxetine on emotional functioning and treatment awareness: a 4-week randomized placebo-controlled study in healthy clinicians. *Psychopharmacology* **207**, 619–629 (2010).
  16. Cooper, J. A., Tucker, V. L. & Papakostas, G. I. Resolution of sleepiness and fatigue: a comparison of bupropion and selective serotonin reuptake inhibitors in subjects with major depressive disorder achieving remission at doses approved in the European Union. *Journal of psychopharmacology* **28**, 118–124 (2014).
  17. Fava, M. *et al.* Clinical relevance of fatigue as a residual symptom in major depressive disorder. *Depression and anxiety* **31**, 250–257 (2014).
  18. Garland, E. J. & Baerg, E. A. Amotivational syndrome associated with selective serotonin reuptake inhibitors in children and adolescents. *Journal of child and adolescent psychopharmacology* **11**, 181–186 (2001).
  19. Hoehn-Saric, R., Lipsey, J. R. & McLeod, D. R. Apathy and indifference in patients on fluvoxamine and fluoxetine. *Journal of clinical psychopharmacology* **10**, 343–345 (1990).
  20. Lee, S. I. & Keltner, N. L. Antidepressant apathy syndrome. *Perspectives in psychiatric care* **41**, 188–192 (2005).
  21. Macoveanu, J. *et al.* Effects of selective serotonin reuptake inhibition on neural activity related to risky decisions and monetary rewards in healthy males. *NeuroImage* **99**, 434–442 (2014).
  22. McCabe, C., Mishor, Z., Cowen, P. J. & Harmer, C. J. Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake

inhibitor treatment. *Biol Psychiatry* **67**, 439–445 (2010).

23. Padala, P. R., Padala, K. P., Monga, V., Ramirez, D. A. & Sullivan, D. H. Reversal of SSRI-associated apathy syndrome by discontinuation of therapy. *The Annals of pharmacotherapy* **46**, e8 (2012).

24. Papakostas, G. I. *et al.* Resolution of sleepiness and fatigue in major depressive disorder: A comparison of bupropion and the selective serotonin reuptake inhibitors. *Biological psychiatry* **60**, 1350–1355 (2006).

25. Reinblatt, S. P. & Riddle, M. A. Selective serotonin reuptake inhibitor-induced apathy: a pediatric case series. *Journal of child and adolescent psychopharmacology* **16**, 227–233 (2006).

26. Rothschild, A. J., Raskin, J., Wang, C.-N., Marangell, L. B. & Fava, M. The relationship between change in apathy and changes in cognition and functional outcomes in currently non-depressed SSRI-treated patients with major depressive disorder. *Comprehensive psychiatry* **55**, 1–10 (2014).

27. Sansone, R. A. & Sansone, L. A. SSRI-Induced Indifference. *Psychiatry* **7**, 14–18 (2010).

28. Schmitt, J. A. J. *et al.* Additional dopamine reuptake inhibition attenuates vigilance impairment induced by serotonin reuptake inhibition in man. *Journal of psychopharmacology* **16**, 207–214 (2002).

29. Wongpakaran, N., Reekum, R. van, Wongpakaran, T. & Clarke, D. Selective serotonin reuptake inhibitor use associates with apathy among depressed elderly: a case-control study. *Annals of general psychiatry* **6**, 7 (2007).

30. Zahodne, L. B. *et al.* Are selective serotonin reuptake inhibitors associated with greater apathy in Parkinson's disease? *The Journal of neuropsychiatry and clinical neurosciences* **24**, 326–330 (2012).

31. Boyer, P., Tassin, J. P., Falissart, B. & Troy, S. Sequential improvement of anxiety, depression and anhedonia with sertraline treatment in patients with major depression. *Journal of clinical pharmacy and therapeutics* **25**, 363–371 (2000).

32. Furlan, P. M., Kallan, M. J., Have, T. T., Lucki, I. & Katz, I. SSRIs do not cause affective blunting in healthy elderly volunteers. *The American journal of geriatric psychiatry* **12**, 323–330 (2004).

33. Mikami, K. *et al.* Prevention of poststroke apathy using escitalopram or problem-solving therapy. *The American journal of geriatric psychiatry* **21**, 855–862 (2013).

34. Raskin, J. *et al.* Apathy in currently nondepressed patients treated with a SSRI for a major depressive episode: outcomes following randomized switch to either duloxetine or escitalopram. *Journal of psychiatric research* **46**, 667–674 (2012).

35. Siddique, H., Hynan, L. S. & Weiner, M. F. Effect of a serotonin reuptake inhibitor on irritability, apathy, and psychotic symptoms in patients with Alzheimer's disease. *The Journal of clinical psychiatry* **70**, 915–918 (2009).

36. Stoy, M. *et al.* Hyporeactivity of ventral striatum towards incentive stimuli

in unmedicated depressed patients normalizes after treatment with escitalopram. *Journal of psychopharmacology* **26**, 677–688 (2012).

37. Tang, T. Z. *et al.* Personality change during depression treatment: a placebo-controlled trial. *Archives of general psychiatry* **66**, 1322–1330 (2009).

38. Tse, W. S., Chow, H., Wing, Y. K. & Bond, A. J. Using a partner’s facial emotion to elucidate social dominance motivation induced by an SSRI. *European neuropsychopharmacology* **24**, 1641–1649 (2014).

39. Wichers, M. C. *et al.* Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology* **34**, 923–931 (2009).

40. Yuen, G. S. *et al.* Neuroanatomical correlates of apathy in late-life depression and antidepressant treatment response. *Journal of affective disorders* **166**, 179–186 (2014).

41. Furlan, P. M. *et al.* Cognitive and psychomotor effects of paroxetine and sertraline on healthy elderly volunteers. *The American journal of geriatric psychiatry* **9**, 429–438 (2001).

42. Nutt, D. *et al.* The other face of depression, reduced positive affect: the role of catecholamines in causation and cure. *Journal of psychopharmacology* **21**, 461–471 (2007).

43. Ossewaarde, L. *et al.* Two-week administration of the combined serotonin-noradrenaline reuptake inhibitor duloxetine augments functioning of mesolimbic incentive processing circuits. *Biol Psychiatry* **70**, 568–574 (2011).

44. Pringle, A., McCabe, C., Cowen, P. J. & Harmer, C. J. Antidepressant treatment and emotional processing: can we dissociate the roles of serotonin and noradrenaline? *Journal of psychopharmacology* **27**, 719–731 (2013).

45. Stahl, S. M. *Stahl’s essential psychopharmacology: neuroscientific basis and practical applications*. (Cambridge university press, 2013).

46. Harmer, C. J. *et al.* Agomelatine facilitates positive versus negative affective processing in healthy volunteer models. *J Psychopharmacol* **25**, 1159–1167 (2011).

47. Martinotti, G. *et al.* Agomelatine versus venlafaxine XR in the treatment of anhedonia in major depressive disorder: a pilot study. *J Clin Psychopharmacol* **32**, 487–491 (2012).

48. Antypa, N., Calati, R. & Serretti, A. The neuropsychological hypothesis of antidepressant drug action revisited. *CNS Neurol Disord Drug Targets* **13**, 1722–1739 (2014).

49. Chamberlain, S. R. *et al.* Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science* **311**, 861–863 (2006).

50. Di Simplicio, M., Norbury, R. & Harmer, C. J. Short-term antidepressant administration reduces negative self-referential processing in the medial prefrontal

- cortex in subjects at risk for depression. *Molecular psychiatry* **17**, 503–510 (2012).
51. Norbury, R. *et al.* Short-term antidepressant treatment modulates amygdala response to happy faces. *Psychopharmacology* **206**, 197–204 (2009).
  52. Murphy, S. E., Yiend, J., Lester, K. J., Cowen, P. J. & Harmer, C. J. Short-term serotonergic but not noradrenergic antidepressant administration reduces attentional vigilance to threat in healthy volunteers. *The international journal of neuropsychopharmacology* **12**, 169–179 (2009).
  53. Norbury, R., Mackay, C. E., Cowen, P. J., Goodwin, G. M. & Harmer, C. J. Short-term antidepressant treatment and facial processing. Functional magnetic resonance imaging study. *The British journal of psychiatry* **190**, 531–532 (2007).
  54. Harmer, C. J., Shelley, N. C., Cowen, P. J. & Goodwin, G. M. Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *The American journal of psychiatry* **161**, 1256–1263 (2004).
  55. National Collaborating Centre for Mental Health (UK). Depression: The Treatment and Management of Depression in Adults (Updated Edition). (2010).
  56. Bari, A. *et al.* Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology* **35**, 1290–1301 (2010).
  57. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & D., F. C. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).
  58. Vinckier, F. *et al.* Confidence and psychosis: a neuro-computational account of contingency learning disruption by NMDA blockade. *Mol Psychiatry* (2016). doi:10.1038/mp.2015.73
  59. Zigmond, A. S. & Snaith, R. P. The hospital anxiety and depression scale. *Acta Psychiatr Scand* **67**, 361–370 (1983).
  60. Millan, M. J. *et al.* Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nat Rev Drug Discov* **11**, 141–168 (2012).
  61. Nagane, A. *et al.* Comparative study of cognitive impairment between medicated and medication-free patients with remitted major depression: class-specific influence by tricyclic antidepressants and newer antidepressants. *Psychiatry Res* **218**, 101–105 (2014).
  62. Harmer, C. J., Duman, R. S. & Cowen, P. J. How do antidepressants work? New perspectives for refining future treatment approaches. *The lancet. Psychiatry* (2017). doi:10.1016/S2215-0366(17)30015-9
  63. Harmer, C. J. & Cowen, P. J. 'It's the way that you look at it'—a cognitive neuropsychological account of SSRI action in depression. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **368**, 20120407 (2013).
  64. Ikemoto, S., Yang, C. & Tan, A. Basal ganglia circuit loops, dopamine and

- motivation: A review and enquiry. *Behav Brain Res* **290**, 17–31 (2015).
65. Schultz, W. Dopamine reward prediction error coding. *Dialogues in clinical neuroscience* **18**, 23–32 (2016).
  66. Niv, Y. Cost, benefit, tonic, phasic: what do response rates tell us about dopamine and motivation? *Ann N Y Acad Sci* **1104**, 357–376 (2007).
  67. Wise, R. A. Dopamine, learning and motivation. *Nat Rev Neurosci* **5**, 483–494 (2004).
  68. Friston, K. *et al.* The anatomy of choice: active inference and agency. *Front Hum Neurosci* **7**, 598 (2013).
  69. Rutledge, R. B., Skandali, N., Dayan, P. & Dolan, R. J. Dopaminergic Modulation of Decision Making and Subjective Well-Being. *J Neurosci* **35**, 9811–9822 (2015).
  70. Cools, R., Barker, R. A., Sahakian, B. J. & Robbins, T. W. L-Dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson’s disease. *Neuropsychologia* **41**, 1431–1441 (2003).
  71. Browne, C. J. & Fletcher, P. J. Decreased Incentive Motivation Following Knockout or Acute Blockade of the Serotonin Transporter: Role of the 5-HT<sub>2C</sub> Receptor. *Neuropsychopharmacology* **41**, 2566–2576 (2016).
  72. Chenu, F., El Mansari, M. & Blier, P. Electrophysiological effects of repeated administration of agomelatine on the dopamine, norepinephrine, and serotonin systems in the rat brain. *Neuropsychopharmacology* **38**, 275–284 (2013).
  73. Millan, M. J. *et al.* The novel melatonin agonist agomelatine (S20098) is an antagonist at 5-hydroxytryptamine<sub>2C</sub> receptors, blockade of which enhances the activity of frontocortical dopaminergic and adrenergic pathways. *J Pharmacol Exp Ther* **306**, 954–964 (2003).
  74. Dewey, S. L. *et al.* Serotonergic modulation of striatal dopamine measured with positron emission tomography (PET) and in vivo microdialysis. *The Journal of neuroscience* **15**, 821–829 (1995).
  75. Dremencov, E., El Mansari, M. & Blier, P. Effects of sustained serotonin reuptake inhibition on the firing of dopamine neurons in the rat ventral tegmental area. *Journal of psychiatry and neuroscience* **34**, 223–229 (2009).
  76. Fervaha, G., Foussias, G., Takeuchi, H., Agid, O. & Remington, G. Motivational deficits in major depressive disorder: Cross-sectional and longitudinal relationships with functional impairment and subjective well-being. *Comprehensive psychiatry* **66**, 31–38 (2016).

## Supplementary Information

Effects of two antidepressants Agomelatine and Escitalopram on learning and cognition :  
a 8 week randomized double blind placebo controlled study in healthy volunteers.

### 1 CONSORT Flowchart

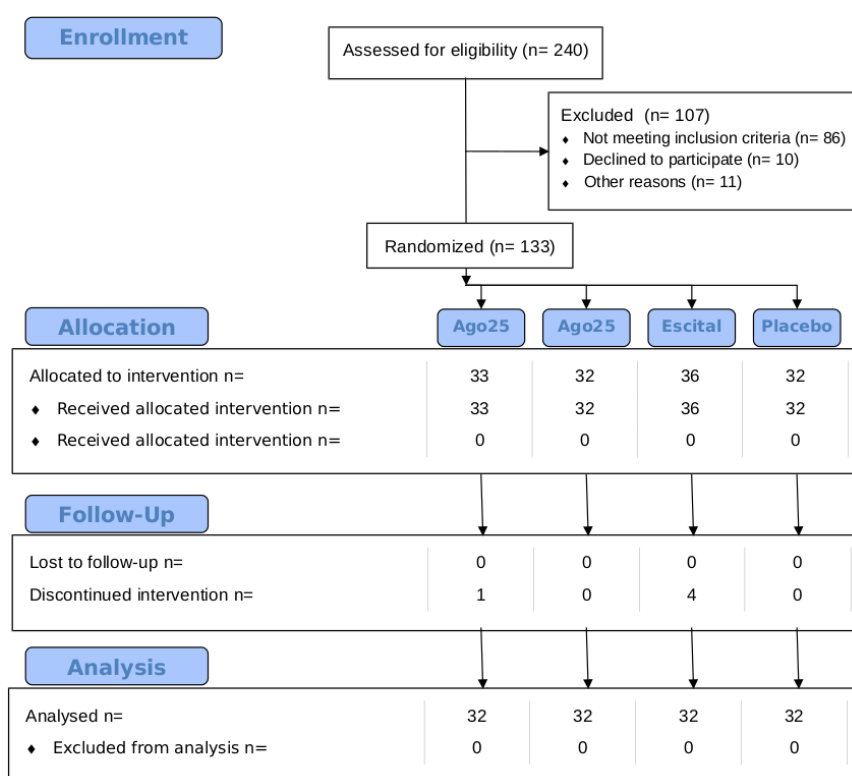


Figure 1: Figure S1 : CONSORT Flowchart of the study

## 2 Accuracy by valence, and treatment group in the sub-chronic condition

Accuracy and reaction time were analysed in the paper using a repeated measure ANOVA with visit and valence as within subject variables, treatment as a between subject variable, and subject as random factor. Visit included all three visits and therefore full statistical analyses have been reported in the paper.

We only present below the graphical representations of Accuracy for the sub-chronic visit along with the corresponding post-hoc tests that were not presented in the main paper. Figure S2 is the equivalent of Figure 2 for the sub-chronic visit.

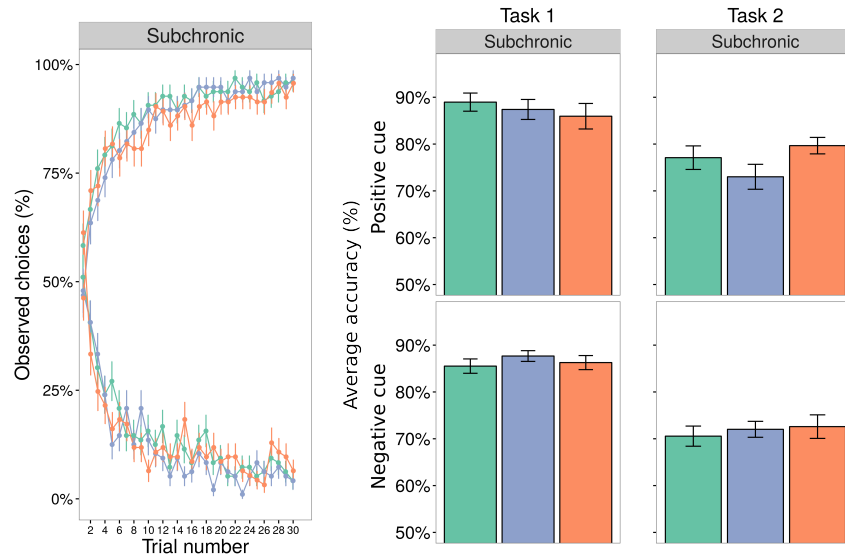


Figure 2: Figure S2 : Accuracy in task 1 and 2, green=placebo, purple=escitalopram, orange=agomelatine. Learning curves by treatment group in task 1 after subchronic (2 weeks) treatment (a). Average accuracy by treatment group as a function of valence for task 1 (b) and task 2 (c). Error bars represent standard error of the mean.

Post-hoc tests show that there was no difference across treatment groups either on the positive or negative cues in task 1 (all  $p > 0.10$ )

### 3 Details of Accuracy by treatment, visit and valence in task 1 and task 2

#### 3.1 Table S1 - Accuracy by Session in Task 1 and Task 2

| Task   | Visit             |        |                     |        |                  |        |              |        |
|--------|-------------------|--------|---------------------|--------|------------------|--------|--------------|--------|
|        | Subacute Accuracy |        | Subchronic Accuracy |        | Chronic Accuracy |        | All Accuracy |        |
|        | mean              | sd     | mean                | sd     | mean             | sd     | mean         | sd     |
| Task 1 | 0.8484            | 0.1112 | 0.8697              | 0.1058 | 0.8814           | 0.1002 | 0.8665       | 0.1066 |
| Task 2 | 0.7013            | 0.1480 | 0.7416              | 0.1296 | 0.7351           | 0.1240 | 0.7260       | 0.1353 |
| All    | 0.7741            | 0.1502 | 0.8053              | 0.1345 | 0.8082           | 0.1343 | 0.7958       | 0.1406 |

#### 3.2 Table S2 - Accuracy by Valence in Task 1 and Task 2

| Task   | Valence           |         |                   |        |              |        |
|--------|-------------------|---------|-------------------|--------|--------------|--------|
|        | Negative Accuracy |         | Positive Accuracy |        | All Accuracy |        |
|        | mean              | sd      | mean              | sd     | mean         | sd     |
| Task 1 | 0.8556            | 0.08715 | 0.8774            | 0.1221 | 0.8665       | 0.1066 |
| Task 2 | 0.6975            | 0.13681 | 0.7544            | 0.1278 | 0.7260       | 0.1353 |
| All    | 0.7761            | 0.13935 | 0.8156            | 0.1392 | 0.7958       | 0.1406 |

Table S3 - Accuracy by Session and Valence and Treatment group in Task 1 and Task 2

| Visit  |                |                   |         |                   |         |                    |         |                   |         |                   |         |                   |         |
|--------|----------------|-------------------|---------|-------------------|---------|--------------------|---------|-------------------|---------|-------------------|---------|-------------------|---------|
| Task   | Treatment      | Subacute Valence  |         |                   |         | Subchronic Valence |         |                   |         | Chronic Valence   |         |                   |         |
|        |                | Negative Accuracy |         | Positive Accuracy |         | Negative Accuracy  |         | Positive Accuracy |         | Negative Accuracy |         | Positive Accuracy |         |
|        |                | mean              | sd      | mean              | sd      | mean               | sd      | mean              | sd      | mean              | sd      | mean              | sd      |
| Task 1 | Placebo        | 0.8204            | 0.10689 | 0.8706            | 0.10310 | 0.8552             | 0.08673 | 0.8896            | 0.10967 | 0.8753            | 0.08251 | 0.8486            | 0.15506 |
|        | Escitalopram20 | 0.8148            | 0.09321 | 0.8296            | 0.13675 | 0.8767             | 0.06487 | 0.8740            | 0.12072 | 0.8826            | 0.08555 | 0.9141            | 0.07905 |
|        | Agomelatine25  | 0.8417            | 0.08907 | 0.9101            | 0.11080 | 0.8627             | 0.08362 | 0.8595            | 0.15179 | 0.8700            | 0.06712 | 0.9004            | 0.09741 |
| Task 2 | Placebo        | 0.6659            | 0.13680 | 0.7102            | 0.14890 | 0.7056             | 0.12170 | 0.7710            | 0.14227 | 0.6805            | 0.12686 | 0.7544            | 0.12196 |
|        | Escitalopram20 | 0.6834            | 0.16779 | 0.7125            | 0.15561 | 0.7202             | 0.09641 | 0.7301            | 0.15072 | 0.7124            | 0.14163 | 0.7557            | 0.08290 |
|        | Agomelatine25  | 0.6528            | 0.15116 | 0.7831            | 0.08945 | 0.7260             | 0.14229 | 0.7967            | 0.09948 | 0.7339            | 0.12959 | 0.7759            | 0.11772 |
| All    | Placebo        | 0.7419            | 0.14475 | 0.7891            | 0.15088 | 0.7804             | 0.12913 | 0.8303            | 0.13947 | 0.7779            | 0.14460 | 0.8015            | 0.14629 |
|        | Escitalopram20 | 0.7480            | 0.15055 | 0.7701            | 0.15694 | 0.7985             | 0.11343 | 0.8020            | 0.15363 | 0.7975            | 0.14431 | 0.8349            | 0.11326 |
|        | Agomelatine25  | 0.7472            | 0.15558 | 0.8466            | 0.11863 | 0.7933             | 0.13512 | 0.8276            | 0.13077 | 0.8019            | 0.12320 | 0.8381            | 0.12415 |



Troisième partie

Discussion générale



# Chapitre 7

## Discussion générale

### 7.1 Critique du choix des tâches d'apprentissage

#### 7.1.1 Le choix des tâches

Dans l'expérience "Effects of acute and chronic antidepressant treatment on reinforcement learning and cognition : a 8 week randomized double blind placebo controlled study in healthy volunteers", la première tâche était une adaptation d'un paradigme développé initialement par Pessiglione et al. (2006). Dans la tâche originale, trois paires de stimuli abstraits étaient utilisées :

- Une paire positive avec une probabilité de gagner £1/£0 de 80%/20% pour un des stimuli, et 20%/80% pour l'autre ;
- Une paire négative avec une probabilité de perdre £1/£0 de 80%/20% pour un des stimuli, et 20%/80% pour l'autre ;
- Une paire neutre sans gain ni perte quelque soit le choix du stimulus.

Dans notre tâche, la paire neutre a été omise, espérant ainsi réduire le temps de passation tout en préservant la possibilité de l'analyse d'un effet valence. Il n'y avait donc que deux paires de symboles, l'une positive et l'autre négative. Ce changement, qui pourrait paraître mineur, a amené des modifications importantes dans le comportement mesuré chez les sujets. De fait, la suppression de la paire neutre peut aboutir à une dichotomie entre paire positive et paire négative qui amène à classer très rapidement les paires selon cette dichotomie au lieu de simplement suivre la valeur de chaque stimulus. Une fois ainsi catégorisées, l'obtention d'un résultat nul (£0) n'a plus le même sens : il s'agit d'un manque à gagner dans la paire positive, ce qui peut être considéré comme une perte, et il s'agit d'un moindre mal dans la paire négative, ce qui peut être considéré comme un gain. Si le sujet utilise une telle stratégie, son comportement amènera très rapidement à des choix optimaux.

Il est possible d’analyser le comportement des sujets en termes d’apprentissage par renforcement. Nous y sommes encouragés par les courbes d’apprentissage moyennes par groupe de traitement. Celles-ci sont tout à fait satisfaisantes : elles ressemblent aux courbes classiquement observées dans les tâches d’apprentissage par renforcement, avec une amélioration progressive de la performance, puis une stabilisation à un niveau plateau. Ces courbes sont visibles dans la Figure 2 de l’article 2 reproduit dans cette thèse.

Analyser le comportement en termes d’apprentissage par renforcement pour un modèle particulier consiste à rechercher le jeu de paramètres pour ce modèle qui maximise la fonction de vraisemblance. Cette fonction de vraisemblance  $P(D|M, \theta_m)$  représente la probabilité d’observer les données comportementales recueillies étant donné le modèle et les paramètres ;  $D$  représente les données,  $M$  le modèle et  $\theta_m$  les paramètres du modèle.

De la même façon qu’il est toujours possible d’ajuster une régression linéaire dans un nuage de points, il est (presque) toujours possible de trouver un jeu de paramètres qui maximise la fonction de vraisemblance. Si l’on souhaite comparer plusieurs modèles, on choisira comme meilleur modèle celui qui minimise un critère comme l’Akaike Information Criterion (AIC) ou le Bayesian Information Criterion (BIC). Dans tous les cas, il est possible de se contenter de cette analyse *relative* pour déterminer et rapporter le meilleur modèle et le meilleur jeu de paramètres pour ce meilleur modèle. Cependant, la qualité *absolue* du modèle et sa capacité à modéliser le comportement de sujets individuels est une information importante, même si elle est en pratique rarement rapportée ou discutée dans les publications.

Dans le cas de notre étude, l’analyse au niveau individuel montre des particularités qui semblent importantes à discuter. Notre première tâche est un paradigme probabiliste de sélection de stimulus. Or si certains sujets ont un comportement qui est apparemment compatible avec un apprentissage par renforcement classique de type model-free, d’autres sujets ont un comportement qu’il paraît difficile de modéliser par ce type d’algorithme. Le comportement de deux sujets individuels dans cette tâche est illustrée Figure 7.1.

Le sujet 92 a un taux d’erreurs initialement assez important qui s’amenuise au fur et à mesure des essais lors de la première visite. Le même comportement s’observe à la dernière visite, avec un bénéfice visible de la visite, le nombre d’erreurs diminuant plus rapidement. Ceci se traduit dans une modélisation computationnelle par un taux d’apprentissage qui s’améliore entre la première et la dernière visite, ce qui fait écho à l’effet statistiquement significatif de la visite sur la performance rapportée en analyse statistique classique dans l’article. La situation est différente pour le sujet 67. Lors de la première visite, son comportement reste compatible avec un apprentissage par renforcement de type model-free, avec un profil d’erreurs comparable au sujet 92. Lors de la dernière visite en revanche, le sujet ne fait aucune erreur. Il a donc immédiatement discriminé tous les stimuli sur la base du premier renforçateur reçu. Il n’a donc pas échantillonné les choix alternatifs.

Ce comportement est révélateur d'un apprentissage model-based où le sujet a une représentation de la tâche et où ses choix sont déterminés sur la base de cette représentation interne.

La modification que nous avons apporté à la tâche la rend plus simple. Ceci a ainsi permis l'émergence de comportements qui ne sont pas ceux que nous escomptions pour certains sujets. La modélisation du comportement par des algorithmes d'apprentissage model-free serait techniquement tout à fait possible, mais ne serait pas légitime au vu des comportements individuels.

Il est important de noter que cette limite n'apparaît que si l'on s'intéresse aux courbes d'apprentissage individuelles. A l'extrême, dans une tâche où chaque sujet aurait un comportement model-based, mais avec un temps nécessaire pour construire le modèle interne différent par sujet, donnerait une courbe d'apprentissage qui serait apparemment compatible avec un apprentissage model-free.

Le choix de la tâche doit donc être effectué avec beaucoup d'attention. Il ne suffit pas que la tâche ressemble à une tâche d'apprentissage par renforcement model-free pour que le comportement de tous les sujets relève de l'apprentissage par renforcement model-free. Les sujets peuvent en effet individuellement développer des stratégies alternatives d'apprentissage. Le niveau de difficulté de la tâche semble particulièrement importante dans notre cas. Une difficulté moins importante a laissé à certains sujets la possibilité de construire une représentation claire de la structure de la tâche et d'utiliser cette représentation pour un apprentissage model-based plus rapide que ce qui est possible en model-free.

### 7.1.2 Cohérence de deux tâches proches dans une même population

L'étude "Specific effect of a dopamine partial agonist on counterfactual learning : evidence from Gilles de la Tourette syndrome" présenté dans cette thèse conclut que l'aripiprazole préserve l'apprentissage simple, mais détériore des phénomènes inférentiels plus complexes comme l'apprentissage contrefactuel.

L'étude "Enhanced habit formation in Gilles de la Tourette syndrome" présenté en Annexe 1, a étudié dans la même population l'équilibre entre contrôle dirigé vers un but et contrôle habituel du comportement avec et sans traitement antipsychotique (principalement de l'aripiprazole). Cette étude conclut que le traitement antipsychotique corrige en partie l'excès de prépondérance du système de décision habituel qui caractérise les patients Gilles de la Tourette non traités. L'imagerie indique un lien entre une plus grande prépondérance du système habituel et un couplage plus important entre striatum dorsal (putamen postérieur / caudé antérieur) et le cortex moteur.

Ces deux résultats apparemment contradictoires sont-ils cohérents et explicables ?

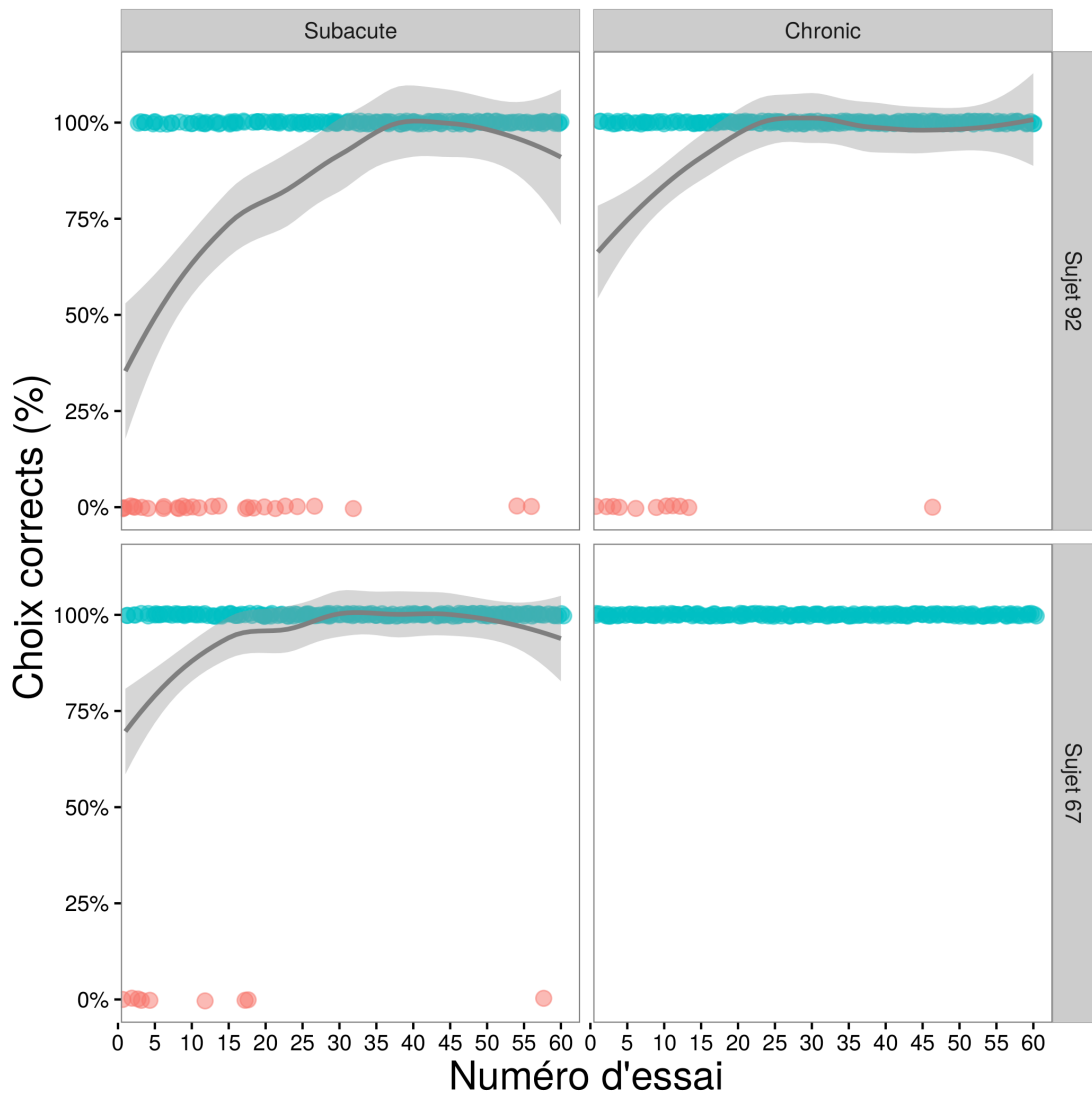


FIGURE 7.1 – Comparaison du comportement essai par essai de deux sujets dans une tâche probabilistique de sélection de stimulus. Chaque point correspond à une réponse individuelle pour un essai donné. Les points verts représentent des réponses correctes, les points rouges des réponses incorrectes. La courbe noire représente un lissage de la courbe moyenne d'apprentissage pour un sujet (haut = sujet 92, bas = sujet 67), pour les trois sessions à l'intérieur d'une visite (gauche = première visite, droit = dernière visite).

Il est possible de proposer une explication. Le syndrome Gilles de la Tourette est associé à une anomalie de la transmission dopaminergique à la fois sous-corticale et corticale (Albin et Mink, 2006 ; Segura et Strafella, 2013). Les symptômes seraient en partie la conséquence d'un couplage excessif entre striatum dorsal et cortex moteur (Wit et al., 2012) qui favoriserait le glissement précoce d'un mode de décision dirigé vers un but vers un comportement habituel (Schultz, 2013). Le traitement antipsychotique pourrait alors régulariser partiellement ce couplage excessif et diminuer en conséquence la prépondérance du contrôle habituel des actions.

Les bases de l'apprentissage contrefactuel sont encore peu précisément comprises mais font probablement intervenir spécifiquement les aires préfrontales, les zones proposées étant le cortex préfrontal dorso-médian et fronto-polaire (Boorman et al., 2011), le cortex fronto-polaire (Koechlin, 2014) ou le cortex orbito-frontal (Camille et al., 2004 ; Sommer et al., 2009). L'aripirazole, initialement présenté comme ayant un effet pro-cognitif, a plus tard démontré un effet de diminution du métabolisme dans le cortex frontal (Bolstad et al., 2015 ; Handley et al., 2013 ; Kim et al., 2013). Le traitement par aripirazole pourrait alors diminuer la capacité des sujets à utiliser le raisonnement contrefactuel sous-tendu par des aires préfrontales.

Cette explication permet d'interpréter les résultats de ces deux paradigmes de façon cohérente. Cette explication est-elle possible ? Oui. Est-elle satisfaisante ? Pas vraiment. La première raison est que notre étude ne comportait pas de mesure des corrélats cérébraux. Toute interprétation de l'effet de l'intervention sur les circuits sous-jacents est donc très hypothétique. La seconde raison est que ce raisonnement est formulé de façon ad-hoc en utilisant des résultats de la littérature qui ne sont pas consensuels. Il est probable qu'une explication aurait pu être trouvée si le phénomène inverse avait été observé. Nous n'avons pas les connaissances théoriques suffisantes pour dire de façon affirmative quelles aires cérébrales sous-tendent chacune des deux tâches utilisées. Nous n'avons pas de certitude sur les particularités du fonctionnement cérébral des patients Gilles de la Tourette. Interpréter l'effet d'un challenge pharmacologique dans ces conditions est difficile.

## 7.2 Nécessité de connaissances fondamentales supplémentaires

La difficulté à interpréter les résultats d'études portant sur des paradigmes d'apprentissage par renforcement avec intervention pharmacologique en population clinique nous incite à nous poser une question. Ces paradigmes sont-ils trop complexes dans l'état actuel des connaissances ?

### 7.2.1 Sur notre connaissance des tâche

Le premier constat est que nous ne connaissons pas suffisamment les aires cérébrales sous-tendant les fonctions cognitives en jeu dans les tâches que nous proposons, leur équilibre en termes de connectivité et l'influence des neuromédiateurs modulant leurs actions.

Prenons l'exemple du contrôle du comportement par le système model-based versus model-free. La présentation de ses mécanismes dans la partie introductive pouvait laisser l'impression de connaissances s'accumulant progressivement, des régions à l'œuvre spécifiquement dans les tâches model-based étant progressivement identifiées comme le cortex préfrontal ventro-médian (O'Doherty et al., 2015) ou le cortex dorso-médian et le cortex fronto-polaire (Doll et al., 2015).

Cependant, le tableau n'est pas si favorable.

Premièrement, le fait que les éléments de base de l'apprentissage par renforcement engagent principalement des structures sous-corticales comme le striatum, qui garde une certaine similitude d'architecture et de fonction entre espèces, a probablement permis son identification plus rapide chez l'homme. À l'inverse, les structures corticales et particulièrement préfrontales sont différentes d'une espèce à l'autre. Ainsi, des études chez le rat montrent qu'une lésion dans le striatum dorso-létéral ou dans le cortex infralimbique détériore l'apprentissage habituel, tandis qu'une lésion dans le striatum dorso-médian ou dans le cortex prélimbique détériore l'apprentissage dirigé vers un but (Dolan et Dayan, 2013). Or, il n'existe pas de consensus sur ce qui constitue l'équivalent du cortex infralimbique et prélimbique du rat chez le primate ou l'homme. Sur la base de la similitude des afférences thalamiques, des auteurs ont suggéré que le cortex infralimbique correspondrait au cortex cingulaire antérieur pré-génual chez l'homme et le cortex prélimbique du rat au cortex cingulaire antérieur sub-génual, mais cette correspondance reste discutée (Gass et Chandler, 2013). L'avancée dans la connaissance des bases cérébrales des fonctions d'apprentissage complexes chez l'homme, comme l'apprentissage model-based, bénéficie donc moins facilement des résultats obtenus chez l'animal.

Deuxièmement, les choix model-based incluent une activité prospective de simulation de différents scénarios ou chemins possibles. Doll et collaborateurs le montrent très élégamment en utilisant, en IRM fonctionnelle, une tâche séquentielle d'apprentissage par renforcement dans laquelle les chemins contenaient des catégories d'objets décodables en IRM (visages, outils, parties du corps, paysages) (Doll et al., 2015). Cette simulation mentale a pour but d'obtenir une évaluation des valeurs des différentes actions possibles. Mais de ce fait, la nature des simulations mentales est très dépendante du type de tâche proposée, et peut être de la navigation spatiale, des calculs de probabilité de transition, de la recherche en arborescence, etc. . . L'identification des bases cérébrales de l'apprentissage model-based est donc probablement rendue plus difficile par la grande variabilité des activités cérébrales suscitées en fonction de la tâche particulière d'apprentissage proposée au sujet (Lee,

2013).

Troisièmement, l'équilibre entre système d'apprentissage model-based et model-free était décrit en introduction de façon relativement simple. La force de connexion entre le cortex moteur et le putamen postérieur prédirait la tendance au contrôle model-free du comportement, tandis que la force de connexion entre le cortex préfrontal ventro-médian et le noyau caudé prédirait la tendance au contrôle model-based du comportement (Wit et al., 2012). Cette formulation donne une impression d'équilibre statique à caractériser chez un individu, qui resterait stable au cours du temps. Or cet équilibre pourrait être en fait variable au cours du temps de façon dynamique et rapide. Une version encore relativement simple serait qu'un arbitrage existe entre les deux systèmes en fonction de l'incertitude relative (Daw et al., 2005), de la fiabilité de leurs prédictions respectives (Lee et al., 2014), ou des besoins en précision versus rapidité du sujet (Keramati et al., 2011). Une version un peu moins simple propose une intégration des deux systèmes, le système habituel déterminant les buts et le système model-based déterminant la façon de les atteindre (Cushman et Morris, 2015). Une version plus complexe propose que ces deux systèmes sont intégrés, collaborent, et que la participation relative de chacun à la prise de décision est modifiée de façon dynamique au cours d'une tâche. C'est la proposition faite par exemple par Keramati et collaborateurs. Les humains utiliseraient de façon simultanée les deux systèmes, en simulant leur environnement jusqu'à un certain point dans le futur, exploitant ensuite les valeurs issues du système habituel pour les prédictions plus avant. Le fait d'accroître la pression temporelle pour obtenir une décision tendrait à diminuer la profondeur de la simulation et à augmenter le rôle des valeurs issues du système habituel de contrôle (Keramati et al., 2016). Il ne suffirait plus alors de parler de connectivité accrue dans tel circuit pour expliquer une prépondérance *globale* d'un système sur l'autre chez un sujet comparé à un autre. Il deviendrait nécessaire de prendre en compte l'équilibre entre les deux systèmes au cours du temps chez un même sujet, de caractériser leur interrelation, de déterminer les circuits ajustant de façon dynamique et rapide leur contribution relative, et de comprendre les perturbations possible de ce système complexe. Nous sommes donc loin d'avoir précisé de façon stable et consensuelle les circuits cérébraux sous-tendant l'équilibre entre système model-based et model-free et leur articulation.

Nous pourrions nous rassurer en considérant que les données semblent solides et claires au moins pour ce qui concerne les bases cérébrales des étapes fondamentales de l'apprentissage par renforcement. Il semble ainsi clairement établi que l'activité phasique des neurones dopaminergiques du tronc cérébral encode l'erreur de prédiction, que cet encodage prend en compte la magnitude de la récompense et la probabilité de la récompense, ce qui est conforme à ce que propose l'implémentation algorithmique de l'apprentissage par renforcement (Fiorillo et al., 2003 ; Hollerman et Schultz, 1998 ; Schultz et al., 1997 ; Sutton et Barto, 1998 ; Tobler et al., 2005).

Cependant, des résultats récents viennent rendre moins simples ces résultats fondamentaux. Chez l'homme, ces données étaient le plus souvent établies sur la base

d'études d'imagerie. Kishida et collaborateurs ont utilisé une technique permettant de mesurer les réponses dopaminergiques avec une résolution temporelle inférieure à la seconde, donc bien meilleure que ce que permet d'obtenir la mesure BOLD, dans une tâche d'apprentissage par renforcement. Ils concluent que les décharges phasiques de dopamine dans le striatum n'encodent pas l'erreur de prédiction de la récompense mais une agrégation d'erreur de prédiction de la récompense et d'erreur de prédiction contrefactuelle (Kishida et al., 2016). Ce résultat pose la question de la possibilité de décrire des briques simples d'apprentissage par renforcement sur lesquelles viendraient se greffer, selon la complexité de la tâche, des briques plus complexes. Il est possible que même les tâches simples fassent intervenir un grand nombre d'aires cérébrales. Pour Collins et Frank (2012), les deux systèmes d'apprentissage model-based et model-free sont à toujours à l'œuvre, y compris dans les tâches d'apprentissage instrumental les plus simples. Le décodage correct du fonctionnement cérébral au plan individuel dans ces tâches d'apprentissage par renforcement, même simples, pourrait donc nécessiter de prendre en compte de nombreux circuits cérébraux en interaction dynamique continue. On peut se poser la question de notre capacité à comprendre et décrire de façon explicite des systèmes complexes dans lesquels interagissent de façon dynamique plusieurs sous-systèmes. Un exemple de description de l'interaction dynamique de plusieurs sous-systèmes est proposé Figure 7.2.

## **7.2.2 Sur l'interaction des tâches avec la pathologie et les challenges pharmacologique**

Les tâches d'apprentissage par renforcement sont utilisées chez le sujet sain comme en population clinique, avec ou sans intervention pharmacologique. Ces deux facteurs interagissent pour complexifier l'interprétation de la tâche et de ses résultats.

### **7.2.2.1 Chez le sujet sain**

Envisageons tout d'abord l'utilisation d'interventions pharmacologiques chez le sujet sain pour caractériser l'effet de traitements. La caractérisation de l'effet des traitements qui en découle sera-t-elle généralisable à des populations cliniques ? Les antidépresseurs nous donnent à penser qu'il faut être prudent dans ce domaine.

Les études sur l'efficacité des antidépresseurs montrent que s'ils sont efficaces dans le traitement des dépressions modérées à sévères, ils n'ont pas un effet significativement différent du placebo pour les dépressions légères (Barbui et al., 2011 ; Fournier et al., 2010 ; Vöhringer et Ghaemi, 2011). Il semble exister à la fois des actions communes et des actions spécifiques des antidépresseurs en fonction du statut déprimé ou sain du sujet. Une méta-analyse s'est intéressée au traitement émotionnel dans la dépression. Elle a montré que l'administration répétée d'antidépresseurs avait pour effet une augmentation de la réponse du cortex préfrontal médian et de certaines



parties du système limbique comme le cortex cingulaire antérieur, l'amygdale et le thalamus aux émotions positives à la fois chez les patients déprimés et chez les sujets sains. En revanche, ces antidépresseurs augmentaient l'activité dans le cortex préfrontal dorso-létéral chez les sujets déprimés, mais diminuaient la réponse cérébrale aux émotions positives dans le nucleus accumbens, le putamen et le cortex préfrontal médian chez les sujets sains (Ma, 2015).

Des médicaments peuvent donc avoir des effets sur le cerveau que l'on peut caractériser chez le sujet malade comme chez le sujet sain, mais également des effets spécifiques qui résulteraient de l'action du médicament sur un état cérébral pathologique particulier, et qu'on ne pourrait donc pas caractériser par l'étude du sujet sain. La validité du sujet sain pour caractériser l'effet des traitements n'est donc pas parfaite.

### 7.2.2.2 Chez le sujet malade

Envisageons maintenant l'utilisation d'interventions pharmacologiques chez le sujet malade pour caractériser l'effet de traitements.

Les modèles physiopathologiques simples comme l'hyperdopaminergie striatale dans la schizophrénie montrent des insuffisances, et les modèles plus récents décrivent des interactions complexes entre système glutamatergique, système dopaminergique et système sérotoninergique, avec des boucles multiples d'interaction à la fois au niveau local et à grande distance (Stahl, 2013).

Pour la maladie Gilles de la Tourette, il n'en va pas différemment. Les mécanismes discutés dans l'article 1 inclus dans cette thèse, principalement une hyperdopaminergie à la fois dans le striatum et dans le cortex préfrontal (Albin et Mink, 2006 ; Segura et Strafella, 2013) sont probablement une simplification extrême des particularités cérébrales impliquées dans le syndrome Gilles de la Tourette. Le striatum reçoit des afférences de nombreuses zones corticales définissant des boucles cortico-basales multiples, qui dessinent des zones fonctionnelles distinctes à l'intérieur du striatum (Tremblay et al., 2015). L'expérimentation animale indique que des dysfonctionnements du contrôle inhibiteur à l'intérieur de sous-régions spécifiques du striatum sous-tendraient les symptômes comportementaux et moteurs dans cette pathologie (Figure 7.3) (Tremblay et al., 2015). L'expérimentation animale dispose dans ce domaine d'outils permettant de tester des hypothèses spécifiques bien plus précisément que chez l'homme. Pour reproduire et tester les hypothèses modernes sur la physiopathologie du syndrome Gilles de la Tourette, il est par exemple possible de rendre inopérant les interneurons cholinergiques du striatum (du moins 50% d'entre eux), et ce de façon très spécifique (Xu et al., 2015).

Dans l'article 1 inclus dans cette thèse, nous évoquons uniquement l'activité agoniste partielle dopaminergique D2 et l'activité agoniste 5HT1A pour rendre compte des résultats observés. Or l'aripiprazole a en fait une affinité pour bien d'autres

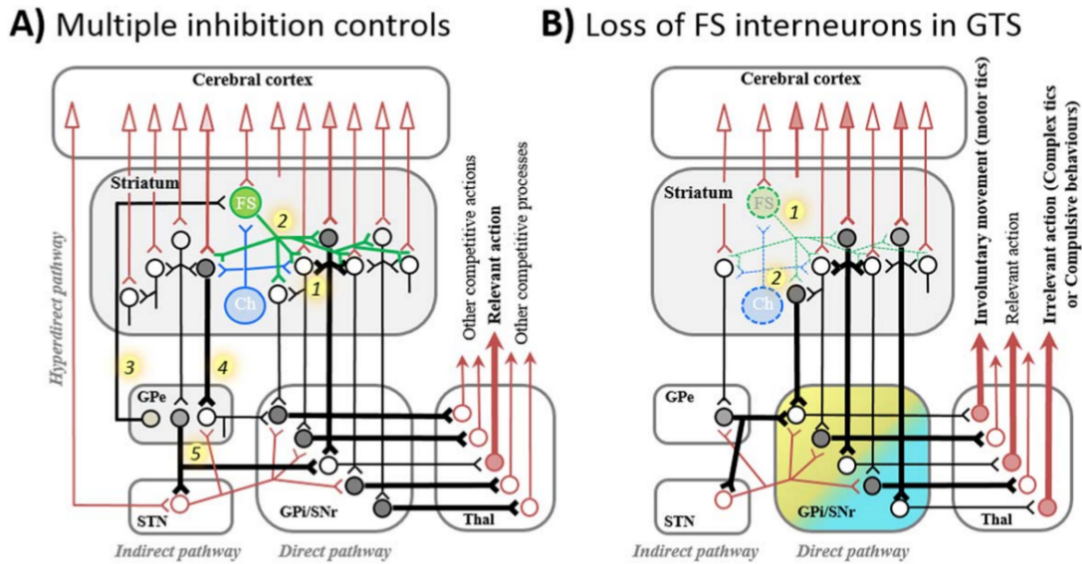


FIGURE 7.3 – Multiples niveaux dans le contrôle inhibiteur à l’intérieur des ganglions de la base et l’hypothèse physiopathologique récente du syndrome Gilles de la Tourette. (A) Représentation schématique de différents niveaux de contrôle inhibiteur à l’intérieur des ganglions de la base qui semblent impliqués dans la sélection d’information pour déterminer l’action appropriée en fonction du but recherché, tout en inhibant les options alternatives. (B) Des travaux récents ont montré que les patients GTS avaient un nombre moins important de neurones à parvalbumine et de neurones cholinergiques dans la partie antérieure du striatum comparé aux sujets sains. La production de mouvements involontaires, de tics complexes et de comportements compulsifs pourraient être expliqués par le nombre réduit de interneurons fast spiking (FS) qui sont de type interneurons à parvalbumine. (Tremblay et al., 2015).

récepteurs. Il a ainsi une affinité aussi forte pour les récepteurs D3 que D2, une affinité très haute pour les récepteurs 5HT2A, 2B et 7 (en plus de l'affinité 5HT1A déjà citée), une affinité modérée pour les récepteurs 5HT1D et 2C, et une affinité limitée pour les récepteurs 5HT1B, 3 et 6 (Shapiro et al., 2003). Est-il licite de ne pas prendre en compte ces affinités dans l'interprétation des résultats ? Probablement pas, car l'activité agoniste pour les récepteurs 5HT2C a un effet de diminution du relargage de la dopamine, notamment en région frontale. L'activité antagoniste 5HT2A a pour effet d'augmenter au contraire le relargage de la dopamine (Stahl, 2013). Il serait nécessaire de prendre en compte ces multiples affinités. Mais nous ne savons pas prédire et interpréter le résultat de l'action cumulée, et pondérée par les affinités réceptorielles différentes, d'une molécule comme l'aripiprazole sur le relargage des monoamines.

Lorsque la tâche d'apprentissage par renforcement est effectuée en population clinique avec intervention pharmacologique, son interprétation bute sur la complexité des phénomènes pathologiques et sur les conséquences complexes de l'intervention pharmacologique, et plus encore sur l'interaction de ces deux facteurs. Ainsi, la manipulation pharmacologique de la dopamine peut avoir des conséquences paradoxales pour des fonctions cognitives différentes, en fonction du niveau basal et optimal de dopamine dans différentes parties du cerveau (Cools et D'Esposito, 2011), et donc également lorsque ces niveaux de dopamine sont altérés par la pathologie.

### 7.2.3 Sur la validité des modèles pharmacologiques de pathologies psychiatriques

Nous avons décrit en introduction comment des substances psychotropes pouvaient être utilisées chez le sujet sain comme modèle pharmacologique de pathologie psychiatrique avec le développement de l'exemple de la kétamine comme modèle pharmacologique de psychose. Si ce modèle a convaincu un grand nombre de chercheurs, avec un nombre de publications conséquent dans cette utilisation, qu'en est-il de la validité de ce modèle ?

En termes de validité apparente, la kétamine n'est pas le seul produit à avoir été utilisé comme modèle pharmacologique de psychose. Si les trois types de symptômes présents dans la schizophrénie, positifs, négatifs et cognitifs ont été rapportés en termes d'effet de groupe avec la kétamine, (Krystal et al., 1994), l'administration de kétamine ne reproduit pas parfaitement l'état pathologique présent dans la schizophrénie chez un sujet donné.

La comparaison des effets de la kétamine avec d'autres substances utilisées comme modèle pharmacologique de psychose va dans ce sens. Carhart-Harris et collaborateurs ont comparé l'effet de cinq produits (le cannabis, la psilocybine, les amphétamines, la kétamine et l'alcool) dans leur capacité à induire des expériences subjectives précises, elles-mêmes corrélées à des symptômes psychiatriques (Carhart-Harris et al., 2013). Ils indiquent que la psilocybine semble être un meilleur modèle

des symptômes positifs. La kétamine induit en effet moins que la psilocybine des expériences de type saillance aberrante avec sentiment de centralité (par exemple : “Je pense que les choses que je vois et entends, que je ne remarquerais pas normalement, ont été mises ici pour me délivrer un message”) - mais plus que le cannabis ou les amphétamines. En revanche, la kétamine induit davantage que la psilocybine des troubles du cours de la pensée (par exemple : “Mes pensées s’arrêtent comme si elles étaient bloquées”) (Figure 7.4).

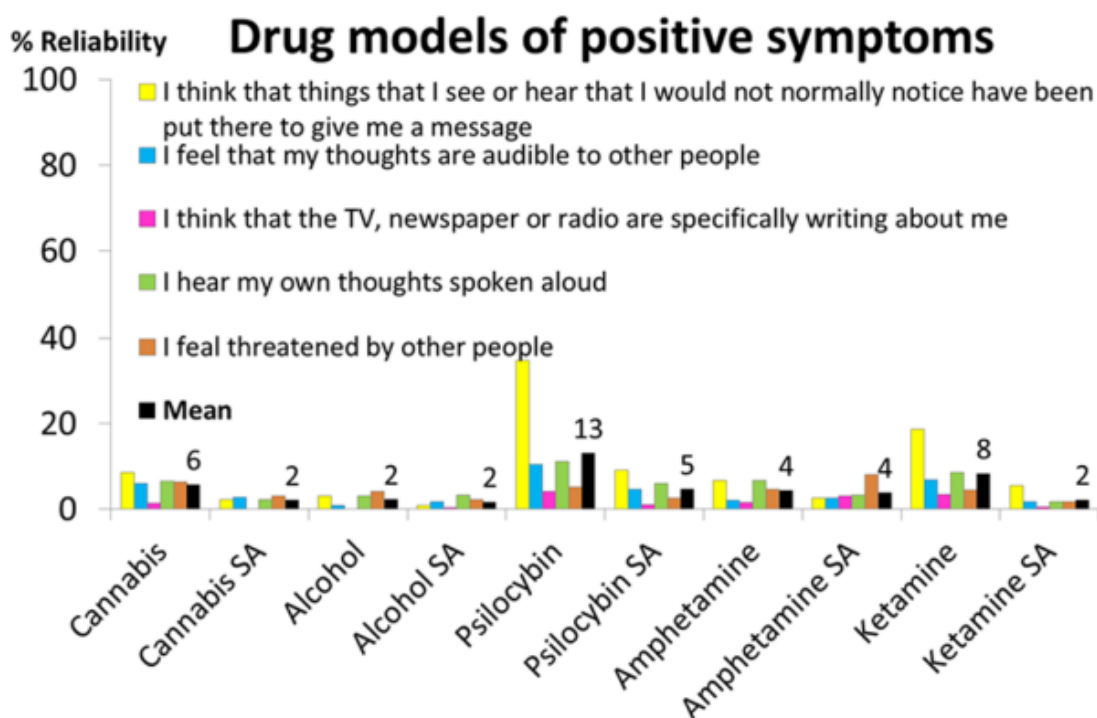


FIGURE 7.4 – Comparaison de la capacité de différentes substances psychotropes à déclencher des phénomènes subjectifs en lien avec des symptômes positifs. En ordonnée, pourcentage d'utilisateurs rapportant avoir eu “toujours ou souvent” l'expérience correspondante (légende couleur sur le graphique) pour chacune des drogues présentées. Enquête menée auprès de 224 utilisateurs expérimentés de drogues. SA = subacute, c'est-à-dire l'état obtenu une fois passée l'intoxication aigüe. L'association entre le groupe de symptômes positifs et les phrases à la première personne est le résultat d'une étude d'association effectuée auprès de 63 psychiatres et psychologues cliniciens. (Carhart-Harris et al., 2013)

Chaque drogue psychotrope utilisée chez l'homme induit donc un profil particulier de symptômes, et aucune ne reproduit parfaitement l'ensemble des symptômes de la maladie modélisée.

Il peut même exister d'apparentes incohérences dans le profil d'effets de certaines substances. Ainsi, une perfusion de faibles doses de kétamine chez le sujet sain augmente la prepulse inhibition (Abel et al., 2003) alors qu'elle déclenche des

symptômes compatibles avec des symptômes psychotiques et que la schizophrénie est associée à une diminution du prepulse inhibition (Greenwood, Light, et al., 2016).

En termes de validité mécanistique, la validité de la kétamine comme modèle de psychose débutante ou prodromale s'est vue renforcée par le développement récent de l'hypothèse que la physiopathologie de la schizophrénie est principalement en lien avec un hypofonctionnement des récepteurs NMDA (Howes et al., 2015 ; Kantrowitz et Javitt, 2010 ; Poels et al., 2014). Son administration a comme pour de nombreuses substances psychotropes une multitudes d'effets, avec des répercussions sur les autres systèmes de neurotransmission (Steeds et al., 2015), glutamatergiques, dopaminergiques et sérotoninergiques, qui entretiennent des influences réciproques (Stahl, 2013).

Comme dans le modèle animal, la validité apparente et mécanistique d'une intervention psychopharmacologique comme modèle de pathologie est limitée si l'on persiste à se placer au niveau de la pathologie syndromique. Le fait qu'aucun modèle pharmacologique ne puisse mimer tous les aspects d'une pathologie aussi complexe que la schizophrénie ne veut pas dire qu'ils ne peuvent pas être une aide précieuse pour la compréhension des mécanismes neurobiologiques sous-tendant des symptômes spécifiques (Curran et al., 2009). Il faut renoncer à essayer de déterminer quelle intervention pharmacologique constitue le meilleur modèle de la schizophrénie. Il est en revanche important d'essayer de comprendre quel aspect particulier de la pathologie mime chacune de ces substances et quel autre aspect elle ne mime pas pour avancer dans notre compréhension des mécanismes cérébraux qui sous-tendent l'expression de symptômes associés à la schizophrénie (Murray et al., 2013).

## **7.3 Apports pour la psychiatrie et limites de l'approche**

### **7.3.1 Apports théoriques pour la psychiatrie**

L'étude de l'apprentissage chez l'homme sain ou dans les pathologies psychiatriques n'est pas né avec la formalisation de l'apprentissage par renforcement. La neuropsychologie classique s'y est évidemment intéressée avec l'évaluation de dimensions nécessaires à l'apprentissage comme la mémoire de travail, la mémoire épisodique, ou des fonctions supérieures comme les fonctions exécutives (voir par exemple Minzenberg et al. (2009) pour une méta-analyse de 41 études d'imagerie fonctionnelle portant sur les fonctions exécutives dans la schizophrénie). Certaines de ces dimensions cognitives ont d'ailleurs été proposées comme des candidats endophénotypes, et des études de liaison de certains de ces candidats avec les

variants génétiques associés avec la schizophrénie ont été effectuées (pour une revue, voir Greenwood et al. (2007) et Greenwood, Lazzeroni, et al. (2016)).

L'étude de l'apprentissage par renforcement a un apport plus spécifique. Il permet une analyse de la façon dont le comportement s'adapte à chaque interaction avec l'environnement, après chaque renforcement. Les études anciennes sur l'apprentissage par renforcement ne tiraient pas entièrement parti de cette particularité et rapportaient des données sur le comportement moyen. En revanche, les études plus récentes se sont appliquées à comparer les prédictions de leurs modèles avec les données comportementales recueillies, essai par essai (Bayer et Glimcher, 2005; O'Doherty et al., 2003).

Cette nuance pourrait paraître anodine, pourtant le fait d'analyser la qualité de prédiction des modèles sur les données, en moyenne, plutôt que la qualité de prédiction des données moyennes ouvre des possibilités nouvelles.

L'analyse essai par essai permet de faire des hypothèses sur la façon dont les choix précédents et les renforcements vont déterminer le choix suivant. Par exemple dans l'apprentissage par renforcement model-free classique comme le formalise l'algorithme de Q-learning, les choix du sujet dans une tâche probabiliste de sélection de stimulus sont déterminés sur la base de la valeur de chaque stimulus, cette valeur étant calculée comme une forme de moyenne glissante des renforcements reçus dans le passé suite au choix de ce stimulus (Sutton et Barto, 1998). A l'inverse, dans un modèle d'apprentissage par renforcement de type model-based, les choix du sujet dans cette même tâche pourront être déterminés non seulement par l'histoire des renforcements reçus, mais également par des données de structure de la tâche, comme la représentation que les stimuli présentés fonctionnent par paire, ou qu'un choix modifie la valeur des options présentées à l'essai suivant. Il est donc possible de faire des hypothèses explicites sur les mécanismes cognitifs et les opérations réalisées par les sujets dans une tâche. Tester ces hypothèses se fait en déterminant quel modèle explique le mieux le comportement observé. Ceci peut se faire en comparant des modèles de complexité différente (Chen et al., 2015) ou à l'intérieur d'un même modèle en comparant les paramètres d'apprentissage permettant d'expliquer le mieux le comportement. On analyse ensuite si ces paramètres sont significativement différents entre deux groupes par exemple des sujets sains et des sujets présentant un trait symptomatique à étudier, ou entre un groupe de sujets symptomatiques traités et non traités. Cette comparaison de paramètres peut être effectuée en supposant un jeu de paramètres fixes par groupe (effet fixe) ou en supposant qu'un jeu de paramètres est à déterminer pour chaque sujet (effet aléatoire) (Daw, 2011).

De plus, cette analyse essai par essai fournit des estimations de paramètres et de grandeurs qui peuvent représenter des dimensions subjectives et non directement mesurables. Ces grandeurs sont obtenues en ajustant le modèle dans son implémentation computationnelle, ce qui fournit une estimation des différentes valeurs et paramètres du modèle à chaque essai. Il est ainsi possible d'estimer des quantités comme la surprise, les corrélats du regret, ou encore la confiance. La surprise

peut être représentée par l'erreur de prédiction, c'est-à-dire la différence entre la valeur attendue et la valeur effectivement reçue. Le regret a un corrélat cognitif, la pensée contrefactuelle, pour l'analyse de laquelle des paradigmes ont été créés et une formalisation algorithmique proposée, permettant d'obtenir des estimations de grandeurs comme l'erreur de prédiction contrefactuelle (Baskin-Sommers et al., 2016 ; Boorman et al., 2011 ; Coricelli et Rustichini, 2010). La confiance, elle aussi, peut être opérationnalisée comme une grandeur pouvant influencer les paramètres d'apprentissage et le fonctionnement des circuits cérébraux à l'œuvre au cours des tâches d'apprentissage par renforcement (Vinckier et al., 2016).

Enfin, le fait que ces modèles estiment des paramètres dont on pense qu'ils ont une représentation dans le cerveau permet de corréler les données estimées par ces modèles à des données d'imagerie fonctionnelle, une technique baptisée "model-based fMRI" (Huettel, 2012 ; O'Doherty et al., 2007), dont le principe est exposé Figure 7.5. La même technique a été utilisée avec l'électroencéphalographie (Larsen et O'Doherty, 2014). Il est donc possible de rechercher les bases neurales sous-tendant l'apprentissage et le comportement. Ceci va bien au-delà de l'analyse de type "activation" ou "réponse" faite dans les études d'imagerie traditionnelle (Chen et al., 2015).

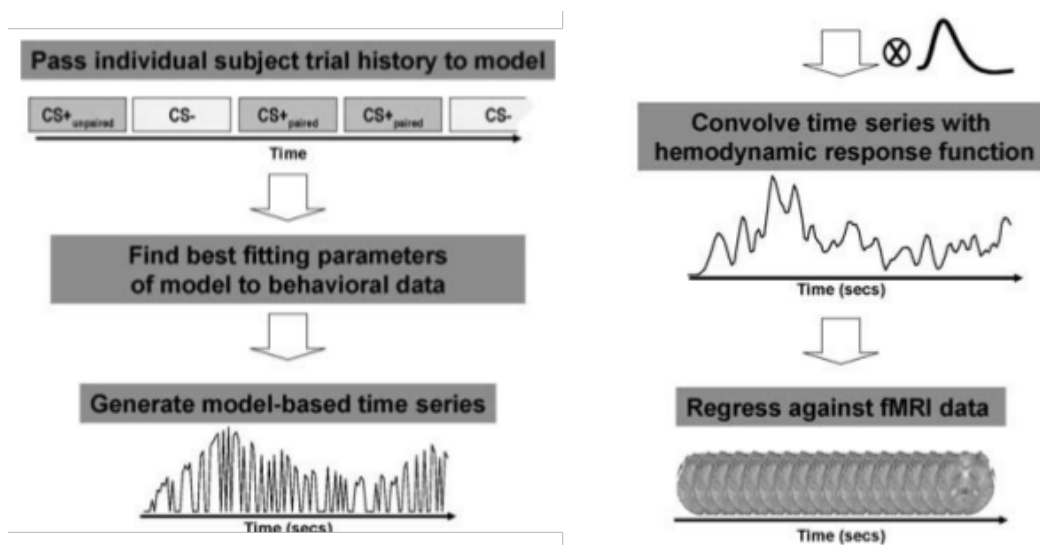


FIGURE 7.5 – Illustration du principe de l'IRM fonctionnelle model-based. Chaque essai est introduit dans le modèle, les paramètres du modèle sont ajustés de façon à minimiser les différences entre les prédictions du modèle et les mesures comportementales. Puis les paramètres du modèle permettant la meilleure adéquation entre les valeurs prédites et les valeurs observées sont utilisés pour générer une série temporelle pour chaque essai. Cette série peut contenir des estimations de valeurs, d'erreurs de prédiction, ou d'autres grandeurs subjectives estimées par le modèle. La série temporelle obtenue est convoluée avec la fonction de réponse hémodynamique pour prendre en compte le retard de la réponse BOLD, et régressée aux données de l'IRM fonctionnelle. (O'Doherty et al., 2007)

Tous ces avantages expliquent que l'analyse computationnelle de mécanismes d'apprentissage par renforcement est un champ de recherche très actif pour l'étude des pathologies psychiatriques (Huys et al., 2011 ; Maia et Frank, 2011 ; Montague et al., 2012 ; Stephan et Mathys, 2014 ; Teufel et Fletcher, 2016 ; Wiecki et al., 2015). Elle donne des outils qui semblent permettre de faire le lien entre mécanismes cognitifs, leur altération et les anomalies du fonctionnement cérébral sous-jacent. Elle pourrait théoriquement constituer une aide précieuse pour la recherche d'endophénotypes (Bogdan et al., 2013 ; Webb et al., 2016). Voyons si les apports de cette approche fournit des candidats qui respectent en pratique les critères d'un bon endophénotype, tels que nous les avons exposés au chapitre 3.

#### 7.3.2 Une caractérisation plus proche des mécanismes génétiques ?

Un des critères importants mis en avant par Gottesman et Gould est que l'architecture génétique sous-tendant le candidat endophénotype soit relativement plus simple que la pathologie étudiée (Gottesman et Gould, 2003). Nous avons vu que la structure génétique sous-tendant l'hérédité établie de certaines pathologies psychiatriques est complexe et hautement polygénique (Consortium et al., 2009), ce qui a été un frein dans les études d'association entre pathologies psychiatriques et variants génétiques. L'endophénotype, en fournissant une mesure supposée plus simple et plus proche de l'action des gènes, devrait ainsi permettre d'améliorer la puissance de ces études génétiques (Lenzenweger, 2013).

Est-ce que la caractérisation du comportement par l'évaluation neuro-cognitive de l'apprentissage par renforcement est réellement moins complexe que la maladie elle-même ?

En théorie oui, car la formalisation computationnelle du contrôle des actions par le système d'apprentissage par renforcement est parcimonieux et intègre le plus souvent un nombre de paramètres restreint, comme le taux d'apprentissage  $\alpha$  ou la température  $\beta$ . Il est possible de caractériser ces paramètres sur la base du comportement observé, et ce de façon quantitative. On obtient donc en théorie un nombre restreint de valeurs quantitatives décrivant des comportements apparemment disparates. Il est alors possible d'injecter ces valeurs quantitatives caractérisant des sujets individuels dans des études d'association génétique.

Cependant, si les comportements recueillis au cours des tâches d'apprentissage par renforcement semblent concorder de façon satisfaisante avec les courbes d'apprentissage théorique issues de la formalisation de l'apprentissage par renforcement, cela est davantage vrai au niveau collectif qu'au niveau individuel. Lorsque l'on s'intéresse au comportement au niveau individuel, ce qui est nécessaire si l'on veut effectuer des corrélations entre ces paramètres d'apprentissage issus de la modélisation du comportement et des données biologiques ou génétiques individuelles, la concordance est plus discutable. Nous l'avons vu avec la tâche d'apprentissage

probabiliste simple dans l'étude de l'effet de l'aripiprazole chez les patients Gilles de la Tourette.

La complexité des mécanismes à l'œuvre est également visible à partir des résultats préliminaires d'une étude de type tâche de Markov chez le sujet sain mesurant la propension de sujets à faire des choix guidés par l'évaluation model-based versus model-free. Le comportement d'un même sujet au cours de la tâche pouvait varier de façon significative, certains sujets commençant par un comportement model-free, puis développent un comportement model-based, avant de revenir finalement à un comportement model-free. La variabilité des comportements inter-sujet et au cours du temps pour un même sujet est illustrée Figure 7.6 par les résultats de cette expérimentation en cours.

Cette tâche illustre que les mécanismes cognitifs mis en œuvre par des sujets pour répondre à une tâche d'apprentissage par renforcement sont non seulement variables d'un individu à l'autre, mais peuvent également varier au cours du temps pour un même sujet.

Cette variabilité inter et intra-individuelle est également identifiable lorsque l'on observe le comportement essai par essai des sujets au cours de la passation de tâches aussi simples qu'une tâche de sélection de stimulus. Il n'est pas rare que l'on soit surpris de certains choix faits par des sujets sains ordinaires lorsque le renforcement est probabiliste, c'est-à-dire lorsque le fait de faire le bon choix n'amène la récompense que dans 70% ou 80% des cas, avec soit une absence de récompense soit une punition dans 20% des cas. Lorsque l'expérience est terminée, interroger le sujet sur ses "stratégies", sur la façon dont il a compris la tâche et sur la raison de certains choix, conduit à un constat clair : les sujets élaborent un nombre varié de modèles de la tâche et recherchent des corrélations contextuelles expliquant les essais pour lesquels ils n'ont pas reçu le renforcement positif. Certains sujets expliquent avoir remarqué que tel symbole leur donnait une récompense seulement s'il se trouvait à droite, ou qu'après 2 fois où ce symbole leur donnait une récompense, il devenait négatif. Ces corrélations sont rapportées spontanément en dépit du fait que d'une part la tâche est contrôlée de façon à ce qu'aucune corrélation de ce type n'existe, et d'autre part qu'il est dit explicitement aux sujets dans les instructions initiales que la position du symbole n'a pas d'importance et que les essais sont indépendants.

La simplicité apparente de la formulation algorithmique de l'apprentissage par renforcement, qui réduit un comportement complexe à un nombre réduit de paramètres, est donc probablement excessive. La plupart des publications ne présentent que des courbes d'apprentissage moyennes, et si ces courbes moyennes peuvent être modélisées par un apprentissage par renforcement simple, elles peuvent en réalité résulter d'une moyenne de courbes d'apprentissage individuelles traduisant des processus cognitifs à la fois variés au plan interindividuel, et variables pour un même sujet au cours du temps. En particulier, on retrouve fréquemment des stratégies d'élaboration de modèles de la tâche, des heuristiques qui prennent

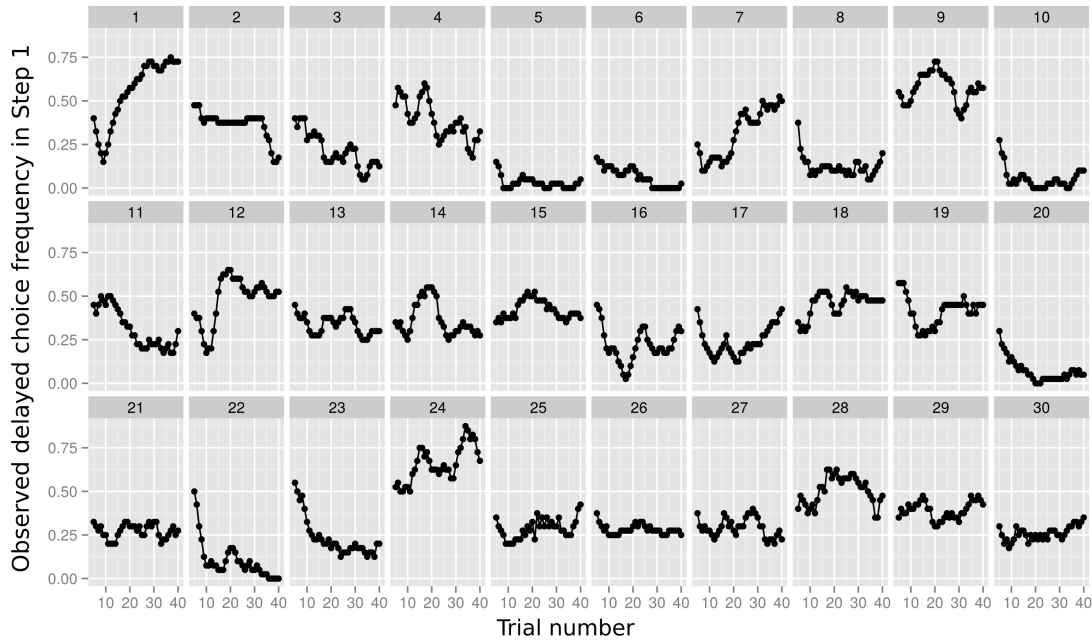


FIGURE 7.6 – Comportement variable entre sujets et, pour un même sujet, au cours du temps dans une tâche de Markov évaluant le comportement model-based versus model-free. Brièvement, ce paradigme est une tâche probabiliste de sélection de stimulus à deux niveaux. Au premier niveau, le choix des sujets amène une récompense probabiliste variable suivant le symbole, mais influence également de manière déterministe ou probabiliste la paire de symbole affichée au second niveau. Or les paires au second niveau n'ont pas les mêmes valeurs moyennes et les choix amenant les récompenses les meilleures au premier niveau peuvent être de mauvais choix car ils favorisent l'apparition des paires défavorables au second niveau, ce qui résulte en un gain global moins élevé que si un symbole moins avantageux avait été sélectionné au premier niveau, favorisant l'apparition des paires plus favorables au second niveau. L'ordonnée représente pour chacun des 30 sujets testés la proportion de choix model-based, c'est-à-dire correspondant à un choix ne maximisant pas le gain au premier niveau mais le gain global en prenant en compte les probabilités de transition entre les deux niveaux. Travail en cours.

parfois la forme de martingales douteuses en dépit d'instructions précises quant à l'absence d'influence de certains facteurs.

Ceci rejoint l'avertissement formulé par Glahn et al. (2014) qui indique que les candidats endophénotypes, qu'ils soient comportementaux, électrophysiologiques ou d'imagerie n'ont pas automatiquement une architecture génétique moins complexe que les pathologies psychiatriques associées. Bien que de nombreuses mesures cognitives ou d'imagerie soient fortement influencées par des facteurs génétiques, il est possible que cette héritabilité traduise l'effet cumulatif génétique global, mais ne révèle pas un effet génétique spécifique de gènes causaux (Glahn et al., 2014).

Ceci est illustré par la recherche des bases génétiques de l'autisme. Cette pathologie psychiatrique a une composante génétique maintenant clairement établie, avec une héritabilité attribuée à des variants communs estimée entre 40% et 60% (Klei et al., 2012). Cependant, aucun polymorphisme nucléotidique (Single Nucleotide Polymorphism ou SNP) n'a été associé de façon fiable aux troubles du spectre autistique, avec des faibles tailles d'effet retrouvés (Buxbaum, 2015). Une explication traditionnellement apportée était que l'hétérogénéité phénotypique dans cette catégorie de diagnostic expliquait les résultats décevants des études génétiques (Veatch et al., 2014). Des auteurs ont alors cherché à voir si le fait de réduire la variabilité phénotypique en stratifiant les sujets au sein d'une cohorte importante (la cohorte Simons, comportant 2576 familles) pouvait avoir un impact sur les tailles d'effet des études génétiques. Ils ont stratifié les phénotypes en onze sous-groupes, sur la base du diagnostic, de mesures d'intelligence (QI), de l'âge du début des troubles, de la sévérité de la maladie, et du profil des symptômes. Les résultats de cette étude sont tranchés : cette réduction de l'hétérogénéité phénotypique n'a pas eu d'impact significatif sur la puissance des associations génétiques (Chaste et al., 2015).

### **7.3.3 Une caractérisation spécifique d'une population clinique ?**

Un autre critère veut que l'endophénotype soit associé à la maladie, et soit davantage associé avec la maladie étudiée qu'avec d'autres troubles psychiatriques (Lenzenweger, 2013). Dans le cas des endophénotypes basés sur la caractérisation de l'apprentissage par renforcement, il est intéressant de discuter cette spécificité.

La modification de l'équilibre entre système de contrôle du comportement habituel versus dirigé vers un but retrouvée chez les patients présentant un syndrome de Gilles de la Tourette, par exemple, n'est pas un élément spécifique de la pathologie étudiée. En effet, une modification du contrôle du comportement avec une prépondérance du système habituel a été montré chez les patients présentant un trouble obsessionnel compulsif (Gillan et al., 2011) ou chez des patients présentant une dépendance à l'alcool (Sjoerds et al., 2013).

Si l'on reste dans le cadre de la classification actuelle des pathologies psychiatriques, ce candidat endophénotype ne respecte donc pas un critère important de son cahier des charges. Cependant, si l'on se place dans le cadre nouveau des Research Domain Criteria (RDoC), alors, on peut considérer qu'on recherche un trait neurocognitif qui a une base neurobiologique et qui serait transnosographique (Insel et al., 2010).

Dans notre exemple, l'addiction pourrait être considérée comme une compulsion comportementale (Everitt et Robbins, 2005) et pourrait partager de ce point de vue des processus physiopathologiques avec le trouble obsessionnel compulsif ou les tics complexes. L'évaluation et la caractérisation quantification de l'équilibre entre système de contrôle dirigé vers un but versus habituel par la formalisation computationnelle de l'apprentissage par renforcement viendrait caractériser de façon dimensionnelle un trait transnosographique.

Mais les travaux actuels restent encore souvent centrés sur une pathologie caractérisée par un diagnostic catégoriel classique. Si des paradigmes identiques sont parfois employés dans différentes populations, les différences de mise en œuvre des expériences par différents laboratoires ne permettent pas de comparer les résultats d'une étude à l'autre en termes quantitatifs. Pour que cela soit possible, il faudrait que soient testés par la même équipe et au sein d'un même dispositif expérimental des patients auxquels des diagnostics différents ont été apposés en termes de sémiologie psychiatrique. Cela n'est pour le moment que rarement fait.

L'utilisation de l'apprentissage par renforcement et de sa modélisation computationnelle ne peut délivrer son plein potentiel que s'il se fait en se plaçant dans le cadre des Research Domain Criteria (RDoC). Cela est techniquement possible d'autant plus que l'apprentissage par renforcement peut participer à définir de nouveaux axes qui permettront de classer les pathologies psychiatriques dans une nosographie dimensionnelle revisitée. Cette possibilité est illustrée Figure 7.7.

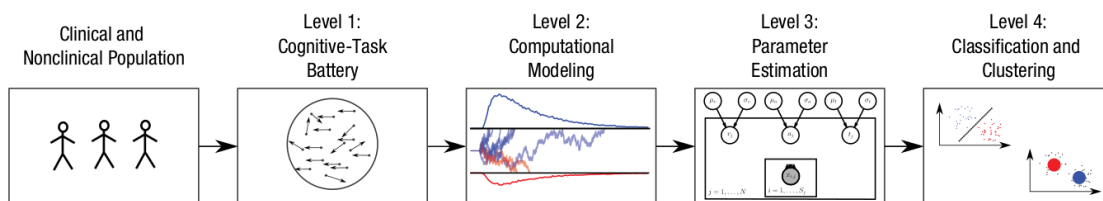


FIGURE 7.7 – Méthodologie permettant un apport de l'approche computationnelle au problème de classification en psychiatrique. Une population clinique et non clinique est testée sur une batterie cognitive. Une modélisation computationnelle est réalisée qui relie des données comportementales et des processus neurocognitifs. Les paramètres de ces modèles sont estimés (différentes méthodes sont possibles). Sur la base d'un profil multidimensionnel issu de l'estimation des paramètres computationnels, des algorithmes d'apprentissage, sont entraînés pour prédire l'état de maladie, découvrir des sous-groupes, ou relier des paramètres du modèle à des symptômes et à leur sévérité. (Wiecki et al., 2015)

### 7.3.4 Une évaluation complémentaire en population non clinique ?

De la même façon, les endophénotypes sont voués à être évalués également chez les apparentés de patients et dans la population générale, avec l'hypothèse que l'endophénotype est plus fréquent chez les sujets malades que chez les apparentés non malades et plus fréquent chez les apparentés non malades que dans la population générale (Lenzenweger, 2013).

Des candidats endophénotypes comme la reconnaissance faciale des émotions, ou la théorie de l'esprit, ont été étudiés dans différentes populations de patients ainsi que chez des apparentés (Bora et Pantelis, 2013 ; Cassetta et Goghari, 2014 ; Spilka et al., 2015). Au plan cognitif, l'évaluation neurocognitive au sens classique, comme l'évaluation de la mémoire, de l'attention soutenue ou partagée, l'inhibition de la réponse ont été étudiées chez les apparentés de sujets présentant des pathologies psychiatriques comme la dépression ou la schizophrénie (Besnier et al., 2009 ; Buoli et Caldiroli, 2015 ; Chkonia et al., 2010 ; Hou et al., 2016 ; Keshavan et al., 2010 ; Kristian Hill et al., 2015).

En revanche, peu d'études utilisant l'apprentissage par renforcement ont été effectuées dans ce sens. Des exemples existent cependant.

Une étude a recherché la présence d'une diminution de la sensibilité à la récompense chez les apparentés au premier degré de patients déprimés dans une tâche de discrimination visuelle renforcée de façon probabiliste (Liu et al., 2016). Les auteurs retrouvent une diminution de la sensibilité aux récompenses, mais seulement chez ceux des apparentés sains au premier degré qui ont des signes sub-syndromiques de dépression et non chez les apparentés qui n'ont pas de symptôme. Une autre étude a recherché des particularités dans le traitement des récompenses par les apparentés au premier degré de patients présentant un épisode dépressif caractérisé, à la fois sur le plan comportemental et au niveau cérébral par l'utilisation de l'IRM fonctionnelle (J Macoveanu et al., 2014). Les auteurs rapportent des particularités dans la réponse du cortex orbito-frontal au fait de manquer une récompense élevée après un choix peu risqué chez les apparentés de sujets présentant un épisode dépressif caractérisé comparé à des apparentés de sujets sains. De plus, cette particularité était corrigée par la prise d'escitalopram comparée à la prise de placebo (J Macoveanu et al., 2014). D'autres études s'intéressent aux endophénotypes cognitifs et d'imagerie cérébrale chez les apparentés au premier degré de patients atteints de trouble bipolaire (Miskowiak et al., 2017) ou à la prise de décision chez les apparentés au premier degré de personnes ayant complété un suicide (Ding et al., 2016). Des exemples existent également chez les enfants non malades mais considérés à risque de dépression du fait d'avoir une mère présentant un trouble dépressif récurrent, chez lesquels une étude a retrouvé une réponse cérébrale anormale à des anticipations de gains ou de perte, comparés à des enfants non à risque (Gotlib et al., 2010).

Pour la question du raisonnement contrefactuel, peu de travaux existent. Une étude s'est intéressée au raisonnement contrefactuel chez les apparentés du premier degré de sujets schizophrènes. Cependant elle n'a pas utilisé d'outil issu de l'apprentissage par renforcement, mais un autoquestionnaire à choix multiple, le "Counterfactual Inference Test" (CIT) testant la capacité des sujets à identifier et prendre en compte les éléments contrefactuel dans de petits scénarios (Albacete et al., 2016).

## 7.4 Approche translationnelle

L'expérimentation clinique chez l'homme, en particulier l'analyse de l'apprentissage par renforcement couplé à un challenge pharmacologique, que ce soit en population clinique ou chez le sujet sain, est donc une approche pleine de promesses. Elle n'est cependant pas dépourvue de limites et de difficultés.

Une difficulté particulière nous l'avons vu vient de la complexité des mécanismes neurobiologiques en jeu, qu'ils soient recrutés par la tâche d'apprentissage, perturbés par la pathologie ou modifiés par l'administration de psychotropes. Cette complexité devient inextricable lorsque ces trois éléments sont mis en jeu simultanément.

Certains auteurs restent optimistes et considèrent que l'insuffisance des connaissances actuelles et la difficulté à interpréter les résultats de tâches qui ne sont pas les plus simples ne doivent pas remettre en cause l'utilité des neurosciences cognitives pour repenser voir révolutionner les diagnostics et les traitements des pathologies psychiatriques (Hasler, 2012 ; Lee, 2013 ; Maia et Frank, 2011 ; Montague et al., 2012).

Les modèles animaux présentent également des limites mais restent indispensables. Ils ne sont pas seulement indispensables pour passer au crible de nouvelles molécules en grand nombre (Griebel et Holsboer, 2012). Ils sont également indispensables car ils disposent d'outils spécifiques permettant de répondre à des questions de façon beaucoup plus précise et affirmative. Les études chez l'homme sont en effet en majorité basées sur des techniques d'imagerie fonctionnelle qui ne donnent un reflet que très indirect des phénomènes à l'œuvre dans le cerveau. Elles ne séparent pas l'action des différents neuromédiateurs, ont une résolution temporelle médiocre et les méthodes d'interprétation statistique des résultats qu'elles fournissent font l'objet de nombreuses critiques (Bennett et al., 2011 ; Turner, 2016). Surtout elles ne permettent pas de conclure à l'implication causale, sauf dans certains cas extrêmes, lorsque des lésions circonscrites permettent de perturber spécifiquement un circuit donné, comme la neuropsychologie classique (méthode anatomo-clinique) a pu l'étudier pour certaines grandes fonctions. A l'inverse, les techniques disponibles chez l'animal pour déterminer les mécanismes cérébraux en jeu sont plus variées et plus spécifiques, avec des techniques comme l'enregistrement de neurone unique, la microdialyse, l'activation de réseaux spécifiques par l'optogénétique... Ils sont également indispensables car ils permettent de tester les conséquences de

manipulations génétiques ciblées avec la création de souris dont certains gènes spécifiques sont inactivés, de façon conditionnelle et tissu-spécifique, permettant de conclure à l'implication causale d'un sous-type récepteuriel par exemple. Les modèles animaux pourraient donc être utilisés pour préciser les mécanismes neurobiologiques sous-tendant des dimensions pathologiques pertinentes et non nécessairement circonscrites à la nosographie actuelle.

Une étude récente illustre cette utilisation. Le locus *GRIA1* est associé à la schizophrénie dans les études d'association pangénomiques, ce locus encodant le récepteur glutamatergique AMPA *GluA1*. Barkus et collaborateurs ont créé des lignées de souris knockout pour *GRIA1*. La caractérisation de leur comportement montre un défaut d'habituation à court terme. Dans certaines conditions, les souris voient leur attention pour des stimuli déjà présentés précédemment non pas diminuer, comme c'est le cas pour les souris type sauvage, mais augmenter (Barkus et al., 2014). Ce comportement est pour les auteurs un équivalent de salience aberrante, un mécanisme présenté par certains comme étant à la base des symptômes positifs de la schizophrénie (Kapur, 2003).

Les auteurs tirent également de leur étude une conclusion sur l'utilisation des modèles animaux :

“The real value of rodent models in the next decade is surely in this domain : not as models of schizophrenia per se but as experimental tools that can help link genomic discoveries to psychological processes and elucidate the underlying neural mechanisms.” (Barkus et al., 2014)

Il serait donc souhaitable de rendre plus complémentaires la recherche préclinique et la recherche clinique, non pas en tentant de complexifier les modèles animaux pour améliorer leur validité à représenter une pathologie psychiatrique complète, mais en profitant de la déconstruction en cours des diagnostics psychiatriques syndromiques pour injecter dans la recherche préclinique des candidats endophénotypes identifiés chez l'homme. Cette nouvelle compréhension de l'utilité des modèles animaux nécessite cependant d'avoir des tâches qui soient translatables ou rétro-translatables entre l'homme et l'animal.

#### **7.4.1 Translation et translation inverse**

La recherche translationnelle n'est pas un domaine récent, et l'adaptation de paradigmes animaux chez l'homme a déjà été proposée. Tout n'est pas adaptable, il semble par exemple difficile éthiquement de tenter un test de nage forcée ou de suspension par la queue chez l'homme. Cependant, de nombreux paradigmes animaux sont adaptables. Un exemple récent est l'adaptation d'une expérience classique chez l'animal, la préférence de place, utilisée cette fois chez l'homme avec administration d'amphétamines pour caractériser les mécanismes en jeu dans la pathologie addictive (Childs et Wit, 2009).

A l'inverse, partir de paradigmes définis chez l'homme et les adapter à l'animal a un avantage important. En effet, l'expérimentation chez l'homme rend possible le lien avec le rapport subjectif, le récit que l'individu peut faire de ses sensations, émotions et pensées. Ainsi, la possibilité d'utiliser la kétamine comme modèle pharmacologique de psychose est née du rapport subjectif de sujets sains (Krystal et al., 2002), que les cliniciens ont pu caractériser comme étant proches de symptômes positifs, avec par exemple des idées de référence, confinant aux idées délirantes (Pomarol-Clotet et al., 2006). La validité de ce modèle pharmacologique de psychose s'est trouvée renforcée par son utilisation chez des sujets psychotiques dont le rapport subjectif, en l'occurrence celui des patients et celui de leurs psychiatres traitants, indiquait que l'expérience vécue sous kétamine ressemblait de façon frappante aux symptômes vécus pendant les phases actives de leur maladie (Lahti et al., 2001).

Mais tous les paradigmes utilisés chez l'homme ne sont pas translatables chez l'animal. En effet, il est évident qu'on ne peut pas attendre qu'une souris nous rapporte un sentiment subjectif de regret ou son niveau de confiance dans le choix qu'elle vient d'effectuer, même si ces représentations pourraient avoir un équivalent dans cette espèce. Pour être utilisable chez l'animal, le paradigme doit mesurer un comportement objectif et mesurable.

De ce point de vue, l'étude de l'apprentissage par renforcement apporte une solution élégante. Elle permet d'opérationnaliser des dimensions subjectives, qu'elles fassent l'objet d'un rapport verbal par l'homme ou non, par la mesure objective du comportement dans des tâches adaptées. L'influence du regret dans la prise de décision peut être opérationnalisée par une tâche d'apprentissage contrefactuel (Baskin-Sommers et al., 2016 ; Boorman et al., 2011 ; Coricelli et Rustichini, 2010). Un sentiment de doute généralisé, sentiment subjectif rapportée par certains sujets dans des phases prodromales de la pathologie psychotique (Vinckier et al., 2016), peut être opérationnalisée dans une tâche d'apprentissage par renforcement probabiliste avec renversements. Sa modélisation computationnelle estimera une variable "confiance", dont il sera possible de rechercher l'influence sur les paramètres d'apprentissage ainsi que les corrélats cérébraux (Vinckier et al., 2016).

L'existence de ce trait d'union entre l'opérationnalisation et le rapport subjectif est un élément rassurant et pourrait être qualifiée de validité apparente, même si pour certains, le rapport subjectif pourrait ne pas être un rapport fiable des modifications cérébrales, mais un rapport influencé par le besoin de cohérence et la résolution de dissonance cognitive (Naccache et al., 2015). Pour autant, ce lien n'est pas requis au sein même de l'expérience pratiquée. Il peut avoir été fait à un autre moment, comme pour la validation de l'opérationnalisation du regret par l'apprentissage contrefactuel par exemple, ou la validation de la capacité de la kétamine à déclencher des symptômes subjectifs proches de la psychose débutante. Les expériences ultérieures peuvent ne pas rechercher ce rapport. Les expériences sous kétamine sont d'ailleurs bien souvent réalisées à des concentrations plasmatiques pour lesquelles les effets subjectifs ne sont pas ressentis par les sujets, afin de permettre des études en double aveugle. Dès lors, de tels paradigmes reposant

sur l'analyse du comportement pur pourraient être translaté chez l'animal.

Voyons en pratique à partir de l'apprentissage le plus simple, l'apprentissage model-free à partir de récompenses, si une telle translation peut se faire simplement ou si elle rencontre des obstacles et des contraintes. Ce type d'apprentissage peut s'analyser comme une suite de différentes étapes, chacune n'ayant pas nécessairement la même capacité à être implémentée de façon similaire chez l'homme et chez l'animal. Ces étapes incluent : le plaisir à la consommation de la récompense, l'anticipation des récompenses futures, la détermination des valeurs relatives, la détermination du coût ou des efforts nécessaires pour les obtenir, une dimension motivationnelle pour effectuer les actions nécessaires à l'obtention de ces récompenses et une dimension d'apprentissage afin de répéter les actions plus enclines à rapporter les récompenses les plus favorables dans le futur.

#### **7.4.1.1 La mesure de l'anticipation de la récompense et du plaisir à la réception de la récompense ne sont pas facilement mesurées de façon similaire entre l'homme et l'animal.**

Commençons par le plaisir à la consommation de la récompense et l'anticipation des récompenses futures. La distinction entre ces deux dimensions est importante car elles seraient sous-tendues par des circuits neuronaux distincts (Berridge et Robinson, 2003; Schultz, 2002), et dépendraient de neuromédiateurs différents. Le plaisir à la réception de la récompense serait sous-tendue par des mécanismes sérotoninergiques et opioïdes, alors que l'anticipation de la récompense serait sous-tendue par des mécanismes principalement dopaminergiques (Barbano et Cador, 2006).

Chez l'homme, la mesure de l'anticipation de la récompense se base essentiellement sur des questionnaires. La "Temporal Experience of Pleasure Scale" (TEPS), échelle d'auto-évaluation à 18 items, mesure à la fois le plaisir à la réception et le plaisir anticipé grâce à une échelle de Likert (Gard et al., 2006). La "Anticipatory and Consummatory Interpersonal Pleasure Scale" (ACIPS), autre échelle d'auto-évaluation, mesure également les deux dimensions avec une échelle de Likert, mais cette fois en examinant spécifiquement les situations d'interaction sociale (Gooding et Pflum, 2014).

Chez le rat, l'évaluation de l'anticipation de la récompense se fait grâce à des mesures comportementales. On habitue un rat à recevoir une récompense à une heure fixe dans la journée. On évalue l'anticipation par les modifications de comportement observées dans les moments qui précèdent le temps où la récompense habituelle lui sera délivré, comme l'augmentation de l'activité motrice (Mistlberger, 1994), ou l'augmentation des vocalisations de haute fréquence (Knutson et al., 2002).

Il n'y a donc pas de mesure commune chez l'homme et chez l'animal de l'anticipation de la récompense.

Chez l'homme, la mesure du plaisir à la réception de la récompense repose sur les deux mêmes échelles d'auto-évaluation décrites pour la mesure de l'anticipation de la récompense. Il existe des échelles supplémentaires recherchant des déficits de plaisir comme la "Snaith-Hamilton Pleasure Scale" (SHAPS) (Snaith et al., 1995) ou la "Fawcett-Clark Pleasure Scale" (Fawcett et al., 1983), ou la "Revised Chapman Physical Anhedonia Scale" (CPAS) (Chapman et al., 1976).

Chez le rat, l'évaluation du plaisir se base le plus souvent sur les tests de préférence au sucrose (Willner et al., 1987). Dans ce test, on mesure la préférence d'un animal pour une solution sucrée plutôt que pour de l'eau. Cette mesure n'a donc rien de commun à priori avec ce que l'on mesure chez l'homme. Pour vérifier si ce test rend compte d'une anhédonie, il a été utilisé chez l'homme. Mais les résultats sont décevants, car des patients présentant un épisode dépressif majeur, ne présentaient pas de particularité dans le test de préférence au sucre comparés à des sujets sains (Dichter et al., 2010).

### **7.4.1.2 La mesure de l'apprentissage des récompenses bénéficie de procédures qui sont utilisables chez l'homme comme chez l'animal.**

Chez l'homme, l'apprentissage des récompenses se mesure le plus souvent dans des tâches probabilistes de sélection de stimulus similaires à celle que nous avons utilisée dans le premier article rapporté dans cette thèse. Dans ces tâches, on présente aux participants successivement différentes paires de stimuli, et le sujet doit choisir à chaque fois un des deux stimuli présentés. On mesure la capacité d'un sujet à identifier progressivement et à choisir de façon consistante le stimulus le plus avantageux de chaque paire. Si des paires apportant des punitions sont incluses, on peut mesurer la capacité du sujet à apprendre des punitions et la comparer à sa capacité à apprendre des récompenses. Une tâche alternative est la "Response Bias Probabilistic Reward Task" (RBPRT). On demande aux sujets de dire si un stimulus ambigu est de type 1 ou 2, les sujets reçoivent un renforcement probabiliste lorsqu'ils ont correctement perçu et classé le stimulus (Pizzagalli et al., 2005). Cependant, sans que le sujet en soit informé, la probabilité de recevoir une récompense pour une réponse de type 1 est beaucoup plus élevée que pour une réponse de type 2. Les sujets sains développent alors un biais de réponse pour la réponse de type 1. On mesure l'influence du renforcement asymétrique sur le comportement du sujet par ce biais de réponse au détriment de la précision. Il est alors possible d'utiliser l'une de ces tâches pour comparer le comportement des sujets sains (Frank et al., 2007, 2004), des patients schizophrènes (Waltz et al., 2007), des patients parkinsoniens avec et sans traitement (Frank et al., 2004), des sujets sains sous agoniste dopaminergique ou antagoniste dopaminergique (Pessiglione et al., 2006).

Chez le rat, des tâches d'apprentissage par renforcement probabiliste proches existent depuis longtemps, mais leur implémentation présente nécessairement des différences avec les dispositifs utilisés chez l'homme. Des résultats concordants ont

été retrouvés entre l'homme et le rat dans la tâche de Response biais probabilistic reward, avec une diminution dans les deux espèces du biais de réponse entre le stimulus fortement renforcé et le stimulus faiblement renforcé dans deux challenges pharmacologiques : l'administration d'un agoniste dopaminergique (le pramipexole) et l'arrêt de l'exposition à la nicotine (Der-Avakian et al., 2013 ; Pergadia et al., 2014).

Récemment, un dispositif de test comprenant un écran tactile a été développé, permettant de faire passer à des souris et des rats des tâches d'apprentissage par renforcement dans des conditions et avec des stimuli visuels similaires à ce qui est utilisé chez l'homme (Horner et al., 2013) (Figure 7.8). Ce dispositif d'ores et déjà été utilisé dans une tâche chez le rat pour étudier les effets de l'administration d'un antagoniste des récepteurs 5-HT<sub>2C</sub> sur l'apprentissage par renversement visuel, et les auteurs montrent une diminution des persévérations avec cette intervention (Alsiö et al., 2015).

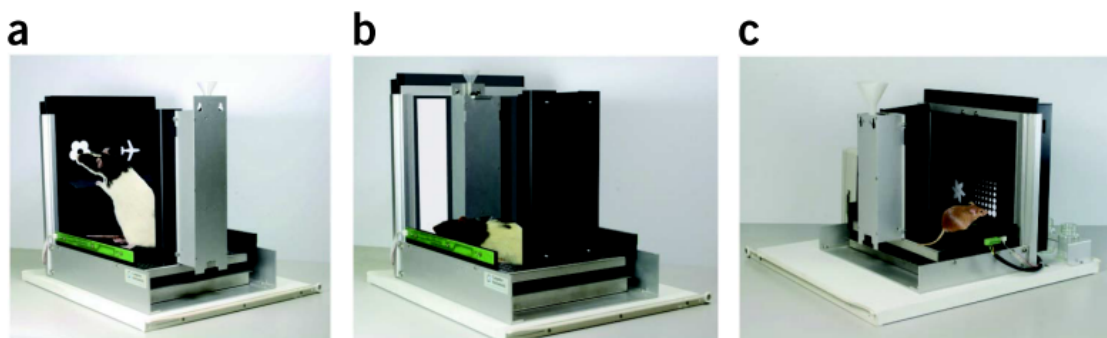


FIGURE 7.8 – Illustration du dispositif à écran tactile permettant l'étude de l'apprentissage par renforcement chez le rat et la souris dans des conditions proches de celles utilisées chez l'homme. (a) Dispositif utilisé chez le rat montrant une tâche d'apprentissage de localisation d'objets associés en paires. (b) Dispositif utilisé chez le rat montrant une tâche d'autoshaping. (c) Dispositif utilisé chez la souris montrant une tâche de discrimination visuelle. (Horner et al., 2013)

#### 7.4.2 Translation et mesures neurophysiologiques

L'adaptation à l'animal de tâches d'apprentissage par renforcement issues de l'expérimentation humaine n'est donc pas triviale et nécessite une analyse précise des mécanismes cognitifs à l'œuvre, étape par étape, pour vérifier si la translation est possible pour chacune d'entre elles (Der-Avakian et al., 2016). Mais au delà de la similitude apparente du dispositifs expérimental utilisé de façon translationnelle entre les deux espèces (validité apparente), il faut vérifier que la tâche recrute des circuits cérébraux proches dans les deux espèces (validité de construit).

Illustrons cette nécessité d'un exemple. L'étude des régions frontales impliquées dans l'apprentissage par renforcement par des récompenses a été mené simultanément

chez l'animal et chez l'homme. Les études chez le singe retrouvaient le plus souvent l'implication du cortex orbito-frontal (OFC), notamment pour coder des signaux de prédiction de récompense indépendamment de la modalité sensorielle ou des actions de l'animal (Padoa-Schioppa et Assad, 2006 ; Tremblay et Schultz, 1999). Les études chez l'homme en revanche attribuaient un rôle équivalent au cortex préfrontal ventro-médian (vmPFC) (Kahnt et al., 2010). Cependant, les études d'anatomie et de connectivité ont montré que l'OFC chez le singe n'était pas l'homologue du vmPFC chez l'homme, qu'il existait une homologie entre l'OFC de l'homme et du singe d'une part et entre le vmPFC de l'homme et du singe d'autre part (Wallis, 2012). Ces différences d'activation ont finalement été attribuées au fait que les tâches rapportées n'étaient pas suffisamment équivalents entre les deux espèces ; elles recrutaient des fonctions cognitives différentes et donc des circuits cérébraux différents (Wallis, 2012). Ces incohérences n'ont plus été constatées lorsque ces différences fines ont été prises en compte et corrigées (Wallis, 2012).

Les mesures électrophysiologiques ou d'imagerie, apportent donc des éléments permettant de valider ou au contraire de questionner la bonne translation de tâches entre l'homme et l'animal. Elles doivent permettre de vérifier que la tâche utilisée de façon translationnelle chez l'homme et l'animal recrute les mêmes circuits cérébraux à l'état normal (Der-Avakian et al., 2016). Les méthodes de mesure de l'activité intracérébrale étaient historiquement différentes entre expérimentation animale et humaine. Cependant l'imagerie par résonance magnétique ou l'enregistrement électroencéphalographique sont maintenant disponibles chez l'animal y compris pour les rongeurs, ce qui rend la comparaison des résultats plus aisée. Ces techniques ne sont bien évidemment pas toujours simples à mettre en place chez l'animal. Dans les tâches d'apprentissage en particulier, l'animal doit être réveillé et ses mouvements créent des artefacts. Certains auteurs développent néanmoins des méthodes pour apprendre à l'animal à ne pas bouger durant certaines séquences et rendre les résultats d'imagerie plus exploitables (Chen et al., 2012).



# Conclusion

La crise que travers l'innovation thérapeutique en psychiatrie semble en partie attribuable à la difficulté qui existe à confirmer les résultats expérimentaux positifs obtenus chez l'animal dans les essais cliniques chez l'homme.

Les modèles animaux sont une ressource importante pour les découvertes scientifiques. Ils permettent de modéliser de multiples aspects des pathologies psychiatriques. Ils souffrent de limites qui ne remettent pas en cause leur utilité en soi, mais qui interrogent la façon dont nous les utilisons.

Premièrement, si nous cherchons à améliorer la validité de modèles de pathologies psychiatriques en tentant de reproduire chez l'animal des tableaux syndromiques issus de la nosographie catégorielle et constitués d'un regroupement de symptômes de surface sans lien clair avec les mécanismes cérébraux sous-jacents, nous courrons à la déconvenue. Le modèle parfait n'existe pas et même si nous connaissions exactement quel ensemble de gènes déclenche la pathologie étudiée, la conséquence fonctionnelle de leur manipulation chez l'animal serait très différente de ce qu'on observe chez l'homme (Millan, 2008). Il est illusoire de chercher un modèle de pathologie psychiatrique "en soi" et préférable de travailler sur des modèles pour des dimensions particulières, potentiellement transnosographiques, de pathologies psychiatriques.

Deuxièmement, sélectionner des candidats médicaments sur la base de leur capacité à passer les tests animaux classiques revient à sélectionner des médicaments aux mêmes modes d'action que ceux que nous possédons déjà et qui ont fait la preuve de leurs insuffisances (Miczek, 2008). Les modèles classiques échouent dans la prédiction de l'efficacité chez l'homme de médicaments dont le mécanisme d'action est nouveau (Nestler et Hyman, 2010).

De nouveaux modèles sont donc nécessaires qui ne cherchent pas à mimer des catégories de diagnostic complètes, mais à approximer des éléments cardinaux présents de façon transnosographique en psychiatrie (Miczek, 2008).

L'expérimentation chez l'homme tout en n'étant pas exempte de difficultés et de limites, pourrait néanmoins fournir à la recherche animale de nouveaux modèles. En effet, le système de classification syndromique actuel semble un frein aux efforts de mise en relation des différents niveaux d'analyse des pathologies

psychiatriques, comportementales, neurobiologiques et génétiques. La recherche d'endophénotypes est une démarche qui vise à corriger cette difficulté, ils peuvent être issus des neurosciences cognitives. L'étude du comportement dans des tâches d'apprentissage par renforcement en particulier permet de faire des hypothèses sur les mécanismes à l'œuvre en termes cognitifs, d'estimer quantitativement des grandeurs potentiellement subjectives impliquées sous cette hypothèse grâce à l'analyse computationnelle, et de tester ces hypothèses et caractériser les bases cérébrales sous-tendant l'apprentissage et le comportement tout au long de la tâche en corrélant ces estimateurs aux données d'imagerie ou d'électrophysiologie. Elle pourrait donc fournir des candidats endophénotypes efficaces.

Cependant, l'interprétation des résultats de l'étude de l'apprentissage par renforcement chez l'homme est limitée notre connaissance encore imparfaite des mécanismes cérébraux qu'elles recrutent chez le sujet normal. Les modalités d'interaction complexes entre cortex préfrontal et ganglions de la base sont par exemple mal connus. Il est possible que même les tâches simples recrutent un nombre important de circuits cérébraux en interaction dynamique. L'utilisation de ces tâches dans des expériences en population clinique et avec intervention pharmacologique ajoute (ou plutôt multiplie par) deux niveaux de difficulté : la complexité des perturbations apportées par la pathologie psychiatrique, et la complexité liée à l'affinité pour une multiplicité de récepteurs de bon nombre des médicaments, avec les cascades d'interaction entre les différents systèmes de neurotransmetteurs qu'ils déclenchent. Nous sommes donc encore loin d'avoir atteint l'exigence de relative simplicité que doivent respecter les marqueurs pour espérer faciliter le lien entre le comportement, ses bases neurobiologiques et les bases génétiques qui les sous-tendent.

Un rapprochement des paradigmes entre recherche animale et humaine, par exemple dans le domaine de l'apprentissage par renforcement, pourrait être bénéfique à plusieurs titres.

Ces approches se complètent. D'un côté, l'expérimentation humaine permet l'exploration des liens entre comportements et circuits neurochimiques en ayant le bénéfice de pouvoir valider l'opérationnalisation comportementale des modèles par le rapport subjectif fait par les sujets. De l'autre, l'expérimentation animale permet plus aisément l'exploration des liens entre des modifications génétiques circonscrites, les circuits cérébraux impliqués et le comportement. Elle apporte des arguments plus forts, voire démontre l'implication causale de tel circuit ou sous-type de récepteur, grâce notamment aux multiples possibilités de manipulations génétiques et à la puissance des analyses spécifiquement utilisables chez l'animal comme les enregistrements de neurone unique ou l'optogénétique. Ce rapprochement favorise donc une accélération de la progression des connaissances fondamentales des circuits cérébraux, ici en lien avec l'apprentissage par renforcement.

De plus, en fournissant à l'expérimentation animale de nouveaux tests issus de l'expérimentation humaine, elle permettrait à l'innovation thérapeutique de sortir de la circularité des tests classiques de validation des candidats médicaments. Une

première étape serait que ces tests soient sélectionnés chez l'homme de façon à caractériser des dimensions pathologiques pertinentes et plus élémentaires, éventuellement transnosographiques, s'affranchissant ainsi des catégorisations syndromiques. Cette étape bénéficierait des spécificités de la recherche chez l'homme, comme l'accès au rapport subjectif, par exemple dans un modèle pharmacologique de pathologie, pour confirmer que ce que ressent l'individu correspond à ce qui est classiquement ressenti par les patients. Une seconde étape consisterait à opérationnaliser les tâches de façon à ce qu'elles ne reposent que sur une mesure objective du comportement et à s'assurer que leur réalisation est possible aussi bien chez l'homme que chez l'animal. Enfin, une troisième étape est l'analyse critique des différences obtenues entre espèces, en s'appuyant sur les mesures d'activité cérébrale. Cette étape doit envisager les conséquences en termes de recrutement de circuits cérébraux, de différences fines dans l'implémentation des tâches pour déterminer leur contribution aux différences d'activité observées entre espèces. On peut faire l'hypothèse que des processus qui seraient très similaires entre les différentes espèces auraient fait l'objet d'une pression de sélection forte au cours de l'évolution, et reposeraient donc sur des caractéristiques génétiques conservées au travers de ces espèces. A vrai dire ce n'est pas tant la démarche translationnelle inverse, de l'homme à l'animal, qui est décisive dans cette démarche, c'est le fait de déterminer des paradigmes expérimentaux les plus proches possibles entre l'homme et l'animal, correspondant chez l'homme à des dimensions pathologiques plutôt qu'à des entités nosographiques, et dans la mesure du possible (c'est-à-dire pas systématiquement) corroborées, au moins à un moment donné, par le récit subjectif que l'homme peut en faire, c'est-à-dire par la similitude entre ce récit et ce qui est considéré comme l'expression habituelle de cette dimension.

Une telle démarche est certes ambitieuse. Elle impose des changements d'habitudes et une collaboration entre chercheurs de différents domaines. Elle se heurte également à des contraintes financières. En effet, les paradigmes décrits sont plus complexes et coûteux à mettre en place que les tests utilisés habituellement en recherche préclinique. Cependant, ces efforts pour l'implémentation de ces nouveaux modèles animaux pour les pathologies psychiatriques devraient permettre, non seulement d'écarter plus tôt les médicaments qui n'ont pas d'action sur des dimensions pertinentes pour l'homme, mais surtout de faire émerger des possibilités thérapeutiques nouvelles sans les risques de les disqualifier par des tests inadaptés.



# Bibliographie

Abel, K. M., Allin, M. P. G., Hemsley, D. R., & Geyer, M. A. (2003). Low dose ketamine increases prepulse inhibition in healthy men. *Neuropharmacology*, 44(6), 729–737.

Abramson, L., & Seligman, M. (1977). Modeling psychopathology in the laboratory : History and rationale. *Psychopathology : Experimental Models*, 1–26.

Albacete, A., Contreras, F., Bosque, C., Gilabert, E., Albiach, Á., Menchón, J. M., ... Ayesa-Arriola, R. (2016). Counterfactual Reasoning in Non-psychotic First-Degree Relatives of People with Schizophrenia. *Frontiers in Psychology*, 7, 665. <https://doi.org/10.3389/fpsyg.2016.00665>

Albin, R. L., & Mink, J. W. (2006). Recent advances in Tourette syndrome research. *Trends Neurosci*, 29(3), 175–182. <https://doi.org/10.1016/j.tins.2006.01.001>

Aleman, A., & Vercammen, A. (2012). The Neuroscience of Hallucinations. In R. Jardri, A. Cachia, P. Thomas, & D. Pins (Eds.) (pp. 107–122). Springer.

Alsiö, J., Nilsson, S. R. O., Gastambide, F., Wang, R. A. H., Dam, S. A., Mar, A. C., ... Robbins, T. W. (2015). The role of 5-HT<sub>2C</sub> receptors in touchscreen visual reversal learning in the rat : a cross-site study. *Psychopharmacology*, 232(21-22), 4017–4031. <https://doi.org/10.1007/s00213-015-3963-5>

American Psychiatric Association, A. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*. (A. P. Association, Ed.) (5th ed.). American Psychiatric Publishing.

Anderson, N. L. (2010). The clinical plasma proteome : a survey of clinical assays for proteins in plasma and serum. *Clinical Chemistry*, 56(2), 177–185. <https://doi.org/10.1373/clinchem.2009.126706>

Antypa, N., Calati, R., & Serretti, A. (2014). The neuropsychological hypothesis of antidepressant drug action revisited. *CNS Neurol Disord Drug Targets*, 13(10), 1722–1739.

Balleine, B. W. (2005). Neural bases of food-seeking : affect, arousal and reward in corticostriatolimbic circuits. *Physiol Behav*, 86(5), 717–730. <https://doi.org/10.1016/j.physbeh.2005.08.011>

1016/j.physbeh.2005.08.061

Barbano, M. F., & Cador, M. (2006). Differential regulation of the consummatory, motivational and anticipatory aspects of feeding behavior by dopaminergic and opioidergic drugs. *Neuropsychopharmacology*, 31(7), 1371–1381. <https://doi.org/10.1038/sj.npp.1300908>

Barbui, C., Cipriani, A., Patel, V., Ayuso-Mateos, J. L., & Ommeren, M. van. (2011). Efficacy of antidepressants and benzodiazepines in minor depression : systematic review and meta-analysis. *The British Journal of Psychiatry*, 198(1), 11–6, sup1. <https://doi.org/10.1192/bjp.bp.109.076448>

Barkus, C., Sanderson, D. J., Rawlins, J. N. P., Walton, M. E., Harrison, P. J., & Bannerman, D. M. (2014). What causes aberrant salience in schizophrenia? A role for impaired short-term habituation and the GRIA1 (GluA1) AMPA receptor subunit. *Molecular Psychiatry*, 19(10), 1060–1070. <https://doi.org/10.1038/mp.2014.91>

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system : a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, 76, 412–427. <https://doi.org/10.1016/j.neuroimage.2013.02.063>

Baskin-Sommers, A., Stuppy-Sullivan, A. M., & Buckholz, J. W. (2016). Psychopathic individuals exhibit but do not avoid regret during counterfactual decision making. *Proceedings of the National Academy of Sciences of the United States of America*, 113(50), 14438–14443. <https://doi.org/10.1073/pnas.1609985113>

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, 47(1), 129–141. <https://doi.org/10.1016/j.neuron.2005.05.020>

Begley, C. G., & Ellis, L. M. (2012). Drug development : Raise standards for preclinical cancer research. *Nature*, 483(7391), 531–533. <https://doi.org/10.1038/483531a>

Belin, D., Belin-Rauscent, A., Murray, J. E., & Everitt, B. J. (2013). Addiction : failure of control over maladaptive incentive habits. *Current Opinion in Neurobiology*, 23(4), 564–572. <https://doi.org/10.1016/j.conb.2013.01.025>

Bell, R. F. (2009). Ketamine for chronic non-cancer pain. *Pain*, 141(3), 210–214. <https://doi.org/10.1016/j.pain.2008.12.003>

Bellman, R. (1957). *Dynamic programming*. Princeton Univ Press.

Belzung, C. (2014). Innovative drugs to treat depression : did animal models fail to be predictive or did clinical trials fail to detect effects? *Neuropsychopharmacology*, 39(5), 1041–1051. <https://doi.org/10.1038/npp.2013.342>

Belzung, C., & Lemoine, M. (2011). Criteria of validity for animal models of

psychiatric disorders : focus on anxiety disorders and depression. *Biology of Mood and Anxiety Disorders*, 1(1), 9. <https://doi.org/10.1186/2045-5380-1-9>

Bennett, C. M., Baird, A. A., Miller, M. B., & Wolford, G. L. (2011). Neural correlates of interspecies perspective taking in the post-mortem atlantic salmon : an argument for proper multiple comparisons correction. *Journal of Serendipitous and Unexpected Results*, 1, 1–5.

Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., & Kinderman, P. (2001). Persecutory delusions : a review and theoretical integration. *Clinical Psychology Review*, 21(8), 1143–1192.

Bentall, R. P., Kinderman, P., & Kaney, S. (1994). The self, attributional processes and abnormal beliefs : towards a model of persecutory delusions. *Behaviour Research and Therapy*, 32(3), 331–341.

Berna, F., Huron, C., Kazes, M., Offerlin-Meyer, I., Willard, D., Verry, P., ... Danion, J.-M. (2014). Chronic persecutory delusion and autobiographical memories in patients with schizophrenia : a diary study. *The Israel Journal of Psychiatry and Related Sciences*, 51(1), 25–33.

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *The Journal of Neuroscience*, 21(8), 2793–2798.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, 26(9), 507–513. [https://doi.org/10.1016/S0166-2236\(03\)00233-9](https://doi.org/10.1016/S0166-2236(03)00233-9)

Besnier, N., Richard, F., Zendjidjian, X., Kaladjian, A., Mazzola-Pomietto, P., Adida, M., & Azorin, J.-M. (2009). Stroop and emotional Stroop interference in unaffected relatives of patients with schizophrenic and bipolar disorders : distinct markers of vulnerability ? *The World Journal of Biological Psychiatry*, 10(4 Pt 3), 809–818. <https://doi.org/10.1080/15622970903131589>

Bespalov, A., Müller, R., Relo, A.-L., & Hudzik, T. (2016a). Drug Tolerance : A Known Unknown in Translational Neuroscience. *Trends in Pharmacological Sciences*, 37(5), 364–378. <https://doi.org/10.1016/j.tips.2016.01.008>

Bespalov, A., Steckler, T., Altevogt, B., Koustova, E., Skolnick, P., Deaver, D., ... Macleod, M. (2016b). Failed trials for central nervous system disorders do not necessarily invalidate preclinical models and drug targets. *Nature Reviews. Drug Discovery*, 15(7), 516. <https://doi.org/10.1038/nrd.2016.88>

Bloch, M. H., & Leckman, J. F. (2009). Clinical course of Tourette syndrome. *Journal of Psychosomatic Research*, 67(6), 497–501. <https://doi.org/10.1016/j.jpsychores.2009.09.002>

Blonk, M. I., Koder, B. G., Bemt, P. M. L. A. van den, & Huygen, F. J. P. M. (2010). Use of oral ketamine in chronic pain management : a review. *European*

*Journal of Pain*, 14(5), 466–472. <https://doi.org/10.1016/j.ejpain.2009.09.005>

Bogdan, R., Nikolova, Y. S., & Pizzagalli, D. A. (2013). Neurogenetics of depression : a focus on reward processing and stress sensitivity. *Neurobiology of Disease*, 52, 12–23. <https://doi.org/10.1016/j.nbd.2012.05.007>

Bolstad, I., Andreassen, O. A., Groote, I. R., Haatveit, B., Server, A., & Jensen, J. (2015). No difference in frontal cortical activity during an executive functioning task after acute doses of aripiprazole and haloperidol. *Frontiers in Human Neuroscience*, 9, 296. <https://doi.org/10.3389/fnhum.2015.00296>

Boorman, E. D., Behrens, T. E., & Rushworth, M. F. (2011). Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biology*, 9(6), e1001093. <https://doi.org/10.1371/journal.pbio.1001093>

Bora, E., & Pantelis, C. (2013). Theory of mind impairments in first-episode psychosis, individuals at ultra-high risk for psychosis and in first-degree relatives of schizophrenia : systematic review and meta-analysis. *Schizophrenia Research*, 144(1-3), 31–36. <https://doi.org/10.1016/j.schres.2012.12.013>

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis : an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91–121. <https://doi.org/10.1146/annurev-clinpsy-050212-185608>

Borsini, F., Evans, K., Jason, K., Rohde, F., Alexander, B., & Pollentier, S. (2002). Pharmacology of flibanserin. *CNS Drug Reviews*, 8(2), 117–142.

Bour. (2016). RetD : Servier se recentre sur 5 axes majeurs pour innover. Retrieved from <http://www.industrie.com/pharma/r-d-servier-se-recentre-sur-5-axes-majeurs-pour-innover>, 79232

Brandon, N. J., Millar, J. K., Korth, C., Sive, H., Singh, K. K., & Sawa, A. (2009). Understanding the role of DISC1 in psychiatric disease and during normal development. *The Journal of Neuroscience*, 29(41), 12768–12775. <https://doi.org/10.1523/JNEUROSCI.3355-09.2009>

Brenner, S. (2003). Nobel lecture : nature's gift to science. *Bioscience Reports*, 23(5), 225–237.

Broome, M. R., Johns, L. C., Valli, I., Woolley, J. B., Tabraham, P., Brett, C., ... McGuire, P. K. (2007). Delusion formation and reasoning biases in those at clinical high risk for psychosis. *The British Journal of Psychiatry*, 51, s38–s42. <https://doi.org/10.1192/bjp.191.51.s38>

Buckley, P. F., Miller, B. J., Lehrer, D. S., & Castle, D. J. (2009). Psychiatric comorbidities and schizophrenia. *Schizophrenia Bulletin*, 35(2), 383–402. <https://doi.org/10.1093/schbul/sbn135>

Buoli, M., & Caldiroli, A. (2015). IQ as a Cognitive Marker of Genetic Liability in Relatives of Schizophrenia Patients. *The American Journal of Psychiatry*, 172(8),

793–794. <https://doi.org/10.1176/appi.ajp.2015.15030387>

Buxbaum, J. D. (2015). DSM-5 and psychiatric genetics - round hole, meet square peg. *Biological Psychiatry*, 77(9), 766–768. <https://doi.org/10.1016/j.biopsych.2015.02.031>

Cade, J. F. J. (1949). Lithium salts in the treatment of psychotic excitement. *The Medical Journal of Australia*, 2(10), 349–352.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J.-R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science*, 304(5674), 1167–1170. <https://doi.org/10.1126/science.1094550>

Carhart-Harris, R. L., Brugger, S., Nutt, D. J., & Stone, J. M. (2013). Psychiatry's next top model : cause for a re-think on drug models of psychosis and other psychiatric disorders. *Journal of Psychopharmacology (Oxford, England)*, 27(9), 771–778. <https://doi.org/10.1177/0269881113494107>

Carmichael, S. T., & Price, J. L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *The Journal of Comparative Neurology*, 346(3), 366–402. <https://doi.org/10.1002/cne.903460305>

Carstensen, M., & Møller, A. M. (2010). Adding ketamine to morphine for intravenous patient-controlled analgesia for acute postoperative pain : a qualitative review of randomized trials. *British Journal of Anaesthesia*, 104(4), 401–406. <https://doi.org/10.1093/bja/aeq041>

Cassetta, B., & Goghari, V. (2014). Theory of mind reasoning in schizophrenia patients and non-psychotic relatives. *Psychiatry Research*, 218(1-2), 12–19. <https://doi.org/10.1016/j.psychres.2014.03.043>

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Lancet*, 374(9683), 86–89. [https://doi.org/10.1016/S0140-6736\(09\)60329-9](https://doi.org/10.1016/S0140-6736(09)60329-9)

Chamberlain, S. R., Müller, U., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2006). Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science*, 311(5762), 861–863. <https://doi.org/10.1126/science.1121218>

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1976). Scales for physical and social anhedonia. *Journal of Abnormal Psychology*, 85(4), 374–382.

Chaste, P., Klei, L., Sanders, S. J., Hus, V., Murtha, M. T., Lowe, J. K., . . . Devlin, B. (2015). A genome-wide association study of autism using the Simons Simplex Collection : Does reducing phenotypic heterogeneity in autism increase genetic homogeneity? *Biological Psychiatry*, 77(9), 775–784. <https://doi.org/10.1016/j.biopsych.2014.09.017>

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforce-

- ment learning in depression : A review of computational research. *Neuroscience and Biobehavioral Reviews*, 55, 247–267. <https://doi.org/10.1016/j.neubiorev.2015.05.005>
- Chen, G., Wang, F., Dillenburg, B. C., Friedman, R. M., Chen, L. M., Gore, J. C., ... Roe, A. W. (2012). Functional magnetic resonance imaging of awake monkeys : some approaches for improving imaging quality. *Magnetic Resonance Imaging*, 30(1), 36–47. <https://doi.org/10.1016/j.mri.2011.09.010>
- Childs, E., & Wit, H. de. (2009). Amphetamine-induced place preference in humans. *Biological Psychiatry*, 65(10), 900–904. <https://doi.org/10.1016/j.biopsych.2008.11.016>
- Chkonia, E., Roinishvili, M., Herzog, M. H., & Brand, A. (2010). First-order relatives of schizophrenic patients are not impaired in the Continuous Performance Test. *Journal of Clinical and Experimental Neuropsychology*, 32(5), 481–486. <https://doi.org/10.1080/13803390903201777>
- Cipriani, A., Furukawa, T. A., Salanti, G., Geddes, J. R., Higgins, J. P., Churchill, R., ... Barbui, C. (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants : a multiple-treatments meta-analysis. *Lancet*, 373(9665), 746–758. [https://doi.org/10.1016/S0140-6736\(09\)60046-5](https://doi.org/10.1016/S0140-6736(09)60046-5)
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur J Neurosci*, 35(7), 1024–1035. <https://doi.org/10.1111/j.1460-9568.2011.07980.x>
- Coltheart, M. (2007). Cognitive neuropsychiatry and delusional belief. *Quarterly Journal of Experimental Psychology*, 60(8), 1041–1062. <https://doi.org/10.1080/17470210701338071>
- Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, 62, 271–298. <https://doi.org/10.1146/annurev.psych.121208.131622>
- Connors, M. H., & Halligan, P. W. (2015). A cognitive account of belief : a tentative road map. *Frontiers in Psychology*, 5, 1588. <https://doi.org/10.3389/fpsyg.2014.01588>
- Consortium, I. S., Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., ... Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 460(7256), 748–752. <https://doi.org/10.1038/nature08185>
- Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., & Pangalos, M. N. (2014). Lessons learned from the fate of AstraZeneca's drug

- pipeline : a five-dimensional framework. *Nat Rev Drug Discov*, 13(6), 419–431. <https://doi.org/10.1038/nrd4309>
- Cools, R., & D’Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*, 69(12), e113–e125. <https://doi.org/10.1016/j.biopsych.2011.03.028>
- Cools, R., Frank, M. J., Gibbs, S. E., Miyakawa, A., Jagust, W., & D’Esposito, M. (2009). Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *The Journal of Neuroscience*, 29(5), 1538–1543. <https://doi.org/10.1523/JNEUROSCI.4467-08.2009>
- Cools, R., Robinson, O. J., & Sahakian, B. (2008). Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharmacology*, 33(9), 2291–2299. <https://doi.org/10.1038/sj.npp.1301598>
- Cooper, J. A., Tucker, V. L., & Papakostas, G. I. (2014). Resolution of sleepiness and fatigue : a comparison of bupropion and selective serotonin reuptake inhibitors in subjects with major depressive disorder achieving remission at doses approved in the European Union. *Journal of Psychopharmacology*, 28(2), 118–124. <https://doi.org/10.1177/0269881113514878>
- Corbett, D., & Wise, R. A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain : a moveable electrode mapping study. *Brain Research*, 185(1), 1–15.
- Coricelli, G., & Rustichini, A. (2010). Counterfactual thinking and emotions : regret and envy learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1538), 241–247. <https://doi.org/10.1098/rstb.2009.0159>
- Coricelli, G., Critchley, H. D., Joffily, M., O’Doherty, J. P., Sirigu, A., & Dolan, R. J. (2005). Regret and its avoidance : a neuroimaging study of choice behavior. *Nature Neuroscience*, 8(9), 1255–1262. <https://doi.org/10.1038/nn1514>
- Corlett, P. R., Honey, G. D., & Fletcher, P. C. (2007). From prediction error to psychosis : ketamine as a pharmacological model of delusions. *J Psychopharmacol*, 21(3), 238–252. <https://doi.org/10.1177/0269881107077716>
- Corlett, P. R., Honey, G. D., Aitken, M. R. F., Dickinson, A., Shanks, D. R., Absalom, A. R., ... Fletcher, P. C. (2006). Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine : linking cognition, brain activity, and psychosis. *Arch Gen Psychiatry*, 63(6), 611–621. <https://doi.org/10.1001/archpsyc.63.6.611>
- Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Prog Neurobiol*, 92(3), 345–369. <https://doi.org/10.1016/j.pneurobio.2010.06.007>
- Correll, C. U., Kishimoto, T., & Kane, J. M. (2011). Randomized controlled trials

in schizophrenia : opportunities, limitations, and trial design alternatives. *Dialogues in Clinical Neuroscience*, 13(2), 155–172.

Costall, B., Naylor, R. J., & Nohria, V. (1979). Hyperactivity response to apomorphine and amphetamine in the mouse : the importance of the nucleus accumbens and caudate-putamen. *The Journal of Pharmacy and Pharmacology*, 31(4), 259–261.

Creese, I., Burt, D. R., & Snyder, S. H. (1976). Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science*, 192(4238), 481–483.

Cromwell, H. C., & Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *Journal of Neurophysiology*, 89(5), 2823–2838. <https://doi.org/10.1152/jn.01014.2002>

Cryan, J. F., Markou, A., & Lucki, I. (2002). Assessing antidepressant activity in rodents : recent developments and future needs. *Trends in Pharmacological Sciences*, 23(5), 238–245.

Cryan, J. F., Sánchez, C., Dinan, T. G., & Borsini, F. (2008). Animal and Translational Models for CNS Drug Discovery, Vol. 1 : Psychiatric Disorders. In R. A. McArthur & F. Borsini (Eds.) (1st ed., pp. 165–198). Academic Press.

Curran, H. V., D’Souza, D. C., Robbins, T. W., & Fletcher, P. (2009). Modeling psychosis. *Psychopharmacology*, 206(4), 513–514. <https://doi.org/10.1007/s00213-009-1663-8>

Cushman, F., & Morris, A. (2015). Habitual control of goal selection in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 112(45), 13817–13822. <https://doi.org/10.1073/pnas.1506367112>

Cuthbert, B., & Insel, T. (2010). The data of diagnosis : new approaches to psychiatric classification. *Psychiatry*, 73(4), 311–314. <https://doi.org/10.1521/psyc.2010.73.4.311>

Daw, N. D. (2011). Decision making, affect, and learning : Attention and performance XXIII. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision making, affect, and learning : Attention and performance XXIII* (Vol. 23, pp. 3–38). Oxford University Press Oxford.

Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, 15(4-6), 603–616.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*, 8(12), 1704–1711. <https://doi.org/10.1038/nn1560>

Dayan, P. (2012). Twenty-five lessons from computational neuromodulation. *Neuron*, 76(1), 240–256. <https://doi.org/10.1016/j.neuron.2012.09.027>

Dayan, P., & Abbott, L. F. (2005). *Theoretical Neuroscience : Computational and*

*Mathematical Modeling of Neural Systems (Computational Neuroscience Series)* (Revised ed.). The MIT Press.

Deakin, J. F. W., Lees, J., McKie, S., Hallak, J. E. C., Williams, S. R., & Dursun, S. M. (2008). Glutamate and the neural basis of the subjective effects of ketamine : a pharmaco-magnetic resonance imaging study. *Arch Gen Psychiatry*, 65(2), 154–164. <https://doi.org/10.1001/archgenpsychiatry.2007.37>

Delay, J., & Buisson, J. F. (1958). Psychic action of isoniazid in the treatment of depressive states. *Journal of Clinical and Experimental Psychopathology*, 19(2, Suppl. 1), 51–55.

Der-Avakian, A., Barnes, S. A., Markou, A., & Pizzagalli, D. A. (2016). Translational Assessment of Reward and Motivational Deficits in Psychiatric Disorders. *Current Topics in Behavioral Neurosciences*, 28, 231–262. [https://doi.org/10.1007/7854\\_2015\\_5004](https://doi.org/10.1007/7854_2015_5004)

Der-Avakian, A., D’Souza, M. S., Pizzagalli, D. A., & Markou, A. (2013). Assessment of reward responsiveness in the response bias probabilistic reward task in rats : implications for cross-species translational research. *Translational Psychiatry*, 3, e297. <https://doi.org/10.1038/tp.2013.74>

Detke, M. J., Rickels, M., & Lucki, I. (1995). Active behaviors in the rat forced swimming test differentially produced by serotonergic and noradrenergic antidepressants. *Psychopharmacology*, 121(1), 66–72.

Di, L., Rong, H., & Feng, B. (2013). Demystifying brain penetration in central nervous system drug discovery. Miniperspective. *Journal of Medicinal Chemistry*, 56(1), 2–12. <https://doi.org/10.1021/jm301297f>

Dichter, G. S., Smoski, M. J., Kampov-Polevoy, A. B., Gallop, R., & Garbutt, J. C. (2010). Unipolar depression does not moderate responses to the Sweet Taste Test. *Depression and Anxiety*, 27(9), 859–863. <https://doi.org/10.1002/da.20690>

Dickinson, A. (1985). Actions and habits : the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London B : Biological Sciences*, 308(1135), 67–78.

Ding, Y., Pereira, F., Hoehne, A., Beaulieu, M.-M., Lepage, M., Turecki, G., & Jollant, F. (2016). Altered brain processing of decision-making in healthy first-degree biological relatives of suicide completers. *Molecular Psychiatry*. <https://doi.org/10.1038/mp.2016.221>

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, 80(2), 312–325. <https://doi.org/10.1016/j.neuron.2013.09.007>

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience*, 18(5),

767–772. <https://doi.org/10.1038/nm.3981>

Dollfus, S., Alary, M., & Razafimandimby, A. (2012). The Neuroscience of Hallucinations. In R. Jardri, A. Cachia, P. Thomas, & D. Pins (Eds.) (pp. 123–136). Springer.

Domino, E. F., Zsigmond, E. K., Domino, L. E., Domino, K. E., Kothary, S. P., & Domino, S. E. (1982). Plasma levels of ketamine and two of its metabolites in surgical patients using a gas chromatographic mass fragmentographic assay. *Anesth Analg*, 61(2), 87–92.

Dournes, C., Beeské, S., Belzung, C., & Griebel, G. (2013). Deep brain stimulation in treatment-resistant depression in mice : comparison with the CRF1 antagonist, SSR125543. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 40, 213–220.

Drevets, W. C., Price, J. L., Simpson, J. R., Todd, R. D., Reich, T., Vannier, M., & Raichle, M. E. (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature*, 386(6627), 824–827. <https://doi.org/10.1038/386824a0>

Ducottet, C., Griebel, G., & Belzung, C. (2003). Effects of the selective nonpeptide corticotropin-releasing factor receptor 1 antagonist antalarmin in the chronic mild stress model of depression in mice. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27(4), 625–631. [https://doi.org/10.1016/S0278-5846\(03\)00051-4](https://doi.org/10.1016/S0278-5846(03)00051-4)

D’Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, 319(5867), 1264–1267. <https://doi.org/10.1126/science.1150605>

D’Souza, D. C., Singh, N., Elander, J., Carbutto, M., Pittman, B., Haes, J. U. de, ... Schipper, J. (2012). Glycine transporter inhibitor attenuates the psychotomimetic effects of ketamine in healthy males : preliminary evidence. *Neuropsychopharmacology*, 37(4), 1036–1046. <https://doi.org/10.1038/npp.2011.295>

Ebisch, S. J. H., & Aleman, A. (2016). The fragmented self : imbalance between intrinsic and extrinsic self-networks in psychotic disorders. *The Lancet Psychiatry*, 3(8), 784–790. [https://doi.org/10.1016/S2215-0366\(16\)00045-6](https://doi.org/10.1016/S2215-0366(16)00045-6)

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction : from actions to habits to compulsion. *Nature Neuroscience*, 8(11), 1481–1489. <https://doi.org/10.1038/nm1579>

Fagiolini, A., Rocca, P., De Giorgi, S., Spina, E., Amodeo, G., & Amore, M. (2016). Clinical trial methodology to assess the efficacy/effectiveness of long-acting antipsychotics : Randomized controlled trials vs naturalistic studies. *Psychiatry Research*, 247, 257–264. <https://doi.org/10.1016/j.psychres.2016.11.044>

Faravelli, C., Castellini, G., Benni, L., Brugnera, A., Landi, M., Lo Sauro, C., ... Ricca, V. (2012). Generalized anxiety disorder : is there any specific symptom ?

- Comprehensive Psychiatry*, 53(8), 1056–1062. <https://doi.org/10.1016/j.comppsy.2012.04.002>
- Farrell, M. S., Werge, T., Sklar, P., Owen, M. J., Ophoff, R. A., O'Donovan, M. C., . . . Sullivan, P. F. (2015). Evaluating historical candidate genes for schizophrenia. *Molecular Psychiatry*, 20(5), 555–562. <https://doi.org/10.1038/mp.2015.16>
- Fawcett, J., Clark, D. C., Scheftner, W. A., & Gibbons, R. D. (1983). Assessing anhedonia in psychiatric patients. *Archives of General Psychiatry*, 40(1), 79–84.
- Fernando, A. B. P., & Robbins, T. W. (2011). Animal models of neuropsychiatric disorders. *Annual Review of Clinical Psychology*, 7, 39–61. <https://doi.org/10.1146/annurev-clinpsy-032210-104454>
- Fervaha, G., Foussias, G., Takeuchi, H., Agid, O., & Remington, G. (2016). Motivational deficits in major depressive disorder : Cross-sectional and longitudinal relationships with functional impairment and subjective well-being. *Comprehensive Psychiatry*, 66, 31–38. <https://doi.org/10.1016/j.comppsy.2015.12.004>
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614), 1898–1902. <https://doi.org/10.1126/science.1077349>
- Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, 79(6), 1243–1255. <https://doi.org/10.1016/j.neuron.2013.07.006>
- Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing : a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci*, 10(1), 48–58. <https://doi.org/10.1038/nrn2536>
- Forbes, N. F., Stewart, C. A., Matthews, K., & Reid, I. C. (1996). Chronic mild stress and sucrose consumption : validity as a model of depression. *Physiology and Behavior*, 60(6), 1481–1484.
- Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., & Fawcett, J. (2010). Antidepressant drug effects and depression severity : a patient-level meta-analysis. *JAMA*, 303(1), 47–53. <https://doi.org/10.1001/jama.2009.1943>
- Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your horses : impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, 318(5854), 1309–1312. <https://doi.org/10.1126/science.1146157>
- Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick : cognitive reinforcement learning in parkinsonism. *Science*, 306(5703), 1940–1943. <https://doi.org/10.1126/science.1102941>
- Freedman, R. (2005). The choice of antipsychotic drugs for schizophrenia. *The New England Journal of Medicine*, 353(12), 1286–1288. <https://doi.org/10.1056/>

NEJMe058200

Freedman, R., Lewis, D. A., Michels, R., Pine, D. S., Schultz, S. K., Tamminga, C. A., ... Yager, J. (2013). The initial field trials of DSM-5 : new blooms and old thorns. *The American Journal of Psychiatry*, 170(1), 1–5. <https://doi.org/10.1176/appi.ajp.2012.12091189>

Frith, C. D. (1992). *The Cognitive Neuropsychology of Schizophrenia (Essays in Cognitive Psychology)* (1st ed.). Hove, UK : Lawrence Erlbaum Associates.

Furey, M. L., & Drevets, W. C. (2006). Antidepressant efficacy of the antimuscarinic drug scopolamine : a randomized, placebo-controlled clinical trial. *Archives of General Psychiatry*, 63(10), 1121–1129. <https://doi.org/10.1001/archpsyc.63.10.1121>

Fusar-Poli, P., Bonoldi, I., Yung, A. R., Borgwardt, S., Kempton, M. J., Valmaggia, L., ... McGuire, P. (2012). Predicting psychosis : meta-analysis of transition outcomes in individuals at high clinical risk. *Archives of General Psychiatry*, 69(3), 220–229. <https://doi.org/10.1001/archgenpsychiatry.2011.1472>

Gard, D. E., Gard, M. G., Kring, A. M., & John, O. P. (2006). Anticipatory and consummatory components of the experience of pleasure : a scale development study. *Journal of Research in Personality*, 40(6), 1086–1102.

Gardner, E. L., & Ashby, C. R. (2000). Heterogeneity of the mesotelencephalic dopamine fibers : physiology and pharmacology. *Neuroscience and Biobehavioral Reviews*, 24(1), 115–118.

Garety, P. A., & Freeman, D. (1999). Cognitive approaches to delusions : a critical review of theories and evidence. *Br J Clin Psychol*, 38 ( Pt 2), 113–154.

Garety, P. A., & Freeman, D. (2013). The past and future of delusions research : from the inexplicable to the treatable. *The British Journal of Psychiatry*, 203(5), 327–333. <https://doi.org/10.1192/bjp.bp.113.126953>

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning : a meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 37(7), 1297–1310. <https://doi.org/10.1016/j.neubiorev.2013.03.023>

Gass, J. T., & Chandler, L. J. (2013). The Plasticity of Extinction : Contribution of the Prefrontal Cortex in Treating Addiction through Inhibitory Learning. *Frontiers in Psychiatry*, 4, 46. <https://doi.org/10.3389/fpsy.2013.00046>

Geoffroy, P. A., Amad, A., & Gaillard, R. (2014). Sciences et psychiatrie. *Annales Médico-Psychologiques*, 172(9), 752–755.

Gershon, E. S., Alliey-Rodriguez, N., & Liu, C. (2011). After GWAS : searching for genetic risk for schizophrenia and bipolar disorder. *The American Journal of Psychiatry*, 168(3), 253–256. <https://doi.org/10.1176/appi.ajp.2010.10091340>

Geyer, M. A., & Ellenbroek, B. (2003). Animal behavior models of the mechanisms

- underlying antipsychotic atypicality. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27(7), 1071–1079. <https://doi.org/10.1016/j.pnpbp.2003.09.003>
- Geyer, M. A., & Markou, A. (1995). Psychopharmacology : The Fourth Generation of Progress. In F. E. Bloom & D. J. Kupfer (Eds.) (4th ed.). Lippincott Williams & Wilkins.
- Giegling, I., Hosak, L., Mössner, R., Serretti, A., Bellivier, F., Claes, S., ... Rujescu, D. (2017). Genetics of schizophrenia : A consensus paper of the WFSBP Task Force on Genetics. *The World Journal of Biological Psychiatry*, 1–14. <https://doi.org/10.1080/15622975.2016.1268715>
- Gilbert, D. L., Christian, B. T., Gelfand, M. J., Shi, B., Mantil, J., & Sallee, F. R. (2006). Altered mesolimbocortical and thalamic dopamine in Tourette syndrome. *Neurology*, 67(9), 1695–1697. <https://doi.org/10.1212/01.wnl.0000242733.18534.2c>
- Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., & Wit, S. de. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *The American Journal of Psychiatry*, 168(7), 718–726. <https://doi.org/10.1176/appi.ajp.2011.10071062>
- Glahn, D. C., Knowles, E. E. M., McKay, D. R., Sprouten, E., Raventós, H., Blangero, J., ... Almasy, L. (2014). Arguments for the sake of endophenotypes : examining common misconceptions about the use of endophenotypes in psychiatric genetics. *American Journal of Medical Genetics*, 165B(2), 122–130. <https://doi.org/10.1002/ajmg.b.32221>
- Goldberg, D. (2011). The heterogeneity of “major depression”. *World Psychiatry*, 10(3), 226–228.
- Gooding, D. C., & Pflum, M. J. (2014). The assessment of interpersonal pleasure : introduction of the Anticipatory and Consummatory Interpersonal Pleasure Scale (ACIPS) and preliminary findings. *Psychiatry Research*, 215(1), 237–243. <https://doi.org/10.1016/j.psychres.2013.10.012>
- Gotlib, I. H., Hamilton, J. P., Cooney, R. E., Singh, M. K., Henry, M. L., & Joormann, J. (2010). Neural processing of reward and loss in girls at risk for major depression. *Arch Gen Psychiatry*, 67(4), 380–387. <https://doi.org/10.1001/archgenpsychiatry.2010.13>
- Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry : etymology and strategic intentions. *The American Journal of Psychiatry*, 160(4), 636–645. <https://doi.org/10.1176/appi.ajp.160.4.636>
- Gottesman, I. I., & Shields, J. (1973). Genetic theorizing and schizophrenia. *The British Journal of Psychiatry*, 122(566), 15–30.
- Gottesman, I. I., Laursen, T. M., Bertelsen, A., & Mortensen, P. B. (2010). Severe

- mental disorders in offspring with 2 psychiatrically ill parents. *Archives of General Psychiatry*, 67(3), 252–257. <https://doi.org/10.1001/archgenpsychiatry.2010.1>
- Gottfried, J. A., & Dolan, R. J. (2004). Human orbitofrontal cortex mediates extinction learning while accessing conditioned representations of value. *Nature Neuroscience*, 7(10), 1144–1152. <https://doi.org/10.1038/nn1314>
- Graaf-Peters, V. B. de, & Hadders-Algra, M. (2006). Ontogeny of the human central nervous system : what is happening when ? *Early Human Development*, 82(4), 257–266. <https://doi.org/10.1016/j.earlhumdev.2005.10.013>
- Graybiel, A. M., & Canales, J. J. (2001). The neurobiology of repetitive behaviors : clues to the neurobiology of Tourette syndrome. *Adv Neurol*, 85, 123–131.
- Greenwood, T. A., Braff, D. L., Light, G. A., Cadenhead, K. S., Calkins, M. E., Dobie, D. J., ... Schork, N. J. (2007). Initial heritability analyses of endophenotypic measures for schizophrenia : the consortium on the genetics of schizophrenia. *Archives of General Psychiatry*, 64(11), 1242–1250. <https://doi.org/10.1001/archpsyc.64.11.1242>
- Greenwood, T. A., Lazzeroni, L. C., Calkins, M. E., Freedman, R., Green, M. F., Gur, R. E., ... Braff, D. L. (2016a). Genetic assessment of additional endophenotypes from the Consortium on the Genetics of Schizophrenia Family Study. *Schizophrenia Research*, 170(1), 30–40. <https://doi.org/10.1016/j.schres.2015.11.008>
- Greenwood, T. A., Light, G. A., Swerdlow, N. R., Calkins, M. E., Green, M. F., Gur, R. E., ... Braff, D. L. (2016b). Gating Deficit Heritability and Correlation With Increased Clinical Severity in Schizophrenia Patients With Positive Family History. *The American Journal of Psychiatry*, 173(4), 385–391. <https://doi.org/10.1176/appi.ajp.2015.15050605>
- Griebel, G., & Beeské, S. (2012). Is there still a future for neurokinin 3 receptor antagonists as potential drugs for the treatment of psychiatric diseases ? *Pharmacology and Therapeutics*, 133(1), 116–123. <https://doi.org/10.1016/j.pharmthera.2011.09.007>
- Griebel, G., & Holsboer, F. (2012). Neuropeptide receptor ligands as drugs for psychiatric diseases : the end of the beginning ? *Nature Reviews. Drug Discovery*, 11(6), 462–478. <https://doi.org/10.1038/nrd3702>
- Group, B. D. W. (2001). Biomarkers and surrogate endpoints : preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics*, 69(3), 89–95. <https://doi.org/10.1067/mcp.2001.113989>
- Gu, X., Wu, X., Liu, Y., Cui, S., & Ma, Z. (2009). Tyrosine phosphorylation of the N-Methyl-D-Aspartate receptor 2B subunit in spinal cord contributes to remifentanyl-induced postoperative hyperalgesia : the preventive effect of ketamine. *Molecular Pain*, 5, 76. <https://doi.org/10.1186/1744-8069-5-76>
- Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus

- valence in decision making. *Trends in Cognitive Sciences*, 18(4), 194–202. <https://doi.org/10.1016/j.tics.2014.01.003>
- Hackam, D. G. (2007). Translating animal research into clinical benefit. *British Medical Journal*, 334(7586), 163–164. <https://doi.org/10.1136/bmj.39104.362951.80>
- Hackam, D. G., & Redelmeier, D. A. (2006). Translation of research evidence from animals to humans. *JAMA*, 296(14), 1731–1732. <https://doi.org/10.1001/jama.296.14.1731>
- Hallak, J. E. C., Paula, A. L. D. de, Chaves, C., Bressan, R. A., & Machado-de-Sousa, J. P. (2015). An Overview on the Search for Schizophrenia Biomarkers. *CNS and Neurological Disorders Drug Targets*, 14(8), 996–1000.
- Halligan, P. W., & David, A. S. (2001). Cognitive neuropsychiatry : towards a scientific psychopathology. *Nature Reviews. Neuroscience*, 2(3), 209–215. <https://doi.org/10.1038/35058586>
- Handley, R., Zelaya, F. O., Reinders, A. A. T. S., Marques, T. R., Mehta, M. A., O’Gorman, R., ... Dazzan, P. (2013). Acute effects of single-dose aripiprazole and haloperidol on resting cerebral blood flow (rCBF) in the human brain. *Hum Brain Mapp*, 34(2), 272–282. <https://doi.org/10.1002/hbm.21436>
- Harris, R. B., Zhou, J., Youngblood, B. D., Smagin, G. N., & Ryan, D. H. (1997). Failure to change exploration or saccharin preference in rats exposed to chronic mild stress. *Physiology and Behavior*, 63(1), 91–100.
- Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal of Neurophysiology*, 95(2), 948–959. <https://doi.org/10.1152/jn.00382.2005>
- Hasler, G. (2012). Can the neuroeconomics revolution revolutionize psychiatry? *Neuroscience and Biobehavioral Reviews*, 36(1), 64–78. <https://doi.org/10.1016/j.neubiorev.2011.04.011>
- Hay, M., Thomas, D. W., Craighead, J. L., Economides, C., & Rosenthal, J. (2014). Clinical development success rates for investigational drugs. *Nature Biotechnology*, 32(1), 40–51. <https://doi.org/10.1038/nbt.2786>
- Hehlmann, R., Hochhaus, A., Baccarani, M., & LeukemiaNet, E. (2007). Chronic myeloid leukaemia. *Lancet*, 370(9584), 342–350. [https://doi.org/10.1016/S0140-6736\(07\)61165-9](https://doi.org/10.1016/S0140-6736(07)61165-9)
- Hemsley, D. R. (1993). A simple (or simplistic?) cognitive model for schizophrenia. *Behaviour Research and Therapy*, 31(7), 633–645.
- Heresco-Levy, U., Gelfin, G., Bloch, B., Levin, R., Edelman, S., Javitt, D. C., & Kremer, I. (2013). A randomized add-on trial of high-dose D-cycloserine for

- treatment-resistant depression. *The International Journal of Neuropsychopharmacology*, 16(3), 501–506. <https://doi.org/10.1017/S1461145712000910>
- Hirst, W. D., Abrahamsen, B., Blaney, F. E., Calver, A. R., Aloj, L., Price, G. W., & Medhurst, A. D. (2003). Differences in the central nervous system distribution and pharmacology of the mouse 5-hydroxytryptamine-6 receptor compared with rat and human receptors investigated by radioligand binding, site-directed mutagenesis, and molecular modeling. *Molecular Pharmacology*, 64(6), 1295–1308. <https://doi.org/10.1124/mol.64.6.1295>
- Hoch, P. H., Catell, J. P., & Pennes, H. H. (1952). Effects of mescaline and lysergic acid (d-LSD-25). *The American Journal of Psychiatry*, 108(8), 579–584. <https://doi.org/10.1176/ajp.108.8.579>
- Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, 1(4), 304–309. <https://doi.org/10.1038/1124>
- Holtzheimer, P. E., & Mayberg, H. S. (2011). Deep brain stimulation for psychiatric disorders. *Annual Review of Neuroscience*, 34, 289–307. <https://doi.org/10.1146/annurev-neuro-061010-113638>
- Horner, A. E., Heath, C. J., Hvoslef-Eide, M., Kent, B. A., Kim, C. H., Nilsson, S. R. O., ... Bussey, T. J. (2013). The touchscreen operant platform for testing learning and memory in rats and mice. *Nature Protocols*, 8(10), 1961–1984. <https://doi.org/10.1038/nprot.2013.122>
- Hou, C.-L., Xiang, Y.-T., Wang, Z.-L., Everall, I., Tang, Y., Yang, C., ... Jia, F.-J. (2016). Cognitive functioning in individuals at ultra-high risk for psychosis, first-degree relatives of patients with psychosis and patients with first-episode schizophrenia. *Schizophrenia Research*, 174(1-3), 71–76. <https://doi.org/10.1016/j.schres.2016.04.034>
- Howard, R. A. (1960). *Dynamic programming and markov processes*. The M.I.T. Press.
- Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia : version III—the final common pathway. *Schizophr Bull*, 35(3), 549–562. <https://doi.org/10.1093/schbul/sbp006>
- Howes, O., McCutcheon, R., & Stone, J. (2015). Glutamate and dopamine in schizophrenia : an update for the 21st century. *Journal of Psychopharmacology*, 29(2), 97–115. <https://doi.org/10.1177/0269881114563634>
- Huettel, S. A. (2012). Event-related fMRI in cognition. *NeuroImage*, 62(2), 1152–1156. <https://doi.org/10.1016/j.neuroimage.2011.08.113>
- Huys, Q. J. M., Moutoussis, M., & Williams, J. (2011). Are computational models of any use to psychiatry? *Neural Netw*, 24(6), 544–551. <https://doi.org/10.1016/j.j>

neunet.2011.03.001

Insel, T. R., & Landis, S. C. (2013). Twenty-five years of progress : the view from NIMH and NINDS. *Neuron*, 80(3), 561–567. <https://doi.org/10.1016/j.neuron.2013.09.041>

Insel, T. R., & Scolnick, E. M. (2006). Cure therapeutics and strategic prevention : raising the bar for mental health research. *Molecular Psychiatry*, 11(1), 11–17. <https://doi.org/10.1038/sj.mp.4001777>

Insel, T. R., Voon, V., Nye, J. S., Brown, V. J., Altevogt, B. M., Bullmore, E. T., ... Sahakian, B. J. (2013). Innovative solutions to novel drug development in mental health. *Neuroscience and Biobehavioral Reviews*, 37(10 Pt 1), 2438–2444. <https://doi.org/10.1016/j.neubiorev.2013.03.022>

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria (RDoC) : toward a new classification framework for research on mental disorders. *Am J Psychiatry*, 167(7), 748–751. <https://doi.org/10.1176/appi.ajp.2010.09091379>

Ioannidis, J. P. A. (2008). Effectiveness of antidepressants : an evidence myth constructed from a thousand randomized trials ? *Philosophy, Ethics, and Humanities in Medicine*, 3, 14. <https://doi.org/10.1186/1747-5341-3-14>

Javitt, D. C., Zukin, S. R., Heresco-Levy, U., & Umbricht, D. (2012). Has an angel shown the way ? Etiological and therapeutic implications of the PCP/NMDA model of schizophrenia. *Schizophr Bull*, 38(5), 958–966. <https://doi.org/10.1093/schbul/sbs069>

Johnson, B. D. (1971). Psychosis and ketamine. *British Medical Journal*, 4(5784), 428–429.

Jones, D. N., Gartlon, J. E., Minassian, A., Perry, W., & Geyer, M. A. (2008). Animal and Translational Models for CNS Drug Discovery, Vol. 1 : Psychiatric Disorders. In R. A. McArthur & F. Borsini (Eds.) (1st ed., pp. 199–262). Academic Press.

Jutkiewicz, E. M., Wood, S. K., Houshyar, H., Hsin, L.-W., Rice, K. C., & Woods, J. H. (2005). The effects of CRF antagonists, antalarmin, CP154,526, LWH234, and R121919, in the forced swim test and on swim-induced increases in adrenocorticotropin in rats. *Psychopharmacology*, 180(2), 215–223. <https://doi.org/10.1007/s00213-005-2164-z>

Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2010). The neural code of reward anticipation in human orbitofrontal cortex. *Proc Natl Acad Sci U S A*, 107(13), 6010–6015. <https://doi.org/10.1073/pnas.0912838107>

Kahnt, T., Park, S. Q., Cohen, M. X., Beck, A., Heinz, A., & Wrase, J. (2009). Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are

- used to guide decisions. *J Cogn Neurosci*, 21(7), 1332–1345. <https://doi.org/10.1162/jocn.2009.21092>
- Kalant, O. J. (1971). Moreau, hashish, and hallucinations. *International Journal of the Addictions*, 6(3), 553–560.
- Kamin, L. (1969). Punishment and aversive behavior. In B. A. Campbell & R. M. Church (Eds.), *New York : Appleton-Century-Crofts* (pp. 279–296). New York : Appleton-Century-Crofts.
- Kantrowitz, J. T., & Javitt, D. C. (2010). N-methyl-d-aspartate (NMDA) receptor dysfunction or dysregulation : the final common pathway on the road to schizophrenia? *Brain Research Bulletin*, 83(3-4), 108–121. <https://doi.org/10.1016/j.brainresbull.2010.04.006>
- Kapur, S. (2003). Psychosis as a state of aberrant salience : a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry*, 160(1), 13–23.
- Kapur, S., Phillips, A. G., & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 17(12), 1174–1179. <https://doi.org/10.1038/mp.2012.105>
- Kendler, K. S. (2015). A joint history of the nature of genetic variation and the nature of schizophrenia. *Molecular Psychiatry*, 20(1), 77–83. <https://doi.org/10.1038/mp.2014.94>
- Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology*, 7(5), e1002055. <https://doi.org/10.1371/journal.pcbi.1002055>
- Keramati, M., Smittenaar, P., Dolan, R. J., & Dayan, P. (2016). Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.1609094113>
- Keshavan, M. S., Kulkarni, S., Bhojraj, T., Francis, A., Diwadkar, V., Montrose, D. M., ... Sweeney, J. (2010). Premorbid cognitive deficits in young relatives of schizophrenia patients. *Frontiers in Human Neuroscience*, 3, 62. <https://doi.org/10.3389/neuro.09.062.2009>
- Kesselheim, A. S., Hwang, T. J., & Franklin, J. M. (2015). Two decades of new drug development for central nervous system disorders. *Nature Reviews. Drug Discovery*, 14(12), 815–816. <https://doi.org/10.1038/nrd4793>
- Kim, E., Howes, O. D., Turkheimer, F. E., Kim, B.-H., Jeong, J. M., Kim, J. W., ... Kwon, J. S. (2013). The relationship between antipsychotic D2 occupancy and change in frontal metabolism and working memory : A dual [(11)C]raclopride and [(18)F]FDG imaging study with aripiprazole. *Psychopharmacology*, 227(2), 221–229.

<https://doi.org/10.1007/s00213-012-2953-0>

Kim, H., Shimojo, S., & O'Doherty, J. P. (2011). Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cerebral Cortex*, 21(4), 769–776. <https://doi.org/10.1093/cercor/bhq145>

Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., ... Vaitl, D. (2003). Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system : an event-related fMRI study. *NeuroImage*, 20(2), 1086–1095. [https://doi.org/10.1016/S1053-8119\(03\)00381-1](https://doi.org/10.1016/S1053-8119(03)00381-1)

Kirson, N. Y., Weiden, P. J., Yermakov, S., Huang, W., Samuelson, T., Offord, S. J., ... Wong, B. J. O. (2013). Efficacy and effectiveness of depot versus oral antipsychotics in schizophrenia : synthesizing results across different research designs. *The Journal of Clinical Psychiatry*, 74(6), 568–575. <https://doi.org/10.4088/JCP.12r08167>

Kishida, K. T., Saez, I., Lohrenz, T., Witcher, M. R., Laxton, A. W., Tatter, S. B., ... Montague, P. R. (2016). Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proceedings of the National Academy of Sciences of the United States of America*, 113(1), 200–205. <https://doi.org/10.1073/pnas.1513619112>

Kishimoto, T., Nitta, M., Borenstein, M., Kane, J. M., & Correll, C. U. (2013). Long-acting injectable versus oral antipsychotics in schizophrenia : a systematic review and meta-analysis of mirror-image studies. *The Journal of Clinical Psychiatry*, 74(10), 957–965. <https://doi.org/10.4088/JCP.13r08440>

Kishimoto, T., Robenzadeh, A., Leucht, C., Leucht, S., Watanabe, K., Mimura, M., ... Correll, C. U. (2014). Long-acting injectable vs oral antipsychotics for relapse prevention in schizophrenia : a meta-analysis of randomized trials. *Schizophrenia Bulletin*, 40(1), 192–213. <https://doi.org/10.1093/schbul/sbs150>

Klei, L., Sanders, S. J., Murtha, M. T., Hus, V., Lowe, J. K., Willsey, A. J., ... Devlin, B. (2012). Common genetic variants, acting additively, are a major source of risk for autism. *Molecular Autism*, 3(1), 9. <https://doi.org/10.1186/2040-2392-3-9>

Knight, A. (2008). Systematic reviews of animal experiments demonstrate poor contributions toward human healthcare. *Reviews on Recent Clinical Trials*, 3(2), 89–96.

Knutson, B., Burgdorf, J., & Panksepp, J. (2002). Ultrasonic vocalizations as indices of affective states in rats. *Psychological Bulletin*, 128(6), 961–977.

Kobayashi, S., Lauwereyns, J., Koizumi, M., Sakagami, M., & Hikosaka, O. (2002). Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. *Journal of Neurophysiology*, 87(3), 1488–1498.

Koechlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical Transactions of the Royal Society of*

*London. Series B, Biological Sciences*, 369(1655). <https://doi.org/10.1098/rstb.2013.0474>

Kola, I., & Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nature Reviews. Drug Discovery*, 3(8), 711–715. <https://doi.org/10.1038/nrd1470>

Koutsouleris, N., Borgwardt, S., Meisenzahl, E. M., Bottlender, R., Möller, H.-J., & Riecher-Rössler, A. (2012a). Disease prediction in the at-risk mental state for psychosis using neuroanatomical biomarkers : results from the FePsy study. *Schizophrenia Bulletin*, 38(6), 1234–1246. <https://doi.org/10.1093/schbul/sbr145>

Koutsouleris, N., Davatzikos, C., Bottlender, R., Patschurek-Kliche, K., Scheuerecker, J., Decker, P., ... Meisenzahl, E. M. (2012b). Early recognition and disease prediction in the at-risk mental states for psychosis using neurocognitive pattern classification. *Schizophrenia Bulletin*, 38(6), 1200–1215. <https://doi.org/10.1093/schbul/sbr037>

Koutsouleris, N., Meisenzahl, E. M., Davatzikos, C., Bottlender, R., Frodl, T., Scheuerecker, J., ... Gaser, C. (2009). Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Archives of General Psychiatry*, 66(7), 700–712. <https://doi.org/10.1001/archgenpsychiatry.2009.62>

Koychev, I., McMullen, K., Lees, J., Dadhiwala, R., Grayson, L., Perry, C., ... Barkus, E. (2012). A validation of cognitive biomarkers for the early identification of cognitive enhancing agents in schizotypy : a three-center double-blind placebo-controlled study. *European Neuropsychopharmacology*, 22(7), 469–481. <https://doi.org/10.1016/j.euroneuro.2011.10.005>

Kreiman, G. (2007). Single unit approaches to human vision and memory. *Current Opinion in Neurobiology*, 17(4), 471–475. <https://doi.org/10.1016/j.conb.2007.07.005>

Kreiman, G., Fried, I., & Koch, C. (2002). Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), 8378–8383. <https://doi.org/10.1073/pnas.072194099>

Kriekhaus, E. E., & Wolf, G. (1968). Acquisition of sodium by rats : interaction of innate mechanisms and latent learning. *Journal of Comparative and Physiological Psychology*, 65(2), 197–201.

Kristian Hill, S., Buchholz, A., Amsbaugh, H., Reilly, J. L., Rubin, L. H., Gold, J. M., ... Sweeney, J. A. (2015). Working memory impairment in probands with schizoaffective disorder and first degree relatives of schizophrenia probands extend beyond deficits predicted by generalized neuropsychological impairment. *Schizo-*

- phrenia Research*, 166(1-3), 310–315. <https://doi.org/10.1016/j.schres.2015.05.018>
- Kronmal, R. A., Cain, K. C., Ye, Z., & Omenn, G. S. (1993). Total serum cholesterol levels and mortality risk as a function of age. A report based on the Framingham data. *Archives of Internal Medicine*, 153(9), 1065–1073.
- Krystal, J. H., Anand, A., & Moghaddam, B. (2002). Effects of NMDA receptor antagonists : implications for the pathophysiology of schizophrenia. *Archives of General Psychiatry*, 59(7), 663–664.
- Krystal, J. H., Karper, L. P., Seibyl, J. P., Freeman, G. K., Delaney, R., Bremner, J. D., ... Charney, D. S. (1994). Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Archives of General Psychiatry*, 51(3), 199–214.
- Krystal, J. H., Perry, E. B., Gueorguieva, R., Belger, A., Madonick, S. H., Abi-Dargham, A., ... D'Souza, D. C. (2005). Comparative and interactive human psychopharmacologic effects of ketamine and amphetamine : implications for glutamatergic and dopaminergic model psychoses and cognitive function. *Arch Gen Psychiatry*, 62(9), 985–994. <https://doi.org/10.1001/archpsyc.62.9.985>
- Kyzas, P. A., Denaxa-Kyza, D., & Ioannidis, J. P. A. (2007). Almost all articles on cancer prognostic markers report statistically significant results. *European Journal of Cancer*, 43(17), 2559–2579. <https://doi.org/10.1016/j.ejca.2007.08.030>
- Laborit, H., Huguenard, P., & Alluaume, R. (1952). Un nouveau stabilisateur végétatif (le 4560 RP). *La Presse Medicale*, 60(10), 206–208.
- Lafeuille, M.-H., Dean, J., Carter, V., Duh, M. S., Fastenau, J., Dirani, R., & Lefebvre, P. (2014). Systematic review of long-acting injectables versus oral atypical antipsychotics on hospitalization in schizophrenia. *Current Medical Research and Opinion*, 30(8), 1643–1655. <https://doi.org/10.1185/03007995.2014.915211>
- Lahti, A. C., Weiler, M. A., Michaelidis, B. A. T., Parwani, A., & Tamminga, C. A. (2001). Effects of ketamine in normal and schizophrenic volunteers. *Neuropsychopharmacology*, 25(4), 455–467. [https://doi.org/10.1016/S0893-133X\(01\)00243-3](https://doi.org/10.1016/S0893-133X(01)00243-3)
- Langreth, & Waldholz. (1999). New era of personalized medicine : targeting drugs for each unique genetic profile. *The Oncologist*, 4(5), 426–427.
- Large, C. H. (2007). Do NMDA receptor antagonist models of schizophrenia predict the clinical efficacy of antipsychotic drugs? *Journal of Psychopharmacology*, 21(3), 283–301. <https://doi.org/10.1177/0269881107077712>
- Larsen, T., & O'Doherty, J. P. (2014). Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain : a combined fMRI-EEG study. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 369(1655). <https://doi.org/10.1098/rstb.2013.0473>
- Lau, B., & Glimcher, P. W. (2007). Action and outcome encoding in the primate

- caudate nucleus. *The Journal of Neuroscience*, 27(52), 14502–14514. <https://doi.org/10.1523/JNEUROSCI.3060-07.2007>
- Leckman, J. F. (2002). Tourette's syndrome. *Lancet*, 360(9345), 1577–1586. [https://doi.org/10.1016/S0140-6736\(02\)11526-1](https://doi.org/10.1016/S0140-6736(02)11526-1)
- Lee, D. (2013). Decision making : from neuroscience to psychiatry. *Neuron*, 78(2), 233–248. <https://doi.org/10.1016/j.neuron.2013.04.008>
- Lee, D., Seo, H., & Jung, M. W. (2012). Neural basis of reinforcement learning and decision making. *Annual Review of Neuroscience*, 35, 287–308. <https://doi.org/10.1146/annurev-neuro-062111-150512>
- Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron*, 81(3), 687–699. <https://doi.org/10.1016/j.neuron.2013.11.028>
- Lencz, T., & Malhotra, A. K. (2015). Targeting the schizophrenia genome : a fast track strategy from GWAS to clinic. *Molecular Psychiatry*, 20(7), 820–826. <https://doi.org/10.1038/mp.2015.28>
- Lenzenweger, M. F. (2013). Thinking clearly about the endophenotype-intermediate phenotype-biomarker distinctions in developmental psychopathology research. *Development and Psychopathology*, 25(4 Pt 2), 1347–1357. <https://doi.org/10.1017/S0954579413000655>
- Lewis, D. A., & Levitt, P. (2002). Schizophrenia as a disorder of neurodevelopment. *Annual Review of Neuroscience*, 25, 409–432. <https://doi.org/10.1146/annurev.neuro.25.112701.142754>
- Li, J., & Daw, N. D. (2011). Signals in human striatum are appropriate for policy update rather than value prediction. *The Journal of Neuroscience*, 31(14), 5504–5511. <https://doi.org/10.1523/JNEUROSCI.6316-10.2011>
- Lichtenstein, P., Yip, B. H., Björk, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F., & Hultman, C. M. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families : a population-based study. *Lancet*, 373(9659), 234–239. [https://doi.org/10.1016/S0140-6736\(09\)60072-6](https://doi.org/10.1016/S0140-6736(09)60072-6)
- Lieberman, J. A., Stroup, T. S., McEvoy, J. P., Swartz, M. S., Rosenheck, R. A., Perkins, D. O., ... Intervention Effectiveness (CATIE) Investigators, C. A. T. of. (2005). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *The New England Journal of Medicine*, 353(12), 1209–1223. <https://doi.org/10.1056/NEJMoa051688>
- Lipska, B. K. (2004). Using animal models to test a neurodevelopmental hypothesis of schizophrenia. *Journal of Psychiatry and Neuroscience*, 29(4), 282–286.
- Lipska, B. K., & Weinberger, D. R. (2000). To model a psychiatric disorder in animals : schizophrenia as a reality test. *Neuropsychopharmacology*, 23(3), 223–239.

[https://doi.org/10.1016/S0893-133X\(00\)00137-8](https://doi.org/10.1016/S0893-133X(00)00137-8)

Lipska, B. K., Jaskiw, G. E., & Weinberger, D. R. (1993). Postpubertal emergence of hyperresponsiveness to stress and to amphetamine after neonatal excitotoxic hippocampal damage : a potential animal model of schizophrenia. *Neuropsychopharmacology*, 9(1), 67–75. <https://doi.org/10.1038/npp.1993.44>

Lipska, B. K., Swerdlow, N. R., Geyer, M. A., Jaskiw, G. E., Braff, D. L., & Weinberger, D. R. (1995). Neonatal excitotoxic hippocampal damage in rats causes post-pubertal changes in prepulse inhibition of startle and its disruption by apomorphine. *Psychopharmacology*, 122(1), 35–43.

Liu, W.-H., Roiser, J. P., Wang, L.-Z., Zhu, Y.-H., Huang, J., Neumann, D. L., ... Chan, R. C. K. (2016). Anhedonia is associated with blunted reward sensitivity in first-degree relatives of patients with major depression. *Journal of Affective Disorders*, 190, 640–648. <https://doi.org/10.1016/j.jad.2015.10.050>

Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages : a meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 35(5), 1219–1236. <https://doi.org/10.1016/j.neubiorev.2010.12.012>

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences of the United States of America*, 104(22), 9493–9498. <https://doi.org/10.1073/pnas.0608842104>

Loomes, G., & Sugden, R. (1982). Regret theory : An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368), 805–824.

Louis, C., Cohen, C., Depoortère, R., & Griebel, G. (2006). Antidepressant-like effects of the corticotropin-releasing factor 1 receptor antagonist, SSR125543, and the vasopressin 1b receptor antagonist, SSR149415, in a DRL-72 s schedule in the rat. *Neuropsychopharmacology*, 31(10), 2180–2187. <https://doi.org/10.1038/sj.npp.1301036>

Luby, E. D., Cohen, B. D., Rosenbaum, G., Gottlieb, J. S., & Keller, R. (1959). Study of a new schizophrenomimetic drug ; sernyl. *A.M.A. Archives of Neurology and Psychiatry*, 81(3), 363–369.

Ma, Y. (2015). Neuropsychological mechanism underlying antidepressant effect : a systematic meta-analysis. *Molecular Psychiatry*, 20(3), 311–319. <https://doi.org/10.1038/mp.2014.24>

Macleod, M. R., Michie, S., Roberts, I., Dirnagl, U., Chalmers, I., Ioannidis, J. P. A., ... Glasziou, P. (2014). Biomedical research : increasing value, reducing waste. *Lancet*, 383(9912), 101–104. [https://doi.org/10.1016/S0140-6736\(13\)62329-6](https://doi.org/10.1016/S0140-6736(13)62329-6)

Macoveanu, J., Fisher, P. M., Haahr, M. E., Frokjaer, V. G., Knudsen, G. M., & Siebner, H. R. (2014). Effects of selective serotonin reuptake inhibition on

- neural activity related to risky decisions and monetary rewards in healthy males. *NeuroImage*, 99, 434–442. <https://doi.org/10.1016/j.neuroimage.2014.05.040>
- Macoveanu, J., Knorr, U., Skimminge, A., Søndergaard, M. G., Jørgensen, A., Fauerholdt-Jepsen, M., . . . Kessing, L. V. (2014). Altered reward processing in the orbitofrontal cortex and hippocampus in healthy first-degree relatives of patients with depression. *Psychological Medicine*, 44(6), 1183–1195. <https://doi.org/10.1017/S0033291713001815>
- Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, 14(2), 154–162. <https://doi.org/10.1038/nn.2723>
- Malison, R. T., McDougle, C. J., van Dyck, C. H., Scahill, L., Baldwin, R. M., Seibyl, J. P., . . . Innis, R. B. (1995). [123I]beta-CIT SPECT imaging of striatal dopamine transporter binding in Tourette’s disorder. *Am J Psychiatry*, 152(9), 1359–1361. <https://doi.org/10.1176/ajp.152.9.1359>
- Matsumoto, K., Suzuki, W., & Tanaka, K. (2003). Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science*, 301(5630), 229–232. <https://doi.org/10.1126/science.1084204>
- Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, 10(5), 647–656. <https://doi.org/10.1038/nn1890>
- McCabe, C., Mishor, Z., Cowen, P. J., & Harmer, C. J. (2010). Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. *Biol Psychiatry*, 67(5), 439–445. <https://doi.org/10.1016/j.biopsych.2009.11.001>
- McCarthy, S. E., Gillis, J., Kramer, M., Lihm, J., Yoon, S., Berstein, Y., . . . Corvin, A. (2014). De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Molecular Psychiatry*, 19(6), 652–658. <https://doi.org/10.1038/mp.2014.29>
- McEvoy, J. P., & Freudenreich, O. (2008). Animal and Translational Models for CNS Drug Discovery, Vol. 1 : Psychiatric Disorders. In R. A. McArthur & F. Borsini (Eds.) (1st ed., pp. 75–95). Academic Press.
- McHugh, P. R. (2005). Striving for coherence : psychiatry’s efforts over classification. *JAMA*, 293(20), 2526–2528. <https://doi.org/10.1001/jama.293.20.2526>
- McKinney, W. T., & Bunney, W. E. (1969). Animal model of depression. I. Review of evidence : implications for research. *Archives of General Psychiatry*, 21(2), 240–248.
- Meyer-Lindenberg, A. (2010). From maps to mechanisms through neuroimaging of

- schizophrenia. *Nature*, 468(7321), 194–202. <https://doi.org/10.1038/nature09569>
- Michael Gazzaniga, G. M., Richard Ivry. (2014). *Cognitive Neuroscience. The Biology of the Mind* (4th ed.). W. W. Norton.
- Miczek, K. A. (2008). Animal and Translational Models for CNS Drug Discovery, Vol. 1 : Psychiatric Disorders. In R. A. McArthur & F. Borsini (Eds.) (1st ed., pp. 97–116). Academic Press.
- Millan, M. J. (2008). Animal and Translational Models for CNS Drug Discovery, Vol. 1 : Psychiatric Disorders. In R. A. McArthur & F. Borsini (Eds.) (1st ed., pp. 1–58). Academic Press.
- Miller, G. (2010). Psychiatry. Beyond DSM : seeking a brain-based classification of mental illness. *Science*, 327(5972), 1437. <https://doi.org/10.1126/science.327.5972.1437>
- Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, 66(8), 811–822. <https://doi.org/10.1001/archgenpsychiatry.2009.91>
- Minzer, K., Lee, O., Hong, J. J., & Singer, H. S. (2004). Increased prefrontal D2 protein in Tourette syndrome : a postmortem analysis of frontal cortex and striatum. *J Neurol Sci*, 219(1-2), 55–61. <https://doi.org/10.1016/j.jns.2003.12.006>
- Miskowiak, K. W., Kjørstad, H. L., Meluken, I., Petersen, J. Z., Maciel, B. R., Köhler, C. A., ... Carvalho, A. F. (2017). The search for neuroimaging and cognitive endophenotypes : A critical systematic review of studies involving unaffected first-degree relatives of individuals with bipolar disorder. *Neuroscience and Biobehavioral Reviews*, 73, 1–22. <https://doi.org/10.1016/j.neubiorev.2016.12.011>
- Mistlberger, R. E. (1994). Circadian food-anticipatory activity : formal models and physiological mechanisms. *Neuroscience and Biobehavioral Reviews*, 18(2), 171–195.
- Moghaddam, B., & Krystal, J. H. (2012). Capturing the angel in “angel dust” : twenty years of translational neuroscience studies of NMDA receptor antagonists in animals and humans. *Schizophr Bull*, 38(5), 942–949. <https://doi.org/10.1093/schbul/sbs075>
- Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends Cogn Sci*, 16(1), 72–80. <https://doi.org/10.1016/j.tics.2011.11.018>
- Moore, J. W., Cambridge, V. C., Morgan, H., Giorlando, F., Adapa, R., & Fletcher, P. C. (2013). Time, action and psychosis : Using subjective time to investigate the effects of ketamine on sense of agency. *Neuropsychologia*, 51(2), 377–384.

<https://doi.org/10.1016/j.neuropsychologia.2012.07.005>

Morgan, C. J. A., Curran, H. V., & Drugs, I. S. C. on. (2012). Ketamine use : a review. *Addiction*, 107(1), 27–38.

Morgan, H. L., Turner, D. C., Corlett, P. R., Absalom, A. R., Adapa, R., Arana, F. S., ... Fletcher, P. C. (2011). Exploring the impact of ketamine on the experience of illusory body ownership. *Biol Psychiatry*, 69(1), 35–41. <https://doi.org/10.1016/j.biopsych.2010.07.032>

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature Neuroscience*, 9(8), 1057–1063. <https://doi.org/10.1038/nn1743>

Moy, S. S., & Nadler, J. J. (2008). Advances in behavioral genetics : mouse models of autism. *Molecular Psychiatry*, 13(1), 4–26. <https://doi.org/10.1038/sj.mp.4002082>

Muenzinger, K. F. (1938). Vicarious trial and error at a point of choice : I. A general survey of its relation to learning efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology*, 53(1), 75–86.

Murray, R. M., Paparelli, A., Morrison, P. D., Marconi, A., & Di Forti, M. (2013). What can we learn about schizophrenia from studying the human model, drug-induced psychosis? *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, 162B(7), 661–670. <https://doi.org/10.1002/ajmg.b.32177>

Naccache, L., El Karoui, I., Salti, M., Chammat, M., Maillet, M., & Allali, S. (2015). [How our subjective coherence is built ? The model of cognitive dissonance]. *Bulletin de L'Academie Nationale de Medecine*, 199(2-3), 253–259.

Nagai, T., Kitahara, Y., Ibi, D., Nabeshima, T., Sawa, A., & Yamada, K. (2011). Effects of antipsychotics on the behavioral deficits in human dominant-negative DISC1 transgenic mice with neonatal polyI :C treatment. *Behavioural Brain Research*, 225(1), 305–310. <https://doi.org/10.1016/j.bbr.2011.07.049>

NAMHC, N. A. M. H. C. W. (2010). *From discovery to cure : accelerating the development of new and personalized interventions for mental illness. Report of the National Advisory Mental Health Council ?s Workgroup.*

National Collaborating Centre for Mental Health (UK). (2010). Depression : The Treatment and Management of Depression in Adults (Updated Edition). British Psychological Society.

Nestler, E. J., & Hyman, S. E. (2010). Animal models of neuropsychiatric disorders. *Nature Neuroscience*, 13(10), 1161–1169. <https://doi.org/10.1038/nn.2647>

Nielsen, D. M., Carey, G. J., & Gold, L. H. (2004). Antidepressant-like activity of corticotropin-releasing factor type-1 receptor antagonists in mice. *European Journal of Pharmacology*, 499(1-2), 135–146. <https://doi.org/10.1016/j.ejphar.2004.07.091>

Nitschke, J. B., Dixon, G. E., Sarinopoulos, I., Short, S. J., Cohen, J. D., Smith,

- E. E., ... Davidson, R. J. (2006). Altering expectancy dampens neural response to aversive taste in primary taste cortex. *Nat Neurosci*, 9(3), 435–442. <https://doi.org/10.1038/nn1645>
- Niv, Y., & Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences*, 12(7), 265–272. <https://doi.org/10.1016/j.tics.2008.03.006>
- Nutt, D. J., & Need, A. C. (2014). Where now for schizophrenia research? *European Neuropsychopharmacology*, 24(8), 1181–1187. <https://doi.org/10.1016/j.euroneuro.2014.05.012>
- Olds, J., & Milner, P. (1954). 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.
- Overstreet, D. H., & Griebel, G. (2004). Antidepressant-like effects of CRF1 receptor antagonist SSR125543 in an animal model of depression. *European Journal of Pharmacology*, 497(1), 49–53. <https://doi.org/10.1016/j.ejphar.2004.06.035>
- Owen, M. J., O'Donovan, M. C., Thapar, A., & Craddock, N. (2011). Neurodevelopmental hypothesis of schizophrenia. *The British Journal of Psychiatry*, 198(3), 173–175. <https://doi.org/10.1192/bjp.bp.110.084384>
- O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, 38(2), 329–337.
- O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, 33(5), 815–826.
- O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, 1104, 35–53. <https://doi.org/10.1196/annals.1390.022>
- O'Doherty, J. P., Lee, S. W., & McNamee, D. (2015). The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences*, 1, 94–100.
- Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441(7090), 223–226. <https://doi.org/10.1038/nature04676>
- Palminteri, S., Clair, A.-H., Mallet, L., & Pessiglione, M. (2012a). Similar improvement of reward and punishment learning by serotonin reuptake inhibitors in obsessive-compulsive disorder. *Biol Psychiatry*, 72(3), 244–250. <https://doi.org/10.1016/j.biopsych.2011.12.028>
- Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., ... Pessiglione, M. (2012b). Critical roles for anterior insula and dorsal striatum in

- punishment-based avoidance learning. *Neuron*, 76(5), 998–1009. <https://doi.org/10.1016/j.neuron.2012.10.017>
- Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A., & Pessiglione, M. (2009). Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proc Natl Acad Sci U S A*, 106(45), 19179–19184. <https://doi.org/10.1073/pnas.0904035106>
- Pardridge, W. M. (2001). Crossing the blood-brain barrier : are we getting it right ? *Drug Discovery Today*, 6(1), 1–2.
- Pearl, J. (1984). *Heuristics : intelligent search strategies for computer problem solving*. Addison-Wesley Pub. Co., Inc., Reading, MA.
- Perel, P., Roberts, I., Sena, E., Wheble, P., Briscoe, C., Sandercock, P., ... Khan, K. S. (2007). Comparison of treatment effects between animal experiments and clinical trials : systematic review. *British Medical Journal*, 334(7586), 197. <https://doi.org/10.1136/bmj.39048.407928.BE>
- Pergadia, M. L., Der-Avakian, A., D'Souza, M. S., Madden, P. A. F., Heath, A. C., Shiffman, S., ... Pizzagalli, D. A. (2014). Association between nicotine withdrawal and reward responsiveness in humans and rats. *JAMA Psychiatry*, 71(11), 1238–1245. <https://doi.org/10.1001/jamapsychiatry.2014.1016>
- Perry, E. B., Cramer, J. A., Cho, H.-S., Petrakis, I. L., Karper, L. P., Genovese, A., ... Group, Y. K. S. (2007). Psychiatric safety of ketamine in psychopharmacology research. *Psychopharmacology*, 192(2), 253–260.
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, 59(4), 561–567. <https://doi.org/10.1016/j.neuron.2008.07.005>
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & D., F. C. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106), 1042–1045. <https://doi.org/10.1038/nature05051>
- Petronis, A., Gottesman, I. I., Crow, T. J., DeLisi, L. E., Klar, A. J., Macciardi, F., ... Sutherland, G. R. (2000). Psychiatric epigenetics : a new focus for the new century. *Molecular Psychiatry*, 5(4), 342–346.
- Phillips, A. G., & Barr, A. M. (1997). Effects of chronic mild stress on motivation for sucrose : mixed messages. *Psychopharmacology*, 134(4), 361–6 ; discussion 371–7.
- Pickard, B. S. (2015). Schizophrenia biomarkers : translating the descriptive into the diagnostic. *Journal of Psychopharmacology*, 29(2), 138–143. <https://doi.org/10.1177/0269881114566631>
- Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype : a signal-detection approach. *Biological*

- Psychiatry*, 57(4), 319–327. <https://doi.org/10.1016/j.biopsych.2004.11.026>
- Pletnikov, M. V., Ayhan, Y., Nikolskaia, O., Xu, Y., Ovanesov, M. V., Huang, H., ... Ross, C. A. (2008). Inducible expression of mutant human DISC1 in mice is associated with brain and behavioral abnormalities reminiscent of schizophrenia. *Molecular Psychiatry*, 13(2), 173–86, 115. <https://doi.org/10.1038/sj.mp.4002079>
- Poels, E. M. P., Kegeles, L. S., Kantrowitz, J. T., Slifstein, M., Javitt, D. C., Lieberman, J. A., ... Girgis, R. R. (2014). Imaging glutamate in schizophrenia : review of findings and implications for drug discovery. *Molecular Psychiatry*, 19(1), 20–29. <https://doi.org/10.1038/mp.2013.136>
- Pomarol-Clotet, E., Honey, G. D., Murray, G. K., Corlett, P. R., Absalom, A. R., Lee, M., ... Fletcher, P. C. (2006). Psychological effects of ketamine in healthy volunteers. Phenomenological study. *The British Journal of Psychiatry*, 189, 173–179. <https://doi.org/10.1192/bjp.bp.105.015263>
- Porsolt, R. D., Anton, G., Blavet, N., & Jalfre, M. (1978). Behavioural despair in rats : a new model sensitive to antidepressant treatments. *European Journal of Pharmacology*, 47(4), 379–391.
- Preskorn, S. H. (2010). CNS drug development. Part I : The early period of CNS drugs. *Journal of Psychiatric Practice*, 16(5), 334–339. <https://doi.org/10.1097/01.pra.0000388628.44405.c0>
- Price, J., Cole, V., & Goodwin, G. M. (2009). Emotional side-effects of selective serotonin reuptake inhibitors : qualitative study. *Br J Psychiatry*, 195(3), 211–217. <https://doi.org/10.1192/bjp.bp.108.051110>
- Rakyan, V. K., Preis, J., Morgan, H. D., & Whitelaw, E. (2001). The marks, mechanisms and memory of epigenetic states in mammals. *The Biochemical Journal*, 356(Pt 1), 1–10.
- Ramboz, S., Oosting, R., Amara, D. A., Kung, H. F., Blier, P., Mendelsohn, M., ... Hen, R. (1998). Serotonin receptor 1A knockout : an animal model of anxiety-related disorder. *Proceedings of the National Academy of Sciences of the United States of America*, 95(24), 14476–14481.
- Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M. C., Lehericy, S., Bergman, H., ... Obeso, J. A. (2010). Goal-directed and habitual control in the basal ganglia : implications for Parkinson's disease. *Nature Reviews. Neuroscience*, 11(11), 760–772. <https://doi.org/10.1038/nrn2915>
- Regoli, D., Boudon, A., & Fauchère, J. L. (1994). Receptors and antagonists for substance P and related peptides. *Pharmacological Reviews*, 46(4), 551–599.
- Rescorla, R. A., Wagner, A. R., & others. (1972). Classical conditioning II : Current research and theory. In A. H. B. [and] William F. Prokasy (Ed.), *Classical conditioning II : Current research and theory* (Vol. 2, pp. 64–99). New York :

Appleton-Century-Crofts.

Ressler, K. J., & Mayberg, H. S. (2007). Targeting abnormal neural circuits in mood and anxiety disorders : from the laboratory to the clinic. *Nature Neuroscience*, 10(9), 1116–1124. <https://doi.org/10.1038/nm1944>

Richter, S. H., Garner, J. P., & Würbel, H. (2009). Environmental standardization : cure or cause of poor reproducibility in animal experiments? *Nature Methods*, 6(4), 257–261. <https://doi.org/10.1038/nmeth.1312>

Richter, S. H., Garner, J. P., Auer, C., Kunert, J., & Würbel, H. (2010). Systematic variation improves reproducibility of animal experiments. *Nature Methods*, 7(3), 167–168. <https://doi.org/10.1038/nmeth0310-167>

Robbins, T. W. (1998). Homology in behavioural pharmacology : an approach to animal models of human cognition. *Behavioural Pharmacology*, 9(7), 509–519.

Robbins, T. W., Gillan, C. M., Smith, D. G., Wit, S. de, & Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity : towards dimensional psychiatry. *Trends in Cognitive Sciences*, 16(1), 81–91. <https://doi.org/10.1016/j.tics.2011.11.009>

Rosenblatt, M. (2017). The Large Pharmaceutical Company Perspective. *The New England Journal of Medicine*, 376(1), 52–60. <https://doi.org/10.1056/NEJMr1510069>

Rossell, S. L. (2012). The Neuroscience of Hallucinations. In R. Jardri, A. Cachia, P. Thomas, & D. Pins (Eds.) (pp. 137–152). Springer.

Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., ... Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps : a STAR\*D report. *The American Journal of Psychiatry*, 163(11), 1905–1917. <https://doi.org/10.1176/ajp.2006.163.11.1905>

Salzer, H. M., & Lurie, M. L. (1953). Anxiety and depressive states treated with isonicotinyl hydrazide (isoniazid). *A.M.A. Archives of Neurology and Psychiatry*, 70(3), 317–324.

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science*, 310(5752), 1337–1340. <https://doi.org/10.1126/science.1115270>

Sams-Dodd, F., Lipska, B. K., & Weinberger, D. R. (1997). Neonatal lesions of the rat ventral hippocampus result in hyperlocomotion and deficits in social behaviour in adulthood. *Psychopharmacology*, 132(3), 303–310.

Santarelli, L., Saxe, M., Gross, C., Surget, A., Battaglia, F., Dulawa, S., ... Hen, R. (2003). Requirement of hippocampal neurogenesis for the behavioral effects

- of antidepressants. *Science*, 301(5634), 805–809. <https://doi.org/10.1126/science.1083328>
- Santos, R., Ursu, O., Gaulton, A., Bento, A. P., Donadi, R. S., Bologa, C. G., ... Overington, J. P. (2017). A comprehensive map of molecular drug targets. *Nature Reviews. Drug Discovery*, 16(1), 19–34. <https://doi.org/10.1038/nrd.2016.230>
- Scahill, L., Sukhodolsky, D. G., Williams, S. K., & Leckman, J. F. (2005). Public health significance of tic disorders in children and adolescents. *Advances in Neurology*, 96, 240–248.
- Scarr, E., Millan, M. J., Bahn, S., Bertolino, A., Turck, C. W., Kapur, S., ... Dean, B. (2015). Biomarkers for Psychiatry : The Journey from Fantasy to Fact, a Report of the 2013 CINP Think Tank. *The International Journal of Neuropsychopharmacology*, 18(10), pyv042. <https://doi.org/10.1093/ijnp/pyv042>
- Scherer, R. W., Langenberg, P., & Elm, E. von. (2007). Full publication of results initially presented in abstracts. *The Cochrane Database of Systematic Reviews*, (2), MR000005. <https://doi.org/10.1002/14651858.MR000005.pub3>
- Schmidt, A., Bachmann, R., Kometer, M., Csomor, P. A., Stephan, K. E., Seifritz, E., & Vollenweider, F. X. (2012). Mismatch negativity encoding of prediction errors predicts S-ketamine-induced cognitive impairments. *Neuropsychopharmacology*, 37(4), 865–875. <https://doi.org/10.1038/npp.2011.261>
- Schönberg, T., Daw, N. D., Joel, D., & O’Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *The Journal of Neuroscience*, 27(47), 12860–12867. <https://doi.org/10.1523/JNEUROSCI.2496-07.2007>
- Schultz, W. (1986). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *Journal of Neurophysiology*, 56(5), 1439–1461.
- Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, 36(2), 241–263.
- Schultz, W. (2013). Updating dopamine reward signals. *Current Opinion in Neurobiology*, 23(2), 229–238. <https://doi.org/10.1016/j.conb.2012.11.012>
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.
- Schwarz, E., Izmailov, R., Spain, M., Barnes, A., Mapes, J. P., Guest, P. C., ... Bahn, S. (2010). Validation of a blood-based laboratory test to aid in the confirmation of a diagnosis of schizophrenia. *Biomarker Insights*, 5, 39–47.
- Segura, B., & Strafella, A. P. (2013). Functional imaging of dopaminergic neurotransmission in Tourette syndrome. *International Review of Neurobiology*, 112, 73–93. <https://doi.org/10.1016/B978-0-12-411546-0.00003-2>
- Selikoff, I. J., & Robitzek, E. H. (1952). Tuberculosis chemotherapy with hydrazine

- derivatives of isonicotinic acid. *Diseases of the Chest*, 21(4), 385–438.
- Seymour, B., Singer, T., & Dolan, R. (2007). The neurobiology of punishment. *Nature Reviews. Neuroscience*, 8(4), 300–311. <https://doi.org/10.1038/nrn2119>
- Shahbazian, M., Young, J., Yuva-Paylor, L., Spencer, C., Antalffy, B., Noebels, J., ... Zoghbi, H. (2002). Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron*, 35(2), 243–254.
- Shapiro, D. A., Renock, S., Arrington, E., Chiodo, L. A., Liu, L.-X., Sibley, D. R., ... Mailman, R. (2003). Aripiprazole, a novel atypical antipsychotic drug with a unique and robust pharmacology. *Neuropsychopharmacology*, 28(8), 1400–1411. <https://doi.org/10.1038/sj.npp.1300203>
- Shih, R. A., Belmonte, P. L., & Zandi, P. P. (2004). A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *International Review of Psychiatry*, 16(4), 260–283. <https://doi.org/10.1080/09540260400014401>
- Singer, H. S. (2005). Tourette's syndrome : from behaviour to biology. *Lancet Neurol*, 4(3), 149–159. [https://doi.org/10.1016/S1474-4422\(05\)01012-4](https://doi.org/10.1016/S1474-4422(05)01012-4)
- Sjoerds, Z., Wit, S. de, Brink, W. van den, Robbins, T. W., Beekman, A. T. F., Penninx, B. W. J. H., & Veltman, D. J. (2013). Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Translational Psychiatry*, 3, e337. <https://doi.org/10.1038/tp.2013.107>
- Skinner, B. F. (2006). *Behavior of Organisms*. Copley Publishing Group. Retrieved from <http://amazon.com/o/ASIN/1583900071/>
- Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., & McGuire, W. L. (1987). Human breast cancer : correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*, 235(4785), 177–182.
- Smith, I., Procter, M., Gelber, R. D., Guillaume, S., Feyereislova, A., Dowsett, M., ... team, H. study. (2007). 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer : a randomised controlled trial. *Lancet*, 369(9555), 29–36. [https://doi.org/10.1016/S0140-6736\(07\)60028-2](https://doi.org/10.1016/S0140-6736(07)60028-2)
- Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *The British Journal of Psychiatry*, 167(1), 99–103.
- Sommer, T., Peters, J., Gläscher, J., & Büchel, C. (2009). Structure-function relationships in the processing of regret in the orbitofrontal cortex. *Brain Structure and Function*, 213(6), 535–551. <https://doi.org/10.1007/s00429-009-0222-8>
- Spilka, M. J., Arnold, A. E., & Goghari, V. M. (2015). Functional activation abnormalities during facial emotion perception in schizophrenia patients and non-

- psychotic relatives. *Schizophrenia Research*, 168(1-2), 330–337. <https://doi.org/10.1016/j.schres.2015.07.012>
- Stahl, S. M. (2013). *Stahl's essential psychopharmacology : neuroscientific basis and practical applications*. Cambridge university press.
- Steeds, H., Carhart-Harris, R. L., & Stone, J. M. (2015). Drug models of schizophrenia. *Therapeutic Advances in Psychopharmacology*, 5(1), 43–58. <https://doi.org/10.1177/2045125314557797>
- Stephan, K. E., & Mathys, C. (2014). Computational approaches to psychiatry. *Current Opinion in Neurobiology*, 25, 85–92. <https://doi.org/10.1016/j.conb.2013.12.007>
- Steru, L., Chermat, R., Thierry, B., & Simon, P. (1985). The tail suspension test : a new method for screening antidepressants in mice. *Psychopharmacology*, 85(3), 367–370.
- Stolerman, I. P., & Price, L. H. (Eds.). (2015). *Encyclopedia of Psychopharmacology* (2nd ed. 2015). Springer. Retrieved from <http://amazon.com/o/ASIN/3642361714/>
- Stoy, M., Schlagenhauf, F., Sterzer, P., Birmahler, F., Hägele, C., Suchotzki, K., ... Ströhle, A. (2012). Hyporeactivity of ventral striatum towards incentive stimuli in unmedicated depressed patients normalizes after treatment with escitalopram. *Journal of Psychopharmacology*, 26(5), 677–688. <https://doi.org/10.1177/0269881111416686>
- Sullivan, P. F., Daly, M. J., & O'Donovan, M. (2012). Genetic architectures of psychiatric disorders : the emerging picture and its implications. *Nature Reviews. Genetics*, 13(8), 537–551. <https://doi.org/10.1038/nrg3240>
- Surget, A., Tanti, A., Leonardo, E. D., Laugeray, A., Rainer, Q., Touma, C., ... Belzung, C. (2011). Antidepressants recruit new neurons to improve stress response regulation. *Molecular Psychiatry*, 16(12), 1177–1188. <https://doi.org/10.1038/mp.2011.48>
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning : An Introduction*. The MIT Press.
- Suzuki, M. (2009). Role of N-methyl-D-aspartate receptor antagonists in post-operative pain management. *Current Opinion in Anaesthesiology*, 22(5), 618–622. <https://doi.org/10.1097/ACO.0b013e32832e7af6>
- Teufel, C., & Fletcher, P. C. (2016). The promises and pitfalls of applying computational models to neurological and psychiatric disorders. *Brain : A Journal of Neurology*, 139(Pt 10), 2600–2608. <https://doi.org/10.1093/brain/aww209>
- Thomas, D. W., Burns, J., Audette, J., Carroll, A., Dow-Hygelund, C., & Hay, M. (2016). *Clinical development success rates 2006 - 2015* (rapport scientifique). Washington, District of Columbia, USA. : Biotechnology Innova-

- tion Organization (BIO). Retrieved from <https://www.bio.org/sites/default/files/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf>
- Thorndike, E. L. (1911). *Animal Intelligence : Experimental Studies*. Macmillan.
- Thorndike, E. L. (1933). A theory of the action of the after-effects of a connection upon it. *Psychological Review*, 40(5), 434.
- Tobia, M. J., Guo, R., Schwarze, U., Boehmer, W., Gläscher, J., Finckh, B., ... Sommer, T. (2014). Neural systems for choice and valuation with counterfactual learning signals. *NeuroImage*, 89, 57–69. <https://doi.org/10.1016/j.neuroimage.2013.11.051>
- Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, 307(5715), 1642–1645. <https://doi.org/10.1126/science.1105370>
- Tolman, E. C., & Honzik, C. H. (1930). Introduction and removal of reward, and maze performance in rats. *University of California Publications in Psychology*.
- Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature*, 398(6729), 704–708. <https://doi.org/10.1038/19525>
- Tremblay, L., Worbe, Y., Thobois, S., Sgambato-Faure, V., & Féger, J. (2015). Selective dysfunction of basal ganglia subterritories : From movement to behavioral disorders. *Movement Disorders*, 30(9), 1155–1170. <https://doi.org/10.1002/mds.26199>
- Turner, R. (2016). Uses, misuses, new uses and fundamental limitations of magnetic resonance imaging in cognitive science. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371(1705). <https://doi.org/10.1098/rstb.2015.0349>
- Valzelli, L. (1973). The “isolation syndrome” in mice. *Psychopharmacologia*, 31(4), 305–320.
- Van Hoeck, N., Watson, P. D., & Barbey, A. K. (2015). Cognitive neuroscience of human counterfactual reasoning. *Frontiers in Human Neuroscience*, 9, 420. <https://doi.org/10.3389/fnhum.2015.00420>
- Van Regenmortel, M. H. V. (2004). Reductionism and complexity in molecular biology. Scientists now have the tools to unravel biological and overcome the limitations of reductionism. *EMBO Reports*, 5(11), 1016–1020. <https://doi.org/10.1038/sj.embor.7400284>
- Veatch, O. J., Veenstra-Vanderweele, J., Potter, M., Pericak-Vance, M. A., & Haines, J. L. (2014). Genetically meaningful phenotypic subgroups in autism spectrum disorders. *Genes, Brain, and Behavior*, 13(3), 276–285. <https://doi.org/10.1111/>

gbb.12117

Vinckier, F., Gaillard, R., Palminteri, S., Rigoux, L., Salvador, A., Fornito, A., ... Fletcher, P. C. (2016). Confidence and psychosis : a neuro-computational account of contingency learning disruption by NMDA blockade. *Mol Psychiatry*. <https://doi.org/10.1038/mp.2015.73>

Vöhringer, P. A., & Ghaemi, S. N. (2011). Solving the antidepressant efficacy question : effect sizes in major depressive disorder. *Clinical Therapeutics*, *33*(12), B49–B61. <https://doi.org/10.1016/j.clinthera.2011.11.019>

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*(6842), 43–48. <https://doi.org/10.1038/35083500>

Wallis, G. G., McHarg, J. F., & Scott, O. C. A. (1949). Acute psychosis caused by dextro-amphetamine. *British Medical Journal*, *2*(4641), 1394.

Wallis, J. D. (2012). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature Neuroscience*, *15*(1), 13–19. <https://doi.org/10.1038/nn.2956>

Waltz, J. A., Frank, M. J., Robinson, B. M., & Gold, J. M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biological Psychiatry*, *62*(7), 756–764. <https://doi.org/10.1016/j.biopsych.2006.09.042>

Wardenaar, K. J., & Jonge, P. de. (2013). Diagnostic heterogeneity in psychiatry : towards an empirical solution. *BMC Medicine*, *11*, 201. <https://doi.org/10.1186/1741-7015-11-201>

Waters, F. (2012). The Neuroscience of Hallucinations. In R. Jardri, A. Cachia, P. Thomas, & D. Pins (Eds.) (pp. 185–199). Springer.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*(3-4), 279–292.

Webb, C. A., Dillon, D. G., Pechtel, P., Goer, F. K., Murray, L., Huys, Q. J. M., ... Pizzagalli, D. A. (2016). Neural Correlates of Three Promising Endophenotypes of Depression : Evidence from the EMBARC Study. *Neuropsychopharmacology*, *41*(2), 454–463. <https://doi.org/10.1038/npp.2015.165>

Wichers, M. C., Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., Peeters, F., Vries, M. de, Mengelers, R., & Os, J. van. (2009). Reduced stress-sensitivity or increased reward experience : the psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology*, *34*(4), 923–931. <https://doi.org/10.1038/npp.2008.66>

Wiecki, T. V., Poland, J., & Frank, M. J. (2015). Model-based cognitive neuroscience approaches to computational psychiatry clustering and classification. *Clinical*

*Psychological Science*, 3(3), 378–399.

Willner, P. (1984). The validity of animal models of depression. *Psychopharmacology*, 83(1), 1–16.

Willner, P. (1986). Validation criteria for animal models of human mental disorders : learned helplessness as a paradigm case. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 10(6), 677–690.

Willner, P. (1997). Validity, reliability and utility of the chronic mild stress model of depression : a 10-year review and evaluation. *Psychopharmacology*, 134(4), 319–329.

Willner, P. (2015). Encyclopedia of Psychopharmacology. In I. P. Stolerman & L. H. Price (Eds.), *Encyclopedia of Psychopharmacology* (2nd ed., pp. 115–121). Berlin, Heidelberg : Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-36172-2\\_336](https://doi.org/10.1007/978-3-642-36172-2_336)

Willner, P., & Belzung, C. (2015). Treatment-resistant depression : are animal models of depression fit for purpose? *Psychopharmacology*, 232(19), 3473–3495. <https://doi.org/10.1007/s00213-015-4034-7>

Willner, P., Borsini, F., & McArthur, R. A. (2008). Animal and Translational Models for CNS Drug Discovery, Vol. 1 : Psychiatric Disorders. In R. A. McArthur & F. Borsini (Eds.) (1st ed., pp. 359–375). Academic Press.

Willner, P., Towell, A., Sampson, D., Sophokleous, S., & Muscat, R. (1987). Reduction of sucrose preference by chronic unpredictable mild stress, and its restoration by a tricyclic antidepressant. *Psychopharmacology*, 93(3), 358–364.

Wit, S. de, Watson, P., Harsay, H. A., Cohen, M. X., Vijver, I. van de, & Ridderinkhof, K. R. (2012). Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *The Journal of Neuroscience*, 32(35), 12066–12075. <https://doi.org/10.1523/JNEUROSCI.1088-12.2012>

Wong, D. F., Brasić, J. R., Singer, H. S., Schretlen, D. J., Kuwabara, H., Zhou, Y., ... Grace, A. A. (2008). Mechanisms of dopaminergic and serotonergic neurotransmission in Tourette syndrome : clues from an in vivo neurochemistry study with PET. *Neuropsychopharmacology*, 33(6), 1239–1251. <https://doi.org/10.1038/sj.npp.1301528>

Wood, S. K., McFadden, K. V., Grigoriadis, D., Bhatnagar, S., & Valentino, R. J. (2012). Depressive and cardiovascular disease comorbidity in a rat model of social stress : a putative role for corticotropin-releasing factor. *Psychopharmacology*, 222(2), 325–336. <https://doi.org/10.1007/s00213-012-2648-6>

Xu, M., Kobets, A., Du, J.-C., Lenington, J., Li, L., Banasr, M., ... Pittenger, C. (2015). Targeted ablation of cholinergic interneurons in the dorsolateral striatum produces behavioral manifestations of Tourette syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 112(3), 893–898.

<https://doi.org/10.1073/pnas.1419533112>

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews. Neuroscience*, 7(6), 464–476. <https://doi.org/10.1038/nrn1919>

Yoon, D. Y., Gause, C. D., Leckman, J. F., & Singer, H. S. (2007). Frontal dopaminergic abnormality in Tourette syndrome : a postmortem analysis. *J Neurol Sci*, 255(1-2), 50–56. <https://doi.org/10.1016/j.jns.2007.01.069>

Zarate, C. A., Singh, J. B., Carlson, P. J., Brutsche, N. E., Ameli, R., Luckenbaugh, D. A., ... Manji, H. K. (2006). A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry*, 63(8), 856–864. <https://doi.org/10.1001/archpsyc.63.8.856>

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, 17(1), 3–18.



# **Annexe 1 - Enhanced habit formation in Gilles de la Tourette syndrome**

## **Annexe 2 - Illusion of agency in patients with Gilles de la Tourette Syndrome**

## **Annexe 3 - Confidence and psychosis : a neuro-computational account of contingency learning disruption by NMDA blockade**

## **Title : Pharmacological challenge of cognition and decision-making : implications for psychiatry**

### **Abstract :**

Successful new drug development has declined in psychiatry in the last decades. This is in part the result of a high failure rate in translating positive preclinical efficacy results to positive clinical trials. Limitations in the validity of animal models and shortcomings in the usefulness of the current categorical diagnostic system.

Cognitive neurosciences and particularly reinforcement learning and its computational analysis might provide biomarkers required to develop new ways of classifying mental disorders on the basis of both observable behaviour and neurobiological measures. Used in conjunction with pharmacological challenges, it may bring new insights into the pathophysiology and brain mechanisms underlying psychiatric disorders. It may also help design new animal models with improved predictive validity for the development of medications relying on innovative mechanisms of action.

We illustrate the use of reinforcement learning and pharmacological challenge in two experimental studies.

In the first experiment, we administered a reinforcement learning task that involves both direct learning from obtained outcomes and indirect learning from forgone outcomes to two groups of Gilles de la Tourette patients, one receiving aripiprazole, one unmedicated and to a group of healthy subjects.

In the second experiment, we administered two probabilistic stimulus selection learning tasks (one simple, one with occasional reversals) to healthy subjects randomly and blindly allocated to either escitalopram, a typical serotonin reuptake inhibitor, agomelatine, an antidepressant with a different mechanism of action, or placebo. The experiment compared the effect of these two classes of antidepressants to placebo after both short term (3 days) and long term (8 weeks) treatment.

These experiments bring insights into the understanding of the clinical condition studied, and the effects of the drugs tested. Implications of this approach for the translational approach to drug development is discussed.

**Mots-clés : Psychiatrie, Psychopharmacologie, Neurosciences cognitives, Apprentissage par renforcement , Modèles animaux**

**Keywords : Psychiatry, Psychopharmacology, Cognitive neuroscience, Reinforcement learning, Animal models**