



**HAL**  
open science

## *Bases cérébrales du compromis coûts/bénéfices*

Nicolas Clairis

► **To cite this version:**

Nicolas Clairis. *Bases cérébrales du compromis coûts/bénéfices*. Psychologie et comportements. Sorbonne Université, 2020. Français. NNT : 2020SORUS026 . tel-03780989

**HAL Id: tel-03780989**

**<https://theses.hal.science/tel-03780989v1>**

Submitted on 20 Sep 2022

**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.

# Sorbonne Université

École doctorale Cerveau Cognition Comportement

*Institut du Cerveau et de la Moelle épinière / Équipe Motivation Cerveau et Comportement*

## **Bases cérébrales du compromis coûts/bénéfices**

Par Nicolas Clairis

Thèse de doctorat de Sciences Cognitives

Dirigée par Mathias Pessiglione

Présentée et soutenue publiquement le 18 septembre 2020

Devant un jury composé de :

PESSIGLIONE	Mathias	Directeur de recherche	Directeur de thèse
PROCYK	Emmanuel	Directeur de recherche	Rapporteur
FOURAGNAN	Elsa	Lecturer	Rapporteuse
ZÉNON	Alexandre	Chargé de Recherche	Examineur
KHAMASSI	Mehdi	Chargé de Recherche	Examineur
THIRION	Bertrand	Directeur de Recherche	Examineur

## Remerciements

ἔν οἶδα, ὅτι οὐδέν οἶδα<sup>1</sup>

Socrate

*Savoir qu'on ne sait pas, c'est la parfaite santé de l'esprit, se figurer qu'on sait, sa pire maladie. Discerner cette maladie, c'est l'éviter : ainsi fait le sage.*

Lao-Tseu

---

<sup>1</sup> « Je sais une chose, c'est que je ne sais rien »

# Remerciements

Je souhaiterai d'abord remercier les membres du jury qui ont accepté d'évaluer les fruits de 4 ans de travail : Emmanuel Procyk, Elsa Fouragnan, Alexandre Zénon, Mehdi Khamassi et Bertrand Thirion. En particulier, j'aimerais remercier grandement Emmanuel Procyk pour m'avoir, avec Céline Amiez, mis sur la route du cortex préfrontal médian (je n'en dirai pas plus de peur de créer une guerre de nomenclatures) et avoir en quelque sorte débuté ma formation et ma rencontre avec l'imagerie par résonance magnétique fonctionnelle quelques mois avant de commencer cette thèse.

Merci aussi aux institutions pour leur financement qui a rendu ce travail possible : l'École Normale Supérieure de Lyon pour mes trois premières années de doctorat et le labex BioPsy pour m'avoir accordé sa confiance pour un financement de quatrième année afin de conclure ces travaux.

Merci en premier lieu à mon encadrant Mathias avec qui j'ai continuellement appris au cours de ces années. Tu m'impressionneras toujours par ta mémoire et ton acuité dans l'identification des atouts ou des failles de certaines études et ta ténacité qui m'a poussé à poursuivre sur certaines pistes que j'aurais probablement abandonné rapidement sans ta persévérance. J'admire toujours ta capacité à sans cesse essayer de nouvelles choses, notamment dans l'organisation de l'équipe, ce qui contribue à améliorer sans cesse la qualité des échanges scientifiques et humains. Tous ces moments d'intenses échanges scientifiques et humains auxquels j'ai pu assister ont été un vrai régal et ont grandement participé à me former scientifiquement. Je me réjouis d'avoir pu bénéficier de ton encadrement tout au long de ces années.

Ironiquement, il n'a pas toujours été facile de garder la motivation à son taux optimal dans cette thèse autour de ladite motivation et je tiens à remercier toutes les personnes qui m'ont soutenues au cours de ces années par leur chaleur et leur amitié. Tout d'abord, je tiens à remercier mes parents pour leur présence et leur soutien indéfectible tout au long de ma thèse et en particulier au cours de ces dernières semaines particulièrement intenses. Cela a été un grand bonheur d'avoir eu l'occasion de partager à nouveau une certaine proximité avec vous après 5 ans loin du foyer familial. Bien évidemment, je tiens aussi à remercier les différentes promotions de la MBB que j'ai pu côtoyées au cours de ces années et qui me donnaient la motivation de revenir le lendemain matin. Merci à Alizée et Delphine, pour votre accueil

## Remerciements

chaleureux dès mon arrivée dans l'équipe, votre amitié qui s'est maintenue depuis et tous les bons moments qu'on a pu passer ensemble à travers le monde, les conférences et les événements. Merci à Stefano, Maël, Vasi, Caroline, Bastien, Chen et Lionel que je n'ai pas tous directement côtoyés au sein de l'équipe mais dont la sympathie m'aura vite fait sentir comme étant membre d'une plus grande famille. Merci à Jules et Emmanuelle dont la compagnie et l'humour m'auront accompagné (presque) tout au long de la thèse. Un grand merci pour votre amitié, les multiples et super voyages faits ensemble, les échanges scientifiques (Jules tu es quasiment mon PI n°2 vu toutes les connaissances que tu m'as transmises pendant cette thèse, un énorme merci !) et humains qu'on a pu avoir au labo, en allant à l'escalade, au bar, au coin d'un feu dans une forêt canadienne et, soyons honnêtes, avant tout et surtout sur WhatsApp, et surtout les multiples conneries qu'on a pu partager pour mettre un peu d'absurde et de bonne humeur pendant ces années pas toujours faciles. Merci à Antonius, avec qui j'ai pu participer à de grands congrès internationaux, mais aussi faire plein de découvertes touristiques et dont la bienveillance éternelle aura aidé à maintenir un climat favorable au sein de l'équipe. Il me faut remercier les « nouveaux » aussi, en particulier, les amis du bunker : Quentin merci pour tous les jeux que tu m'auras fait découvrir, nos énervements mutuels et toutes nos discussions geeky, Claire et Pablo merci pour votre éternelle bonne humeur et aussi merci à l'ami de la salle 4.027, Jacob, pour ta sérénité, ton écoute, tes conseils (dans la vie et dans AoE surtout bien évidemment) et nos petits moments de pétage de câble en plus-ou-moins fin de journée quand l'inhibition sautait. Je n'oublierai jamais toutes les supers parties de jeux que nous avons pu faire tous ensemble et les bons moments passés avec vous, (surtout dans notre cher ~~bunker~~ salle 4.027!). Merci aussi à William et Cynthia pour votre compagnie et les bons moments passés ensemble. Fabien, un énorme merci pour le temps passé à discuter science, mais aussi pour tous les chocolats et multiples mets excellents que tu as offerts à l'équipe et toutes les splendides découvertes de whisky faites grâce à toi au cours des retraites de l'équipe et surtout je te remercie d'avoir su m'aider quand j'en avais le plus besoin. Raphaël, Sébastien et surtout Jean, merci pour tous les échanges scientifiques et les connaissances que vous avez pu me transmettre. Cela a été une vraie émulation de faire partie de cette équipe avec des profils aussi différents (économistes, psychologues, psychiatres, biologistes, physiciens, informaticiens, etc.) et d'avoir eu la chance de bénéficier de moments d'échanges scientifiques un peu informels qui sont parmi ceux où je pense avoir le plus appris, notamment à la retraite d'équipe et aux « cerveaux-tempête ».

## Remerciements

Merci aux manips du CENIR, Mélanie et Antoine notamment, dont la compagnie agréable allégeait significativement la pénibilité du travail en weekend en plein été ! Et merci aux membres de la plateforme Prisme, notamment Natacha Logerot, dont l'aide pour le recrutement des participants a été un atout considérable lors de mes débuts dans l'équipe.

Merci aussi à Baba, Julie, Pupuche, Alex Bacq, Lucas, Nico Brard, Carton, Elise Liu, Lindsay et Oriana dont l'amitié et les pauses café (ou mezcal pour Alex) me manqueront assurément dans mon avenir professionnel. NicoBrard et Julie notre petite équipe de marche-Montparnasse va me manquer ! NicoBrard un énorme merci pour ta relecture aussi ! Enfin, merci aux amis hors de l'Institut, en particulier les amis de Lyon, Mike, Momo, Milou, Jojo, Etienne, Thibault et Judith et à ceux de Grèce Δήμητρα, Κατερίνα, Μαριγούλα et mes cousins Τατιάνα et Αλέξης. Votre amitié a permis à mes batteries de se recharger un nombre incalculable de fois au cours de ses quatre années, je ne sais pas ce que je ferai sans vous ! Merci en particulier à toi Δήμητρα pour ton soutien dans ces mois de confinement que nous avons tous les deux dû traverser avec un rapport à rendre au bout. J'ai beaucoup de chance de tous vous avoir et j'espère bien que nous continuerons à garder contact au fil des pérégrinations des uns et des autres à travers le monde.

## Résumé

« Est-ce que je vais aller acheter du pain ? » Cette question de la vie quotidienne qui paraît pourtant simple engage tout un ensemble de processus cérébraux. En effet, elle requiert d'effectuer un compromis entre les bénéfices escomptés de l'action, c'est-à-dire obtenir une baguette bien chaude pour accompagner le repas, les pertes qu'elle va engendrer, c'est-à-dire le coût monétaire de la baguette, et les coûts de l'action, c'est-à-dire la pénibilité de l'effort impliqué par le déplacement vers la boulangerie et le fait de devoir faire la queue. On peut d'ailleurs passer un temps plus ou moins long à réfléchir à cette question avant de se décider, tant que la confiance dans cette décision n'est pas assez élevée. Cette décision pourrait alors s'accompagner de la sensation d'un effort mental.

Au travers de trois études menées avec des participants sains à l'aide de la technique d'imagerie par résonance magnétique fonctionnelle (IRMf), cette thèse cherche à mieux caractériser les corrélats cérébraux impliqués dans l'estimation du compromis coûts/bénéfices, mais aussi des corrélats cérébraux des variables régulant le calcul de ce compromis.

La première étude visait à identifier des corrélats cérébraux stables, chez le même groupe de participants, de l'estimation du compromis coûts/bénéfices au travers de trois tâches. Le but était également d'essayer de dissocier les structures directement impliquées dans le calcul au premier ordre du compromis coûts/bénéfices, des structures impliquées dans un calcul au second ordre des coûts et des bénéfices de la décision en elle-même. Une décision peut en effet impliquer un coût en termes de temps passé à délibérer et d'effort mental associé à la délibération et des bénéfices en termes de confiance gagnée suite à la délibération. L'étude se décomposait en une tâche où les participants devaient évaluer différents stimuli, une tâche où les participants devaient choisir entre deux stimuli de même nature sur la base de leurs préférences et une tâche où ils devaient choisir d'accepter ou refuser des options impliquant d'effectuer un effort en vue d'une récompense. Cette étude a aussi permis de vérifier si l'estimation du coût subjectif des efforts était basée sur les mêmes bases cérébrales que celles impliquées dans l'estimation de la valeur subjective des récompenses ou si ces deux estimations avaient des bases cérébrales distinctes.

Il a été proposé que le compromis coûts/bénéfices soit soumis à un biais comportemental qui a été caractérisé comme une aversion à la perte. La deuxième étude visait à étudier l'existence éventuelle de l'aversion à la perte dans des tâches impliquant une performance

## Résumé

d'ordre physique ou mental. L'aversion à la perte implique que la perspective d'une perte monétaire a un poids nettement plus important subjectivement que la perspective d'un gain monétaire équivalent. Ce biais, principalement étudié dans le cadre de tâches de choix, engendrerait ainsi une réponse comportementale plus intense en réponse à des pertes équivalentes. À l'inverse, la performance dans cette expérience était légèrement supérieure pour des gains équivalents suggérant plutôt l'influence d'un effet pavlovien, c'est-à-dire une facilitation d'efforts plus vigoureux en réponse à des stimuli appétitifs par rapport à des stimuli aversifs. Cet effet se reflétait aussi dans l'activité du striatum.

La troisième étude visait à étudier si une version simple d'une tâche d'apprentissage par renforcement impliquait l'activité des zones liées à l'exertion d'un effort mental. L'apprentissage par renforcement consiste au fait que la majorité des animaux a tendance à réitérer les actions qui ont mené à l'obtention d'une récompense et au contraire à éviter les actions qui les ont menés à subir un événement aversif. Ce type de tâche pourrait se faire de manière totalement automatique dans l'ensemble du règne animal, y inclus chez l'homme, ou, comme les résultats de cette expérience ont permis de le montrer, elle pourrait requérir un effort au moment de la délibération chez les espèces, comme l'homme, qui sont capables d'exercer du contrôle cognitif.

Ces trois études nous ont ainsi permis de caractériser différents aspects du compromis coûts/bénéfices et de mieux comprendre quand ce compromis nécessitait l'allocation de ressources neurales, exprimées sous la forme de l'utilisation de contrôle cognitif.



# Table des matières

Remerciements .....	2
Résumé .....	5
Table des matières .....	7
Chapitre 1: Introduction .....	11
I    L'estimation des valeurs positives : le réseau des valeurs .....	12
I.1    L'évolution des théories de la valeur .....	13
I.2    Bases neurales de la valeur .....	34
II   Les dévaluateurs de la valeur subjective .....	55
II.1   L'incertitude.....	55
II.2   Le délai.....	64
II.3   L'effort physique .....	71
II.4   L'effort mental .....	80
II.5   Les punitions .....	92
II.6   Le dmPFC : un système de réaction ?.....	99
Résumé intermédiaire .....	112
III   La dynamique de la prise de décision.....	113
III.1   Décomposer la motivation .....	113
III.2   Choix binaires et modèles .....	120
IV   Les variables méta-décisionnelles : le coût décisionnel et la confiance.....	129
IV.1   La confiance dans la décision .....	130
IV.2   Les coûts de la prise de décision.....	136
Chapitre 2: Études expérimentales.....	138
I    La valeur, la confiance et le temps de délibération au moment du compromis coûts/bénéfices ont des supports neuraux distincts dans le cortex préfrontal médian.....	140
I.1   Introduction.....	140

## Table des matières

I.2	Méthodes.....	143
I.3	Résultats.....	158
I.4	Discussion.....	171
I.5	Conclusion.....	174
II	L'effort, tant physique que mental, est plus motivé par l'attrait de gagner que par la peur de perdre.....	176
II.1	Introduction.....	176
II.2	Méthodes.....	178
II.3	Résultats.....	193
II.4	Discussion.....	206
II.5	Conclusion.....	213
III	L'apprentissage par renforcement implique de l'effort mental.....	213
III.1	Introduction.....	213
III.2	Méthodes.....	216
III.3	Résultats.....	224
III.4	Discussion.....	231
III.5	Conclusion.....	233
Chapitre 3:	Discussion générale.....	234
I	Le réseau des valeurs.....	234
I.1	Le réseau des valeurs calcule un compromis coûts/bénéfices.....	234
I.2	Un réseau motivant l'action.....	235
I.3	Qu'est-ce que le vmPFC ?.....	238
I.4	Conclusion sur le réseau des valeurs et la valeur nette.....	240
II	La confiance.....	241
II.1	Valeur et confiance.....	241
II.2	La confiance peut être dissociée de la valeur.....	243

## Table des matières

III	Les ressources neurales allouées au compromis coûts/bénéfices .....	244
III.1	Le dmPFC et la prise de décision.....	244
III.2	Le dmPFC : la vigie du cerveau ?.....	249
III.3	Nature des ressources engagées .....	252
III.4	Le dmPFC <i>versus</i> l'insula antérieure .....	252
Chapitre 4: Conclusion.....		255
Bibliographie.....		258
Appendices .....		296
I	Article en anglais en lien avec la première étude.....	297
II	Enregistrement de l'activité du vmPFC à l'aide de différentes séquences d'IRMf....	329
II.1	Introduction.....	329
II.2	Méthodes.....	329
II.3	Résultats.....	330
II.4	Discussion.....	333
III	Utiliser un modèle computationnel pour inférer les valeurs cachées associées au comportement .....	334
III.1	Introduction.....	334
III.2	Méthodes.....	335
III.3	Résultats.....	337
III.4	Discussion.....	338
IV	Valeur et attention .....	339
IV.1	Introduction.....	339
IV.2	Méthodes d'acquisition dans chaque étude.....	339
IV.3	Résultats.....	342
IV.4	Discussion.....	349
Liste des figures .....		350

## Table des matières

# Chapitre 1: Introduction

Dans le Trésor de la Langue Française<sup>2</sup>, on retrouve plusieurs définitions de la motivation selon les champs d'expertise. On peut trouver ci-dessous les trois principales :

- Philosophie : Action de motiver, d'alléguer les (l'ensemble des) considérations qui servent de motif(s) avant l'acte et de justification à cet acte, a posteriori.
- Psychologie : Ensemble des facteurs dynamiques qui orientent l'action d'un individu vers un but donné, qui déterminent sa conduite et provoquent chez lui un comportement donné ou modifient le schéma de son comportement présent.
- Économie : Ensemble des facteurs qui déterminent le comportement d'un agent économique (acheteur/consommateur, vendeur, etc.) face à un produit ou un service donné.

La définition philosophique est peut-être plus restreinte que les autres en ce qu'elle se rapporte aux motifs qui justifient l'action avant ou après que celle-ci ait eu lieu. Ils doivent donc pouvoir être rapportés. Ceci implique que ces justifications soient conscientes là où les autres définitions incluent autant les facteurs explicites et conscients que les facteurs implicites et non-conscients par l'individu. En résumé, la plupart de ces considérations convergent vers une définition simple qui reviendrait à dire que la motivation est l'ensemble des facteurs qui déterminent la mise en mouvement, c'est-à-dire les actions d'un individu. Le cerveau doit ainsi pouvoir effectuer un rapport bénéfices/coûts incluant l'ensemble de ces facteurs en vue de déterminer quand cela vaut ou non la peine d'engager des ressources en vue de l'atteinte d'un but. Le calcul de ce rapport implique une prise de décision : l'objectif vaut-il la peine de se mettre en action ? Cette prise de décision est ainsi basée sur un compromis entre d'un côté la valeur subjective associée aux coûts des actions à mettre en place pour atteindre un but donné, et de l'autre la valeur subjective associée aux bénéfices escomptés à la suite de la mise en mouvement. Les bénéfices peuvent correspondre à des récompenses à consommer ou à des punitions à éviter. Cette prise de décision constitue une des étapes clés de ce que nous considérons être le processus motivationnel. Elle est ensuite suivie, ou non, d'une autre étape : la mise en action. Enfin, les conséquences de l'action, notamment en fonction de si l'objectif a ou non été atteint et de la quantité de ressources qu'il aura fallu engager pour ce faire entraînent par la suite un

---

<sup>2</sup> <http://stella.atilf.fr>

## Introduction

mécanisme d'apprentissage qui permet d'améliorer les prises de décision futures. Au sein de cette thèse, nous allons ainsi nous intéresser aux différentes étapes de ce processus motivationnel en élucidant, au moins en partie, les mécanismes cérébraux à l'aide de l'imagerie fonctionnelle par résonance magnétique. Plus précisément, nous allons nous concentrer sur les bases cérébrales des différents aspects du rapport entre les bénéfices et les coûts dirigeant la motivation et des bases cérébrales des concepts cognitifs qui régulent ce rapport pendant la prise de décision, et après la prise de décision.

Parce que la motivation et la prise de décision sont aussi centrales à notre vie quotidienne, il est remarquable de constater combien de penseurs de domaines très divers ont pu s'y intéresser. Sans aucunement revendiquer d'être exhaustif, je tenterai de résumer d'abord avec quelques points majeurs les différentes réflexions qui ont eu lieu dans plusieurs domaines quant à la perception subjective des valeurs et l'idée que c'est ces valeurs subjectives qui nous motivent à agir. Nous verrons aussi comment ces réflexions ont évolué au cours du temps. Cela nous emmènera d'abord dans un voyage à travers les domaines aussi larges que la philosophie, la psychologie, les mathématiques ou encore l'économie. Le domaine central de cette thèse sera cependant bien évidemment les sciences cognitives, en particulier dans l'étude de la prise de décision. Nous essaierons ainsi de voir comment l'étude de la prise de décision a pu permettre de percer les bases neurales de la motivation et du compromis coûts/bénéfices qui sous-tend la prise de décision et la motivation. Par la suite, nous verrons par quels mécanismes les valeurs positives peuvent être dévaluées en nous penchant sur les cas de l'incertitude, du délai, de l'effort physique, de l'effort mental et des punitions. Nous essaierons aussi de voir dans quelles mesures ces différentes mesures dévaluant les récompenses peuvent, ou non, être englobées dans un seul mécanisme. Ensuite nous essaierons d'étudier les différents modèles qui ont été proposés pour décrire la dynamique des bases cérébrales du compromis coûts/bénéfices. Enfin, nous nous concentrerons sur deux variables que nous avons appelé des variables « méta-décisionnelles », c'est-à-dire des variables qui régulent la décision sans prendre une part active dans la décision en elle-même. Il s'agira ainsi d'étudier d'un côté la confiance subjective dans la décision et, de l'autre, le coût subjectif lié au compromis coûts/bénéfices.

## **I L'estimation des valeurs positives : le réseau des valeurs**

*« Philosophes écoutez cette phrase est pour vous*

*Le bonheur est un astre volage*

*Qui s'enfuit à l'appel de bien des rendez-vous*

## Introduction

*Il s'efface il se meurt devant nous*

*Quand on croit qu'il est loin il est là tout près de vous*

*Il voyage il voyage il voyage*

*Puis il part il revient il s'en va n'importe où*

*Cherchez-le il est un peu partout... »*

*Le soleil a rendez-vous avec la lune, Charles Trénet*

Dans cette partie, nous allons chercher à mieux comprendre les bases théoriques, comportementales et cérébrales de la motivation, en particulier sous le prisme des récompenses, du plaisir et du bonheur. Cela va nous amener à traverser différentes contrées épistémiques dont j'essaierai de donner un bref aperçu. J'espère que vous apprécierez lire ce cheminement à travers le temps et les disciplines autant que j'ai apprécié en refaire tout l'historique pour les besoins de cette thèse.

### **I.1 L'évolution des théories de la valeur**

#### **I.1.a Philosophie**

*« Le plaisir est le principe et la fin de la vie bienheureuse. »*

*Lettre à Ménécée, Epicure (Laërce, 1999)*

Au moins depuis l'Antiquité gréco-romaine, les philosophes s'accordent sur le fait que **la plupart des hommes cherchent à mener une vie heureuse**. Le stoïcien Sénèque au I<sup>er</sup> siècle après Jésus-Christ écrit par exemple : *« Vivre heureux, Gallio mon frère, c'est ce que veulent tous les hommes, mais, quant à discerner ce qui rend la vie heureuse, ils sont dans les ténèbres. »* (Sénèque, 2005). Quelques siècles plus tôt, Aristote écrivait aussi : *« est simplement final le bien digne de choix en lui-même en permanence et jamais en raison d'un autre. Or ce genre de bien, c'est dans le bonheur surtout qu'il consiste, semble-t-il. Nous le voulons, en effet, toujours en raison de lui-même et jamais en raison d'autre chose. L'honneur, en revanche, le plaisir, l'intelligence et n'importe quelle vertu, nous les voulons certes aussi en raison d'eux-mêmes (car rien n'en résulterait-il, nous voudrions chacun d'entre eux), mais nous les voulons encore dans l'optique du bonheur, dans l'idée que, par leur truchement, nous pouvons être heureux, tandis que le bonheur, nul ne le veut en considération de ces biens-là, ni globalement, en raison d'autre chose. »* (Aristote, 2004).

## Introduction

La plupart des philosophes de l'Antiquité semblent s'accorder sur le fait que mener une vie bienheureuse est un objectif pour la majorité des hommes. Cependant, une question va les préoccuper pendant plusieurs siècles, et il s'agit de savoir **quels sont les critères pour mener une vie heureuse**. En effet, comme le montre Aristote dans l'extrait suivant, chacun semble associer le bonheur à différentes choses : « *Mais le bonheur, qu'est-ce que c'est ? [...] Pour certains, en effet, la réponse est claire et évidente : c'est quelque chose comme le plaisir, la richesse ou l'honneur, quoique la réponse varie des uns aux autres – et souvent même un individu identique change d'avis, puisque tombé malade, il dit que c'est la santé, et dans l'indigence que c'est la richesse.* » (Aristote, 2004). De même, selon les écoles de pensées philosophiques, les courants de pensées vont parfois converger et parfois diverger sur la manière d'atteindre le bonheur. Beaucoup de philosophes vont écrire des sortes de manuels qui regroupent un ensemble de principes visant à mener la vie la plus heureuse qui soit. Un peu comme en cuisine, chacun va proposer ses propres recettes pour conseiller les autres hommes sur les meilleurs moyens d'atteindre le bonheur. Le *Manuel d'Épictète* (II<sup>ème</sup> siècle ap. JC), tout comme les *Pensées pour moi-même* de **Marc-Aurèle** (II<sup>ème</sup> siècle ap. JC), *La Vie heureuse* de Sénèque (I<sup>er</sup> siècle ap. JC), les apophtegmes des **Sept Sages** de la Grèce antique (VII<sup>ème</sup>-VI<sup>ème</sup> siècle av. JC), et la liste est loin d'être exhaustive, se présentent tous comme des manuels d'instruction pour indiquer à tout un chacun comment mener une vie bienheureuse. En Chine, aux VI<sup>ème</sup>-V<sup>ème</sup> siècle av. JC, de même, **Lao-Tseu** écrira le *Tao Tö King (Livre de la Voie et de la Vertu)* et **Confucius** et ses disciples écriront plusieurs écrits dans le but, là encore, de fournir des préceptes pour mener une vie heureuse.

Parmi les philosophes de l'Antiquité, la plupart distinguent d'un côté le **bonheur** et de l'autre le **plaisir**. La distinction principale entre ces deux concepts semble principalement liée à des temporalités différentes : le plaisir est bref et lié à un évènement donné, alors que le bonheur serait une sorte d'état interne qui se maintiendrait sur un temps plus long. **Épicure** est probablement l'un des penseurs de l'Antiquité les plus connus sur ce sujet. Il est souvent présenté, à tort, comme quelqu'un qui verrait le bonheur comme la somme des plaisirs. Il a en effet donné lieu à l'adjectif « épicurien » en français qui signifie, d'après Le Petit Larousse illustré de 2012, « *qui recherche les plaisirs de la vie ; sensuel* ». Certains penseurs, comme **Aristippe de Cyrène** (IV<sup>ème</sup> siècle av. JC) ou **Eudoxe de Cnide** (IV<sup>ème</sup> siècle av. JC), vont revendiquer cette approche du bonheur. Eudoxe de Cnide aurait été l'un des premiers à proposer cette idée selon laquelle, puisque le plaisir est ce que recherchent tous les êtres vivants, il est le souverain bien (Aristote, 2004; Platon, 2018). Aristippe de Cyrène, ancien disciple de Socrate,



## Introduction

fondera, lui, l'**hédonisme** comme doctrine philosophique et l'école de pensée des Cyrénaïques. Le fondement de cette doctrine est la recherche des plaisirs et l'évitement des souffrances comme but de l'existence humaine. Même si plusieurs philosophes tels que Platon ou Speusippe s'y opposent, on voit que cette thèse était relativement populaire dans l'Antiquité où plusieurs jeunes, tel Philèbe, qui s'oppose à Socrate dans le dialogue homonyme de Platon, étaient séduits par cette idée (Platon, 2018). On trouve aussi une résurgence de cette pensée dans l'**utilitarisme hédoniste de Jeremy Bentham** (XVIII<sup>ème</sup> siècle ap. JC). Pour Jeremy Bentham, est utile ce qui participe à augmenter le bien-être en augmentant le plaisir et en réduisant la douleur.

Ironiquement, Épicure se distingue de cette approche qui assimile le bonheur à la somme des plaisirs, bien que ce soit ce que le langage courant en ait retenu plus de 2000 ans après son existence. Il s'agit en effet d'une caricature qu'en faisaient déjà ses opposants dans l'Antiquité de son vivant. L'extrait suivant de sa *Lettre à Ménécée* (Laërce, 1999) permet de montrer qu'il s'est toujours opposé à cette vision caricaturale de sa pensée : « *lorsque nous disons que le plaisir est la fin, nous ne voulons pas parler des « plaisirs des fêtards » ni des « plaisirs qui se trouvent dans la jouissance », comme le croient certains qui, par ignorance, sont en désaccord avec nous ou font à nos propos un mauvais accueil, mais de l'absence de douleur en son corps, et de trouble en son âme.* ». Épicure précise bien que ce qu'il entend par « *absence de douleur en son corps, et de trouble en son âme* » c'est de **ne pas être dans un état de manque** (comme la faim, la soif, etc.). On peut voir cela dans ce passage : « *c'est à ce moment que nous avons besoin d'un plaisir, lorsque nous souffrons par suite de l'absence du plaisir ; < mais lorsque nous ne souffrons pas, nous n'avons plus besoin du plaisir >.* ». Épicure définit donc le bonheur comme l'absence de douleur au niveau physique, état qu'il appellera l'**aponie**, et l'absence de troubles au niveau psychologique, c'est-à-dire l'absence de désirs ou de frustrations en général et de souffrance, état qu'il nommera l'**ataraxie**.

Une autre doctrine prône la **maîtrise des désirs par la raison**, afin d'atteindre le bonheur. Cette doctrine est probablement la plus connue. On la retrouve dans plusieurs courants de pensée. L'un des apophtegmes les plus célèbres des **Sept Sages** de l'Antiquité « *Μηδὲν ἄγαν* » (*Aucun excès*) l'illustre bien. On retrouve d'ailleurs souvent cet apophtegme dans la bouche de **Socrate** via les écrits de **Platon** et cette idée est aussi très présente chez les **Stoïciens**, comme **Épictète**. Platon, dans la conclusion du *Philèbe*, résumera ainsi la doctrine hédoniste à une vision vulgaire : « *tous les bœufs et les chevaux et toutes les bêtes à l'envi témoignent du contraire par leur chasse à la jouissance ; le vulgaire s'y fie, comme les devins aux oiseaux, pour juger que les plaisirs sont les facteurs les plus puissants de la vie bonne, et*

## Introduction

*regarde les amours des bêtes comme des témoins plus autorisés que ne sont les amours nourris aux intuitions rationnelles de la muse philosophique.* » (Platon, 2018). Épicure explique que « *parmi les désirs, certains sont naturels, d'autres vides, et que parmi les désirs naturels, certains sont nécessaires, d'autres seulement naturels ; et parmi les désirs nécessaires, certains sont nécessaires au bonheur, d'autres à l'absence de perturbations du corps, d'autres à la vie même.* » (Laërce, 1999). Il classifie ainsi les désirs en trois catégories : les désirs naturels et nécessaires qui servent une fonction vitale du corps (manger, boire), les désirs naturels non-nécessaires qui sont innés mais ne sont pas nécessaires à la survie du corps (pour la reproduction par exemple) et les désirs non-naturels non-nécessaires. Plus récemment, **Descartes**, en France, au XVII<sup>ème</sup> siècle, dans son *Discours de la méthode* prônera, de même, dans toute une partie consacrée à la maîtrise des désirs, l'idée qu'il faut maîtriser l'objet de ses désirs plutôt que de leur laisser libre cours. L'idée principale des doctrines que nous venons d'évoquer est d'apprendre à maîtriser ses désirs afin de ne laisser libre cours qu'à ceux qu'Épicure appellerait les désirs naturels et nécessaires, et de juger, par la raison, dans quelle mesure les autres désirs peuvent nous apporter du bonheur. Dans ce passage du *Manuel* d'Épictète, Épictète explique bien cela : « *Souviens-toi que le vœu du désir est d'obtenir ce dont il a désir, que le vœu de l'aversion est de ne pas tomber sur l'objet de son aversion. Or, celui qui n'obtient pas ce qu'il désire est infortuné, et celui qui tombe sur l'objet qu'il a en aversion est malheureux. Si donc tu n'as en aversion, dans ce qui dépend de toi, que ce qui est contraire à la nature, tu ne tomberas sur aucun objet d'aversion. Mais si tu as de l'aversion pour la maladie, la mort ou la pauvreté, tu seras malheureux.* » (Épictète, 2007). En résumé, cela revient à la doctrine d'Épicure de **réguler ses désirs** pour limiter, autant que faire se peut, le « trouble de l'âme », c'est-à-dire l'état de frustration lié à la non-satisfaction de certains désirs.

Plus récemment, **le marquis de Sade** (XVIII<sup>ème</sup> siècle ap. JC), à l'opposé direct de la plupart des philosophes de l'Antiquité, revendiquera une **recherche permanente du plaisir et de l'excès sous toutes ses formes**, du moment qu'il y a émotion et excitation. « *Nous voulons être émus [...] c'est le but de tout homme qui se livre à la volupté, et nous voulons l'être par les moyens les plus actifs. En partant de ce point, il ne s'agit pas de savoir si nos procédés plairont ou déplairont à l'objet qui nous sert, il s'agit seulement d'ébranler la masse de nos nerfs par le choc le plus violent possible : or il n'est pas douteux que la douleur affectant bien plus vivement que le plaisir, les chocs résultatifs sur nous de cette sensation produite sur les autres seront essentiellement d'une vibration plus vigoureuse [...] embraseront aussitôt les organes de la volupté et les disposeront au plaisir* » (Sade, 2012). Ce passage, en plus de nous

## Introduction

aider à comprendre d'où vient le mot « *sadisme* » en français, permet de bien comprendre que chez Sade, ce qui compte avant tout, c'est le plaisir et l'excitation, quitte à en passer par la douleur. Là où beaucoup de philosophes, tel Épicure cité précédemment, partaient du principe que le bonheur est lié à l'absence de douleur physique et de désirs psychologiques, le marquis de Sade revendique au contraire de chercher en permanence à renouveler ses désirs et l'excitation associée à leur assouvissement, quitte à passer par la douleur physique pour cette fin. Chez lui, comme dans les philosophies hédonistes, le bonheur passe par l'assouvissement des plaisirs. Cependant, contrairement à ces philosophes qui intègrent généralement l'idée de fuir la douleur, notamment sur le plan physique, Sade voit la douleur au niveau physique comme un moyen supplémentaire d'atteindre le plaisir, et donc le bonheur, qu'il faut cultiver.

Ne pouvant bien sûr pas faire un tour d'horizon exhaustif de la philosophie sur le sujet, je me bornerai à évoquer un dernier grand penseur du désir et des affects qu'est **Baruch Spinoza**. Spinoza, dans son traité *Éthique* publié à sa mort en 1677, voudra avoir une **approche mathématique** de la philosophie et organise tout son ouvrage sur la forme d'un traité mathématique. Là où la plupart des philosophes de l'Antiquité émettaient un jugement de valeur sur les désirs, soit pour les blâmer et dire qu'il faut les réprimer quand ils ne sont pas nécessaires et naturels, soit au contraire pour dire qu'il faut chercher à les satisfaire autant que faire se peut, Spinoza aura une approche différente en cherchant simplement à les analyser, plutôt qu'à les juger. « *Certes n'ont pas manqué les hommes éminents (au labeur et à l'industrie desquels nous avouons devoir beaucoup) pour écrire sur la conduite droite de la vie beaucoup de belles choses, et donner aux mortels des conseils pleins de prudence ; mais, quant à déterminer la nature et la force des Affections, et ce que peut l'Âme de son côté pour les gouverner, nul, que je sache, ne l'a fait. [...] à ceux qui aiment mieux détester ou railler les Affections et les actions des hommes que les connaître [...] il paraîtra surprenant que j'entreprenne de traiter des vices des hommes et de leurs infirmités à la manière des Géomètres et que je veuille démontrer par un raisonnement rigoureux ce qu'ils ne cessent de proclamer contraire à la Raison, vain, absurde et digne d'horreur. Mais voici quelle est ma Raison. Rien n'arrive dans la Nature qui puisse être attribué à un vice existant en elle [...] je considérerai les actions et les appétits humains comme s'il était question de lignes, de surfaces et de solides.* » (Spinoza, 1965) En quelque sorte, Spinoza pose déjà ici les fondements d'une approche rationnelle et mathématique des comportements humains et des désirs humains, un siècle avant les débuts de l'économie avec Adam Smith, et plusieurs siècles avant l'économie comportementale.

## Introduction

Spinoza posera aussi les bases de l'idée de **désirs conscients et inconscients** qui sera plus tard énoncée par Freud en distinguant l'appétit et le désir : *« il n'y a nulle différence entre l'Appétit et le Désir, sinon que le Désir se rapporte généralement aux hommes, en tant qu'ils ont conscience de leurs appétits, et peut, pour cette raison, se définir ainsi : le Désir est l'Appétit avec conscience de lui-même. Il est donc établi par tout cela que nous ne nous efforçons à rien, ne voulons, n'appétons ni ne désirons aucune chose, parce que nous la jugeons bonne ; mais, au contraire, nous jugeons qu'une chose est bonne parce que nous nous efforçons vers elle, la voulons, appétons et désirons. »* (Spinoza, 1965). Spinoza place donc d'abord l'appétit (par exemple le besoin de manger ou de boire pour réguler son homéostasie) avant la conscience (un appétit conscient est un désir, sinon c'est juste un appétit) et le jugement de valeur subjectif. Cette pensée est précisée ailleurs : *« Ce qu'on appelle cause finale n'est d'ailleurs rien que l'appétit humain en tant qu'il est considéré comme le principe ou la cause primitive d'une chose. Quand, par exemple, nous disons que l'habitation a été la cause finale de telle ou telle maison, certes nous n'entendons rien d'autre sinon qu'un homme, ayant imaginé les avantages de la vie de maison, a eu l'appétit de construire une maison. L'habitation donc, en tant qu'elle est considérée comme une cause finale, n'est rien de plus qu'un appétit singulier, et cet appétit est en réalité une cause efficiente, considérée comme première parce que les hommes ignorent communément les causes de leurs appétits. Ils sont en effet, je l'ai dit souvent, conscients de leurs actions et appétits, mais ignorants des causes par où ils sont déterminés à appéter quelque chose. »* Spinoza explique que, derrière chaque action humaine peut se cacher un appétit, c'est-à-dire un but, un besoin à satisfaire, et que l'origine de ce besoin peut souvent être ignorée par le sujet.

Si on résume en quelques points majeurs les pensées des philosophes présentées ici, au risque de caricaturer et déformer légèrement leur pensée, on pourrait dire qu'ils sont tous d'accord pour dire que les humains cherchent à atteindre le bonheur, les règles d'une vie bienheureuse mais qu'ils diffèrent par la manière d'y accéder. Les obstacles au bonheur sont principalement la douleur ou la maladie au niveau physique et la frustration psychologique liée au non-assouvissement de certains désirs. Les désirs naissent d'appétits qui ne sont pas toujours conscients mais qui poussent l'homme à agir pour les assouvir. Certaines écoles de pensée, comme l'hédonisme, prônent la recherche de la satisfaction des désirs et l'évitement des déplaisirs pour atteindre le bonheur. A l'opposé, d'autres écoles de pensée, comme le stoïcisme, prônent la maîtrise des désirs afin de limiter la frustration en limitant les désirs de choses qui ne dépendent pas de soi.

### I.1.b Psychanalyse

En 1901, Sigmund **Freud**, fondateur de la psychanalyse, dans son ouvrage *Sur le rêve*, explique que le rêve aurait un rôle biologique de « *gardien du sommeil* » en nous montrant « *la figuration d'un rêve accompli* » (Freud, 1988). L'idée simple derrière cela est que le corps biologique a besoin du sommeil pour recouvrer des forces alors qu'« *il est clair que les désirs et les besoins qui se manifestent sont les inhibitions qui s'opposent à l'endormissement.* ». Les désirs et les besoins inciteraient les individus à agir en vue de satisfaire leurs désirs. **Le rêve faciliterait donc l'endormissement en montrant que les désirs en cours ont été satisfaits.** L'évolution aurait donc favorisé les rêves pour faciliter la fonction de repos liée à l'endormissement en donnant au sujet l'impression qu'il est en train de satisfaire ses désirs et n'a donc pas besoin de se mettre en action pour ce faire.

Ce cheminement va amener Freud à développer l'idée que l'homme a des **désirs refoulés** du fait de différents tabous sociaux intégrés depuis la plus tendre enfance. De la même manière que Spinoza définissait les appétits comme la cause primitive et inconsciente d'une chose, Freud va postuler l'existence d'une entité, le *Ça*, qui serait le siège de toutes les pulsions inconscientes, de tous les désirs. Chez les enfants, la plupart de ces désirs remonteraient à la conscience. C'est l'obtention de l'objet de ces désirs qui serait principalement observée au cours des rêves des enfants. Cependant, à l'âge adulte, les tabous sociaux, qui seraient stockés dans le *Surmoi* empêcheraient certains de ces désirs de remonter à la conscience, le *Moi*. L'interprétation des rêves, des lapsus et des oublis par un psychanalyste permettrait justement d'accéder à ces désirs refoulés dans l'inconscient. L'idée de Freud est qu'il existe des désirs qui dirigent le comportement mais auxquels on ne peut pas accéder simplement en demandant au sujet conscient lesquels ils sont ou en observant son comportement parce qu'ils sont refoulés du fait de tabous sociaux. Le but du psychanalyste serait donc d'identifier les désirs refoulés pour aider le patient à aller mieux en les acceptant.



**Figure 1 : Célèbre caricature anonyme du début du <sup>XX</sup><sup>ème</sup> siècle censée résumer la pensée de Freud.** Titre original : *What's on a man's mind (Ce qu'il y a dans l'esprit d'un homme)*.

Beaucoup de caricatures présentent Freud comme quelqu'un qui verrait des désirs sexuels dans toutes les actions humaines (voir **Figure 1**). Néanmoins, « *Freud donne à la sexualité une signification plus large. La sexualité est l'ensemble des pulsions et des comportements qui visent, directement ou indirectement, un plaisir* » (Bégorre-Bret, 2011). En résumé, là encore Freud voit dans la recherche des plaisirs la composante principale qui motiverait le comportement. Freud, comme les philosophes qui l'ont précédé, voit dans le plaisir et le désir du plaisir un facteur primordial qui cause les comportements humains.

Par la suite, influencé par l'expérience de la première guerre mondiale de 1914-18, il va rajouter une nouvelle notion, l'idée de la pulsion de Mort. Il propose que les individus soient non seulement dirigés par la recherche de satisfaction de leurs plaisirs « sexuels » qu'il va appeler l'Eros ou libido, mais aussi par une force destructrice, la pulsion de mort qui les pousserait à agir en vue de détruire tout ce qui est vivant à l'état inorganique.

Sans aller dans l'héritage de Freud avec les différents embranchements qui ont suivi et aussi les critiques sur l'irréfutabilité de sa théorie (Popper, 2002), il faut concéder à Freud d'avoir mis en avant ce que Spinoza évoquait déjà c'est-à-dire que certains désirs relèvent du domaine de l'inconscient mais dirigent cependant nos comportements.

### I.1.c Psychologie et physiologie

*« Si l'âme a été affectée une fois de deux affections en même temps, sitôt que plus tard elle sera affectée de l'une, elle sera affectée aussi de l'autre. »*

*Proposition XIV, Livre III, Éthique, Spinoza (Spinoza, 1965)*

L'apprentissage par renforcement consiste à renforcer la valeur d'actions et/ou de stimuli qui ont été associés avec l'obtention de récompenses et, au contraire, à dévaluer des stimuli ou des actions associés au fait de subir une punition qui peut être d'ordre physique, comme la douleur, ou d'ordre psychologique, comme la perte d'argent. L'apprentissage par renforcement est une capacité relativement universelle. On l'observe chez la plupart des Vertébrés et il a même été proposé que les plantes aussi sont capables d'apprendre par renforcement en associant un flux d'air neutre issu d'un ventilateur à la présence de lumière (Gagliano et al., 2016). Cette idée s'est développée en deux grands embranchements.

D'un côté, elle a donné lieu au **conditionnement « classique »**, développé par le physiologiste russe Ivan **Pavlov** à partir de 1889. Il montrera, avec sa fameuse expérience, que des chiens ayant régulièrement associé le son d'une cloche à l'obtention de nourriture, finissent par saliver non seulement à la vue de la nourriture mais dès qu'ils entendent le son de cloche, par anticipation de la nourriture qui va arriver. Même une fois que le son de cloche n'est plus suivi de l'obtention de nourriture, les chiens continuent de saliver quand ils entendent le son de cloche. Les chiens associent un stimulus initialement neutre (le son de cloche) à un événement positif (la nourriture). On dira du son de cloche que c'est un **stimulus conditionné**, parce qu'il a été conditionné à être perçu comme positif et à déclencher une réaction physiologique, alors que la nourriture est un **stimulus inconditionné** car elle déclenche une réaction de manière réflexe (ici la salivation), sans avoir besoin d'apprentissage. Le psychologue américain **John Watson** répliquera ces résultats chez l'homme, avec une expérience éthiquement discutable, en reprenant les principes du conditionnement chez un enfant âgé de 11 mois, le petit Albert, chez lequel il va associer la réaction aversive à un son bruyant (stimulus inconditionné) à la présence d'animaux ou d'objets à fourrure (Watson and Rayner, 1920). Dans le conditionnement « classique », ce qui est donc renforcé est l'association entre un stimulus observé passivement et un événement positif ou négatif.

De l'autre côté, elle a donné lieu au développement du conditionnement opérant et au behaviorisme. L'idée centrale de cette approche est que ce n'est pas un stimulus neutre, comme dans le conditionnement « classique », mais une action du sujet, qui va être renforcée, dans le

## Introduction

cas où elle est suivie d'une récompense, ou, au contraire, dévaluée, si elle est suivie d'une punition. Entre la fin du XIX<sup>ème</sup> et le début du XX<sup>ème</sup> siècle, le psychologue américain Edward Lee **Thorndike** s'intéresse à l'apprentissage. En 1898, il démontrera dans sa thèse que des chats et des chiens devant s'échapper d'un labyrinthe pour avoir accès à de la nourriture, sortent de plus en plus vite de la cage car ils apprennent quelles sont les actions qui leur ont permis de s'échapper du labyrinthe (Thorndike, 1898). Plus tard, plusieurs expériences viendront confirmer que l'homme aussi est capable d'apprendre de manière similaire (Thorndike, 1913). Cela permettra à Edward Lee Thorndike d'énoncer, en 1911, sa **loi de l'effet**: un comportement accompagné ou rapidement suivi d'une récompense sera associé à la situation qui l'a déclenché (Thorndike, 1911). Cette loi sera un des fondements du concept de **conditionnement opérant**, ou **conditionnement instrumental**, développé plus tard par Burrhus Frederic **Skinner**. Skinner développera toute une série d'expériences chez des chats, des pigeons et des rats qui va confirmer cette idée. L'idée basique de cette théorie est qu'une action, un comportement *actif*, est renforcée positivement (ou négativement) si elle a été associée temporellement à un événement positif (ou négatif).

Ces principes donneront lieu plus tard au développement des modèles d'apprentissage par renforcement que nous verrons plus loin (voir *Mathématiques*).

### I.1.d Économie

*« Le plaisir est le principe et la fin de la vie bienheureuse. Car c'est le plaisir que nous avons reconnu comme le bien premier et congénital, et c'est à partir de lui que nous commençons à choisir et refuser, et c'est à lui que nous aboutissons, en jugeant tout bien d'après l'affection prise comme règle. »*

*Lettre à Ménécée, Épicure (Laërce, 1999)*

Beaucoup voient la date de la naissance de l'économie dite classique dans la publication de l'ouvrage *La richesse des nations* d'Adam Smith en 1776. Comme on l'a vu précédemment, dès le IV<sup>ème</sup> siècle av. JC, Épicure distinguait les désirs naturels et nécessaires des désirs naturels non-nécessaires et des désirs non-naturels. Dans une logique similaire, **Adam Smith** en 1776 pose le **paradoxe de l'eau et du diamant** sous la forme suivante : *« Il n'y a rien de plus utile que l'eau, mais elle ne peut presque rien acheter ; à peine y a-t-il moyen de rien avoir en échange. Un diamant, au contraire, n'a presque aucune valeur quant à l'usage, mais on trouvera fréquemment à l'échanger contre une très grande quantité d'autres marchandises. »* (Smith, 1776). L'un des deux biens, l'eau, qui répond à un désir nécessaire, a un prix très faible.



## Introduction

L'eau ne peut être échangée que contre très peu de choses. Les diamants qui correspondent à un désir ni naturel, ni nécessaire, peuvent au contraire être échangés contre énormément de biens. Adam Smith en tirera deux concepts : la valeur d'échange et la valeur d'usage. La **valeur d'usage** d'un bien correspond à l'utilité de posséder ou de consommer un bien. La **valeur d'échange** d'un bien correspond à la quantité d'autres biens que celui-ci permet d'obtenir.

**David Ricardo** et **Adam Smith** lui-même répondent en partie à ce paradoxe de l'eau et du diamant en postulant que la valeur d'échange dépend des **coûts de production**. Adam Smith écrit déjà en effet : « *En tout temps et en tous lieux, ce qui est difficile à obtenir ou ce qui coûte beaucoup de travail à acquérir est cher, et ce qu'on peut se procurer aisément ou avec peu de travail est à bon marché.* » (Smith, 1776). David Ricardo va aussi reprendre la même idée quand il écrit « *tout augmente ou baisse de valeur à proportion de la facilité ou de la difficulté de production, ou, en d'autres mots, à proportion de la quantité de travail employée dans la production* » (Ricardo, 2002). Dans la vie quotidienne, de nombreuses personnes ont des difficultés avec ces concepts quand elles s'offusquent qu'un plat dans un restaurant coûte nettement plus cher que le même plat cuisiné chez soi. Elles ignorent les coûts de production (salaire du personnel en cuisine et au service, aménagement du lieu, etc.). Cela donnera naissance à la « *labor theory of value* ». Cette théorie se traduit en français par le concept de « *valeur-travail* ». L'idée principale de cette théorie, qui sera notamment reprise par Karl Marx, est que la valeur d'un bien dépend du nombre d'heures et de l'effort qui a été dépensé pour produire le bien en question. Cela expliquerait que les diamants coûtent plus chers que l'eau. Les diamants nécessitent en effet plus d'efforts pour être produits que l'eau qu'on trouve, en tout cas à l'époque, naturellement à l'état buvable dans l'environnement. De même, cela expliquerait pourquoi le même plat au restaurant ou chez soi n'a pas le même prix puisque dans le second cas, c'est vous-même qui produisez l'effort pour cuisiner le plat en question. Adam Smith propose que ce phénomène explique non seulement les différences des prix entre différents biens mais aussi la fluctuation des prix pour un bien donné. « *Dans le seizième siècle, la découverte des mines fécondes de l'Amérique réduisit la valeur de l'or et de l'argent, en Europe, à un tiers environ de ce qu'elle avait été auparavant. Ces métaux, coûtant alors moins de travail pour être apportés de la mine au marché, ne purent plus acheter ou commander, quand ils y furent venus, qu'une moindre quantité de travail, et cette révolution dans leur valeur, quoique peut-être la plus forte, n'est pourtant pas la seule dont l'histoire nous ait laissé des témoignages.* » (Smith, 1776). Dans la même veine d'idée, David Ricardo propose que l'invention de machines qui réduisent la quantité de travail nécessaire pour produire un bien

## Introduction

donné va nécessairement amener à une baisse des prix du fait de la concurrence : « *Sans doute, celui qui découvre une machine ou qui en fait le premier l'application, doit, pendant quelques années, jouir d'avantages spéciaux et légitimes et de profits énormes ; mais l'emploi de sa machine se généralisant peu à peu, le prix de la marchandise produite descendrait, sous la pression de concurrence, au niveau des frais de production, et le capitaliste verrait baisser ses profits.* » (Ricardo, 2002). En résumé, l'une des composantes principales expliquant la variation des prix d'échange serait liée à la quantité d'effort nécessaire pour produire un bien.

Très vite, les limites de cette théorie sont apparues. En effet, si l'on regarde les prix de certains items à notre époque par exemple, on peut voir que le gilet en laine de Kurt Cobain, ex-chanteur de Nirvana décédé en 1994, jamais lavé et avec un trou de cigarette, a été vendu aux enchères 334.000 dollars à New York en octobre 2019. De même, un fan a vendu un faux-ongle de la chanteuse Lady Gaga 9.000 euros en juin 2013<sup>3</sup>. A l'inverse, aujourd'hui, il est facilement possible d'obtenir un téléphone portable pour un prix dérisoire de quelques dizaines à quelques centaines d'euros. Or un téléphone portable implique un coût de production conséquent en termes d'heures de travail, si l'on considère l'extraction de tous les minerais rares nécessaires à sa production, les trajets effectués pour acheminer les différents composants, le nombre de personnes ayant travaillé à sa fabrication, etc.<sup>4</sup>. Ces exemples montrent bien que le coût de production ne suffit pas à lui seul à expliquer l'évolution des prix. En effet, dans un cas, les prix sont extrêmement élevés pour un travail quasi-nul. Dans l'autre cas, le prix est relativement faible pour un travail au contraire très élevé. Un nouveau courant de pensée qui contribuera à la formation des écoles dites néo-classiques et autrichiennes va résoudre ce paradoxe quelques dizaines d'années plus tard grâce à ce qu'on appelle la « **révolution marginaliste** ». David Ricardo, en 1817, a déjà l'intuition que la disponibilité d'un bien va déterminer son prix quand il écrit : « *le brasseur, le distillateur, le teinturier, emploient continuellement l'air et l'eau dans la fabrication de leurs produits ; mais comme la source de ces agents est inépuisable, ils n'ont point de prix. Si la terre jouissait partout des mêmes propriétés, si son étendue était sans bornes, et sa qualité uniforme, on ne pourrait rien exiger pour le droit de la cultiver, à moins que ce ne fût là où elle devrait à sa situation quelques avantages particuliers. C'est donc uniquement parce que la terre varie dans sa force*

---

<sup>3</sup> <https://www.msn.com/fr-fr/divertissement/celebrities/stars-et-ench%C3%A8res-les-objets-vendus-les-plus-improbables/ar-AAJXr7I>

<sup>4</sup> [https://www.lemonde.fr/planete/article/2016/10/01/le-sulfureux-parcours-du-tel%C3%A9phone-portable-des-mines-aux-fili%C3%A8res-clandestines-de-dechets\\_5006655\\_3244.html](https://www.lemonde.fr/planete/article/2016/10/01/le-sulfureux-parcours-du-tel%C3%A9phone-portable-des-mines-aux-fili%C3%A8res-clandestines-de-dechets_5006655_3244.html)

## Introduction

*productive, et parce que, dans le progrès de la population, les terrains d'une qualité inférieure, ou moins bien situés, sont défrichés, qu'on en vient à payer une rente pour avoir la faculté de les exploiter.* » (Ricardo, 2002). Trois économistes européens, William Stanley Jevons en Angleterre (1871), Carl Menger en Autriche (1871) et Léon Walras en France (1874), ont tous écrit presque en même temps pour formaliser cette idée que la valeur d'échange dépend de la disponibilité du bien. L'idée est très simple, si l'accessibilité de l'eau est grande, je refuserai de payer cher pour un bien que j'ai déjà en grande quantité. A l'inverse, le gilet de Kurt Cobain est unique au monde, puisqu'aucun autre gilet ne pourra avoir eu son historique. Posséder ce gilet revient donc à posséder quelque chose d'unique au monde que personne d'autre ne pourra posséder, même si la valeur d'usage en est extrêmement faible. Léon Walras reprend l'idée de David Ricardo en écrivant : « *L'utilité ne suffit donc pas pour créer la valeur : il faut encore que la chose utile n'existe pas en quantité illimitée, qu'elle soit rare. Ce raisonnement est confirmé par les faits. L'air respirable, le vent qui enfle les voiles des navires et fait tourner les moulins, la lumière du soleil qui nous éclaire et sa chaleur qui mûrit les moissons et les fruits, l'eau et la vapeur qu'elle fournit une fois échauffée, beaucoup d'autres forces de la nature sont utiles, mêmes nécessaires. Et cependant elles n'ont pas de valeur. Pourquoi ? Parce qu'elles sont illimitées en quantité, parce que chacun de nous peut en obtenir, quand il y en a, autant qu'il en peut désirer sans rien donner, sans faire aucun sacrifice en échange.* » (Walras, 1874). Les membres de la révolution marginaliste vont formaliser cette idée en termes mathématiques pour montrer comment la valeur varie en fonction de la disponibilité d'un bien.

Ces économistes ont aussi postulé que les individus, quand ils agissent, doivent chercher à maximiser leur satisfaction, aussi appelée *utilité*. Ils proposent que le plaisir retiré d'un litre supplémentaire d'eau quand un individu en possède déjà plusieurs centaines de litre est faible, alors que le plaisir retiré de posséder un gilet en laine de Kurt Cobain quand on en possède aucun doit être élevé. De la même manière que les philosophes partaient du principe que « *vivre heureux, [...] c'est ce que veulent tous les hommes* » (Sénèque, 2005), les économistes postulent que maximiser son utilité, c'est ce que veulent tous les hommes.

**Vilfredo Pareto** initiera une nouvelle révolution dans les sciences économiques en 1906 (Pareto, 1909). En effet, il remarquera que la plupart des théories économiques jusque-là étaient basées sur 2 postulats : 1) l'utilité existe et 2) les individus agissent pour la maximiser. Or, aucune preuve n'étayait ce fait jusque-là. En effet, Pareto signale que les seules variables auxquelles les économistes ont accès sont les prix et les choix faits par les individus. En se basant uniquement sur leurs choix on peut classer et ordonner par ordre croissant de

## Introduction

préférence les items choisis. Par exemple, si entre A et B, le sujet choisit A et qu'entre B et C, le sujet choisit B, je peux ordonner ces items de sorte que  $A > B > C$ . Cependant, cette classification ne me dit absolument rien sur l'utilité supposée de chacun de ces items. On pourrait avoir  $A = 3 > B = 2 > C = 1$  (en unité d'utilité), tout comme  $A = 5000 > B = 2 > C = 1$  et l'ordre de préférence serait le même. Pareto insistera donc sur le fait que, pour les différentes théories économiques visant à expliquer les choix, seule la valeur ordinale des options est nécessaire (c'est-à-dire leur rang dans l'ordre des préférences) et non la valeur cardinale associée à chaque item. Tout ce qu'il faudrait savoir pour expliquer le choix d'un individu entre deux options A et B serait si A est supérieur à B, et il n'est pas besoin de savoir à quel point A est supérieur à B. Cette notion aura un impact conséquent et long sur le domaine, encore aujourd'hui, puisqu'il change radicalement le paradigme en passant d'une recherche de l'identification de l'utilité cardinale qui expliquerait les choix (quelle est la satisfaction associée à un objet), à une **utilité ordinale**. Les sciences économiques, à partir de là, ne s'intéressent plus tant à l'utilité associée à un objet en terme quantitatif (j'identifie les utilités de A et B,  $u_A$  et  $u_B$ , donc j'en déduis  $u_A > u_B$ ), mais plutôt en terme qualitatif (est-ce que  $u_A > u_B$  ?).

**Paul Samuelson**, dans son article de 1938 (Samuelson, 1938), poussera les choses plus loin dans sa recherche de débarrasser définitivement la théorie économique des éléments inobservables. Dans son questionnement, il poursuivra la démarche de Pareto de remettre en question l'utilité comme fondement des choix. Pareto, on l'a vu, avait déjà remis en question deux postulats qui sous-tendaient la science économique jusque-là : 1) l'utilité existe et 2) les individus cherchent à la maximiser. Observer les choix des individus serait un moyen indirect de voir comment ils maximisent leur utilité. Samuelson ira plus loin en inversant le paradigme : plutôt que de partir du postulat que la maximisation de l'utilité est ce qui cause les choix, il va proposer de voir si les choix faits par les individus sont cohérents avec une théorie de la maximisation de l'utilité. Il transforme le postulat en hypothèse testable. Cela consiste en un changement majeur de point de vue qui perdure encore de nos jours. Il va ainsi proposer, en se concentrant uniquement sur les choix, une série d'axiomes permettant de voir si, oui ou non, les comportements sont cohérents avec une utilité qu'il s'agirait de maximiser. Ces différents axiomes que nous allons voir maintenant font partie de la **théorie de la préférence révélée** développée par Samuelson. Le premier axiome est l'**axiome faible de la préférence révélée** (*Weak Axiom of Revealed Preference*, WARP). Cet axiome consiste simplement à dire que si un individu choisit une option A plutôt qu'une option B, il faut que  $u_A \geq u_B$ . Mettons un individu qui choisit l'option A lors d'un premier choix. Une autre fois, en face des mêmes

## Introduction

options, il choisit B au lieu de A, ou, pire, en face de A avec un supplément monétaire et B sans supplément, il choisit B au lieu de A. Dans ces deux situations, on pourra dire que l'individu n'a pas respecté cet axiome et on ne pourra donc pas dire qu'il maximise une utilité interne. Il faut préciser à ce stade que cet axiome part du principe que les préférences, c'est-à-dire l'utilité associée à chaque option, sont stables. Nous verrons plus tard que ceci est complètement faux mais cela marque, encore aujourd'hui, certains domaines de la science économique et de la prise de décision.

L'économiste **Hendrik Houthakker** complètera la théorie de Samuelson en 1950 (Houthakker, 1950) avec l'**axiome généralisé de la préférence révélée** (*Generalized Axiom Of Revealed Preference*, GARP) qui dit simplement que si un individu choisit A plutôt que B et B plutôt que C, alors, face à un choix entre A et C, il devrait choisir A (car  $u_A > u_B$  et  $u_B > u_C$ , donc  $u_A > u_C$ ). Houthakker montrera que ce critère de **transitivité** des préférences constituera une condition nécessaire et suffisante pour justifier une maximisation de l'utilité. Autant Samuelson que Houthakker ne prétendent pas démontrer que les individus se comportent en maximisant une utilité interne s'ils obéissent à ces axiomes, mais plutôt démontrer que la proposition est falsifiable puisque, si les individus violent ces axiomes dans leur comportement, ce serait incompatible avec la proposition selon laquelle ils maximisent une utilité. Il faut garder en tête que, selon la même logique qu'en statistiques où montrer que l'hypothèse nulle  $H_0$  n'est pas fausse, n'en démontre pas la véracité, mais juste que l'hypothèse  $H_1$  est fausse, les auteurs ont posé des conditions pour invalider la théorie. Ces axiomes formeront la base de l'idée d'agent économique rationnel. On dira en effet d'un individu qui n'obéit pas au WARP ou au GARP qu'il est un individu irrationnel économiquement parlant.

### I.1.e Mathématiques

En 1654, le chevalier de Méré aurait demandé de l'aide à son ami **Blaise Pascal** pour savoir comment répartir les gains si on interrompait une partie d'un jeu de hasard avant qu'un des deux joueurs ait gagné. Ceci amènera Pascal à une série d'échanges épistolaires avec un autre mathématicien de renom, Pierre de Fermat, qui donnera lieu à la naissance de la théorie des probabilités d'après plusieurs autres mathématiciens<sup>5</sup>. C'est à la suite de ces échanges épistolaires en effet que Pascal proposera la formule de l'**espérance** :

$$E = p * x$$

---

<sup>5</sup> <https://www.pourlascience.fr/sd/histoire-sciences/une-partie-de-des-avec-pascal-2432.php>

## Introduction

où  $E$  est l'espérance, c'est-à-dire la valeur attendue compte tenu des conditions,  $x$  les gains en jeu et  $p$  la probabilité d'obtenir ces gains. On appelle aussi cette variable, la **valeur attendue**. Dans le cas où il faut choisir entre 2 options de probabilités et de montants différents, il suffirait de comparer l'espérance de chacune des options et choisir en conséquence.

Une des limites de cette approche purement mathématique pour expliquer les comportements humains est qu'elle n'apparaît pas réaliste par certains aspects. En effet, Nicolas Bernoulli présentera en 1713 un paradoxe lié à ce théorème que l'on appelle le **paradoxe de Saint-Pétersbourg**. Si l'on considère un jeu où le montant obtenu est proportionnel au nombre de lancers d'une pièce qu'il faut faire avant qu'elle retombe sur face (par exemple  $2^k$ € au  $k$ -ième lancer) et que cette pièce n'est pas truquée (donc la probabilité étant de  $\frac{1}{2}$  pour chaque lancer, on a une probabilité de  $(\frac{1}{2})^k$  d'obtenir le montant correspondant au  $k$ -ième lancer), quel serait le montant maximum qu'il faudrait être prêt à investir pour jouer ? L'espérance de gain du jeu tend vers l'infini car elle est égale à la somme des gains espérés pour chaque lancer  $E = \sum_{i=1}^{\infty} \frac{1}{2^i} * 2^i = \sum_{i=1}^{\infty} 1 = \infty$ . Sur la base de ce calcul, il faudrait donc être prêt à investir tout l'argent possédé pour jouer puisque l'espérance tend vers l'infini. Le cousin de Nicolas, **Daniel Bernoulli** résoudra ce paradoxe dans un article du journal de l'Académie Impériale de la Science de Saint-Pétersbourg en 1738 (Bernoulli, 1738) en montrant que ce que font les individus n'est pas tant de maximiser l'espérance de gains (la valeur attendue), mais plutôt **l'utilité espérée**. Dans cet article, Daniel Bernoulli expliquera « *la détermination de la VALEUR d'un article ne sera pas fondée sur son prix, mais sur l'UTILITE qu'il procure. Le prix d'un article ne dépend que de l'objet lui-même ; il est le même pour toute personne ; en revanche l'utilité dépend des conditions particulières où se trouve la personne qui a la charge de l'évaluation. Ainsi il n'est pas douteux que le gain de mille ducats est plus important pour un homme pauvre que pour un riche, bien que le montant soit le même pour l'un et pour l'autre.* »<sup>6</sup>. Pour résoudre le paradoxe, il proposera d'utiliser la fonction logarithmique de sorte qu'au lieu de  $E = \sum_{i=1}^{\infty} \frac{1}{2^i} * 2^i = \infty$ , on a  $E = \sum_{i=1}^{\infty} \frac{1}{2^i} * \ln(2^i) = \ln(2) \ll \infty$ . On obtient donc bien une fonction bornée qui ne tend plus vers l'infini et qui permet d'avoir une idée du montant à investir dans le jeu, bien que la fonction choisie soit arbitraire. La formule générale s'écrit :

$$E = \sum P(x) * u(x)$$

---

<sup>6</sup> <https://www.lesmotivations.net/spip.php?article118>

## Introduction

où  $P(x)$  est la probabilité de  $x$  et  $u(x)$  est la transformation de  $x$  par une fonction inconnue qui donne l'utilité de  $x$  et elle remplace l'ancienne formule:  $E = \sum P(x) * x$ . Le plus intéressant dans l'approche de Bernoulli est qu'il ajoute la subjectivité et la psychologie dans l'équation via la transformation par la fonction d'utilité. En somme, un siècle avant la « révolution marginaliste » (voir *Économie*), Bernoulli posait déjà les bases de l'idée via sa théorie de l'utilité espérée que les marginalistes reprendront sous le terme d'utilité marginale.

Deux siècles plus tard, les mathématiciens reviennent sérieusement s'intéresser aux choix économiques via deux mathématiciens, **John Von Neumann** et **Oskar Morgenstern**. Reprenant l'approche de Samuelson de 1938 sur la théorie des préférences révélées (voir *Économie*), ils voulurent, comme leurs prédécesseurs mathématiciens avant eux, réintroduire les probabilités dans l'équation. En effet, les économistes jusque-là s'étaient peu intéressés aux probabilités. Von Neumann et Morgenstern, dans leur article de 1944, proposent d'imaginer deux situations : d'un côté, Robinson Crusoé, isolé sur son île, confronté à certaines données matérielles (des besoins et différentes ressources dans son environnement) cherchant à maximiser sa satisfaction ; de l'autre, un individu dans une société avec des échanges économiques. Ils expliquent que dans le cas de Robinson Crusoé, aucun des éléments avec lesquels il doit maximiser sa situation n'a à voir avec la volonté ou l'intention de nature économique d'un autre individu qui aurait des motivations similaires aux siennes. Dans l'autre cas, au contraire, l'individu est confronté à des données matérielles, notamment les prix, qui varient non seulement en fonction des actions de l'individu, mais avant tout en fonction des actions et désirs des autres participants. Les actions de l'individu seront de ce fait altérées par son attente sur l'évolution des autres individus et aussi sur l'attente des autres individus sur ses actions à lui (Von Neumann and Morgenstern, 1944). Ils développent ainsi une **théorie des jeux en milieu compétitif** qui donnera lieu au développement, d'un côté de la **théorie des jeux** qui sera reprise ensuite dans les sciences sociales, les sciences politiques ou la biologie évolutionniste, de l'autre ils apportent leur grain à la **théorie de l'utilité espérée**. Leur apport à cette seconde théorie se résume en quatre axiomes fondamentaux qui incluent le GARP (voir *Économie*) et qui permettent encore une fois de préciser les conditions à remplir pour pouvoir dire d'un sujet qu'il maximise une fonction d'utilité interne. Leur apport va aussi permettre de montrer comment prendre en compte la probabilité associée à chacune des options. Leurs apports fondamentaux sont, d'un côté, de rendre la théorie de leur prédécesseur mathématicien

## Introduction

Bernoulli falsifiable (Caplin and Glimcher, 2014), et, de l'autre, d'inclure les probabilités dans les théories économiques.

Von Neumann et Morgenstern avaient rajouté les probabilités dans la théorie de l'utilité espérée, mais uniquement dans des circonstances où les probabilités sont connues des individus. Or, en 1921, l'économiste **Frank Knight** fera la distinction entre le **risque**, au sens d'une incertitude liée à une probabilité mesurable et estimable, et l'incertitude (*uncertainty*) qu'on nomme depuis **l'ambiguïté** dans le jargon des économistes et qui n'est pas estimable (Knight, 1921). Dix ans après la proposition de Von Neumann et Morgenstern, en 1954, le mathématicien **Léonard Jimmie Savage** complètera leur proposition, en rajoutant une composante permettant de prendre en compte le cas d'incertitude soulevé par Frank Knight quand les probabilités ne sont pas connues. La composante en question est l'idée que les individus possèdent une **fonction interne pour se représenter les probabilités** qui est indépendante de l'utilité associée à un objet et qui est subjective, et qui peut donc radicalement dévier par rapport à la réalité. Il identifiera sept axiomes qui permettent, s'ils sont vérifiés, de dire si un agent économique se comporte bien comme s'il maximisait une utilité interne (Savage, 1954) sur la base de ces ajouts. Cette extension de la théorie sera nommée la **théorie de l'utilité espérée subjective**.

En parallèle à ces mathématiciens qui se concentrent sur les choix et les théories économiques, d'autres mathématiciens s'intéressent à la psychologie et aux théories du conditionnement (voir *Psychologie et physiologie*). En 1951, les mathématiciens américains Robert R. **Bush** et Frederik **Mosteller** vont essayer de développer un **modèle mathématique de l'apprentissage**, en particulier dans le cas du conditionnement opérant (Bush and Mosteller, 1951). Plus tard, d'autres mathématiciens américains, Allan R. **Wagner** et Robert A. **Rescorla** appliqueront le même formalisme au conditionnement classique (Rescorla and Wagner, 1972). Dans ces deux modèles, l'idée basique est qu'une action, dans le premier cas, ou un stimulus, dans le second cas, est associé à une valeur  $Q(t)$  à un essai  $t$  donné et que cette valeur est mise à jour en fonction d'une erreur de prédiction  $\Delta Q(t) = \Delta Q = r(t) - Q(t)$  où  $r(t)$  est la récompense (ou la punition) reçue à cet essai. La valeur mise à jour vaut donc  $Q(t + 1) = Q(t) + \alpha * \Delta Q(t)$  où  $\alpha$  est le taux d'apprentissage, spécifique à un individu.

### I.1.f L'économie comportementale

Les sciences économiques, grâce notamment à l'aide des mathématiciens, avaient fait de grands progrès dans l'étude des comportements humains comme nous l'avons vu



## Introduction

précédemment (voir *Mathématiques*). Dans les années 1970, les psychologues **Daniel Kahneman** et **Amos Tversky** vont provoquer une nouvelle révolution dans l'étude des comportements humains. Daniel Kahneman dira plus tard qu'un tournant pour lui a été de lire un extrait de l'économiste suisse Bruno Frey expliquant que l'être économique est un individu rationnel, égoïste et que ses goûts ne changent pas. En tant que psychologue en contact direct avec la réalité des comportements humains, cette vision très théorique de l'homme lui sembla tellement éloignée de la réalité qu'il décida progressivement, avec son collègue Amos Tversky, de mener une série d'expériences pour montrer dans quelle mesure les propositions établies en économie jusque-là n'étaient pas respectées et d'essayer d'en rendre compte (Kahneman, 2012). Ils aboutiront ainsi à une nouvelle théorie : **la théorie des perspectives**. L'article fondateur de cette théorie date de 1979 (Kahneman and Tversky, 1979). Il reprend les théories économiques connues en y ajoutant des principes expliquant pourquoi les humains ne se comportent pas comme les théories le prédisaient jusqu'alors. Kahneman et Tversky ont, en choisissant de publier dans un journal d'économie plutôt que de psychologie, réussi à créer un pont entre les deux disciplines et à créer ce nouveau champ d'étude qu'est l'économie comportementale, inspirée de ces deux domaines.

Un de leurs premiers ajouts est l'idée d'un **point de référence subjectif**. Pour expliquer ce point, Daniel Kahneman donne l'exemple de deux individus, Jack et Jill, qui ont tous les deux aujourd'hui 5 millions de dollars. Si l'utilité était purement équivalente à la richesse, ils devraient être également heureux tous deux. Seulement, hier, Jack avait 1 million de dollars alors que Jill en avait 9 millions. Kahneman nous explique que, dans ce cas, on s'attend à ce que Jack soit plus heureux que Jill. En effet si Jack a vu ses gains augmenter de 4 millions de dollars, Jill en revanche a perdu 4 millions de dollars. En résumé, leur perception de la somme qu'ils possèdent ne dépend pas seulement du montant possédé à un moment  $t$ , mais du point de référence par rapport auquel cette somme est comparée.

Un autre point qu'ils soulèvent dans la théorie des perspectives est l'idée d'une **sensibilité décroissante au changement** par rapport au point de référence. Par exemple, si vous êtes chercheur en neurosciences, votre point de référence est probablement tel que vous êtes sensible à la différence entre 100€ et 500€, mais par contre vous aurez potentiellement beaucoup plus de mal à distinguer 1.000.000€ de 10.000.000€. A l'inverse si vous êtes Bernard Arnault, première fortune mondiale en 2019, vous serez probablement bien plus sensible à la différence entre les deux derniers montants. Ceci s'illustre sur la courbe de transformation des

## Introduction

montants objectifs en valeur subjective par une saturation quand on s'éloigne du point de référence subjectif (voir **Figure 2**).

Un autre point essentiel soulevé par Kahneman et Tversky est la différence comportementale dans la considération des probabilités entre une situation liée à des gains et une situation liée à des pertes. Pour illustrer ce point, ils proposent un cas où il faut choisir entre :

- A) Gagner 900\$ avec une probabilité de 100%
- B) Gagner 1000\$ avec une probabilité de 90%

La majorité des gens tend à choisir l'option A dans ce cas. Mais, si on remplace la perspective de gains par des pertes, c'est l'inverse qui se produit. Les participants préfèrent alors l'option B. Kahneman et Tversky constatent ainsi que dans le domaine des gains, les individus présentent une **aversion pour le risque** (*risk aversion*), c'est-à-dire une préférence pour les options avec une probabilité plus élevée. Cependant, dans le domaine des pertes, cette aversion pour le risque s'inverse en **recherche du risque** (*risk seeking*), c'est-à-dire une préférence pour les options avec une probabilité plus faible, dans la mesure où celles-ci sont associées à une perte moins grave.

Enfin, Kahneman et Tversky ont découvert un autre principe fondamental de la psychologie humaine. Ils proposent à des participants de jouer à pile ou face. Si la pièce tombe sur pile, les participants gagnent 150€, si elle tombe sur face les participants perdent 100€. Ils constateront avec plusieurs choix de ce type que l'équivalent, dans le domaine des pertes, de la perspective de gagner une certaine somme, comme 150€, n'est pas la perspective de perdre une somme équivalente dans le domaine des pertes mais un montant 1.5 à 2.5 fois plus élevé en valeur objective. Ce principe s'illustre par la différence de la pente transformant la valeur objective en valeur subjective entre le domaine des gains et celui des pertes (voir **Figure 2**). En effet, on constate alors que la pente de la courbe est plus raide dans le cas des pertes. Ceci implique qu'un montant équivalent dans le domaine des gains et dans le domaine des pertes au niveau objectif n'aura pas la même valeur psychologique. C'est ce que Kahneman et Tversky appelleront **l'aversion à la perte**.



**Figure 2: Illustration des grands principes de la théorie des perspectives.** Au centre est le point de référence subjectif. La pente de la courbe est plus importante dans le cas des pertes que dans le cas des gains ce qui illustre l'aversion à la perte. Enfin, la discrimination entre deux points proches du point de référence est plus importante qu'entre deux points éloignés du point de référence à cause de la forme en S de la courbe. Extrait de (Kahneman, 2012)

### I.1.g La neuroéconomie

A la suite de deux congrès marquants en 1997 et en 2001, visant à regrouper des économistes, des neuroscientifiques et des psychologues s'intéressant aux sciences cognitives, la neuroéconomie est apparue (Glimcher and Fehr, 2014). En quelque sorte, là où l'économie comportementale marquait la liaison entre les psychologues et les économistes, la neuroéconomie marque la liaison entre ces deux disciplines et les neurosciences issues de la biologie. La problématique fondamentale de la neuroéconomie est probablement de mieux comprendre les facteurs qui expliquent les choix effectués par les individus, autant sur le plan comportemental que sur le plan biologique. Nous étudierons l'aspect neural plus loin, mais un des apports fondamentaux de la neuroéconomie est de permettre l'observation de l'activité

## Introduction

cérébrale et donc de pouvoir voir s'il existe dans le cerveau des structures qui vérifient ou non les théories économiques de l'utilité.

### Résumé

En somme, nous avons vu que plusieurs disciplines se sont intéressées à la valeur subjective, parfois en complément les unes des autres, parfois complètement en parallèle. La plupart de ces disciplines s'accordent cependant sur quelques points :

- **Tous les individus cherchent à maximiser une variable.** Les philosophes l'appellent le **bonheur** et les économistes **l'utilité**. Cependant, ces deux notions sont très proches puisque l'utilité associée à une option l'est dans la mesure où elle permet d'augmenter le bonheur global de l'individu ;
- Cette variable est, au moins en partie, une fonction des désirs, des appétits dirait Spinoza, et de leur satisfaction ;
- Pour maximiser son bonheur, on peut donc soit limiter les buts à atteindre, c'est ce que prônent la plupart des philosophes, soit, comme Sade ou les Cyrénaïques, chercher à les satisfaire tous ;
- En règle générale, la recherche de satisfaction de ces désirs ainsi que la recherche d'évitement des peines semblent être deux des moteurs principaux du comportement et des choix des individus ;
- Le comportement des individus diffère selon qu'ils agissent pour atteindre un des objets de leur désir ou pour essayer d'éviter de subir quelque chose d'avérsif. Ils donnent, notamment, un point plus important aux événements avérsifs. Ils sont aussi prêts à prendre plus de risques s'ils agissent en vue d'éviter un événement avérsif par rapport à s'ils agissent en vue de garantir l'obtention d'un objet de leurs désirs ;
- L'utilité dépend d'un point de référence subjectif ;
- Le lien entre un appétit, un besoin, et le moyen de le satisfaire est lié à un apprentissage tout au long de la vie par association entre des événements temporellement reliés.

## I.2 Bases neurales de la valeur

### I.2.a Histoire de la découverte des centres cérébraux du plaisir

*« ce n'est pas l'odeur des lièvres qui fait le plaisir des chiens ; ce qu'ils aiment, au contraire, c'est de les manger ; mais l'odeur leur fait ordinairement percevoir la proie à croquer. Et ce qu'aime le lion, ce n'est pas non plus la voix du bœuf ; son plaisir, au contraire, c'est de le*

## Introduction

*dévoré ; mais il perçoit d'habitude la proximité de celui-ci par son beuglement et donc il paraît aimer cette voix. De la même façon, il n'a pas non plus, d'ailleurs, de plaisir à voir « soit une biche, soit une chèvre sauvage », mais bien à découvrir qu'il aura de quoi manger. »*

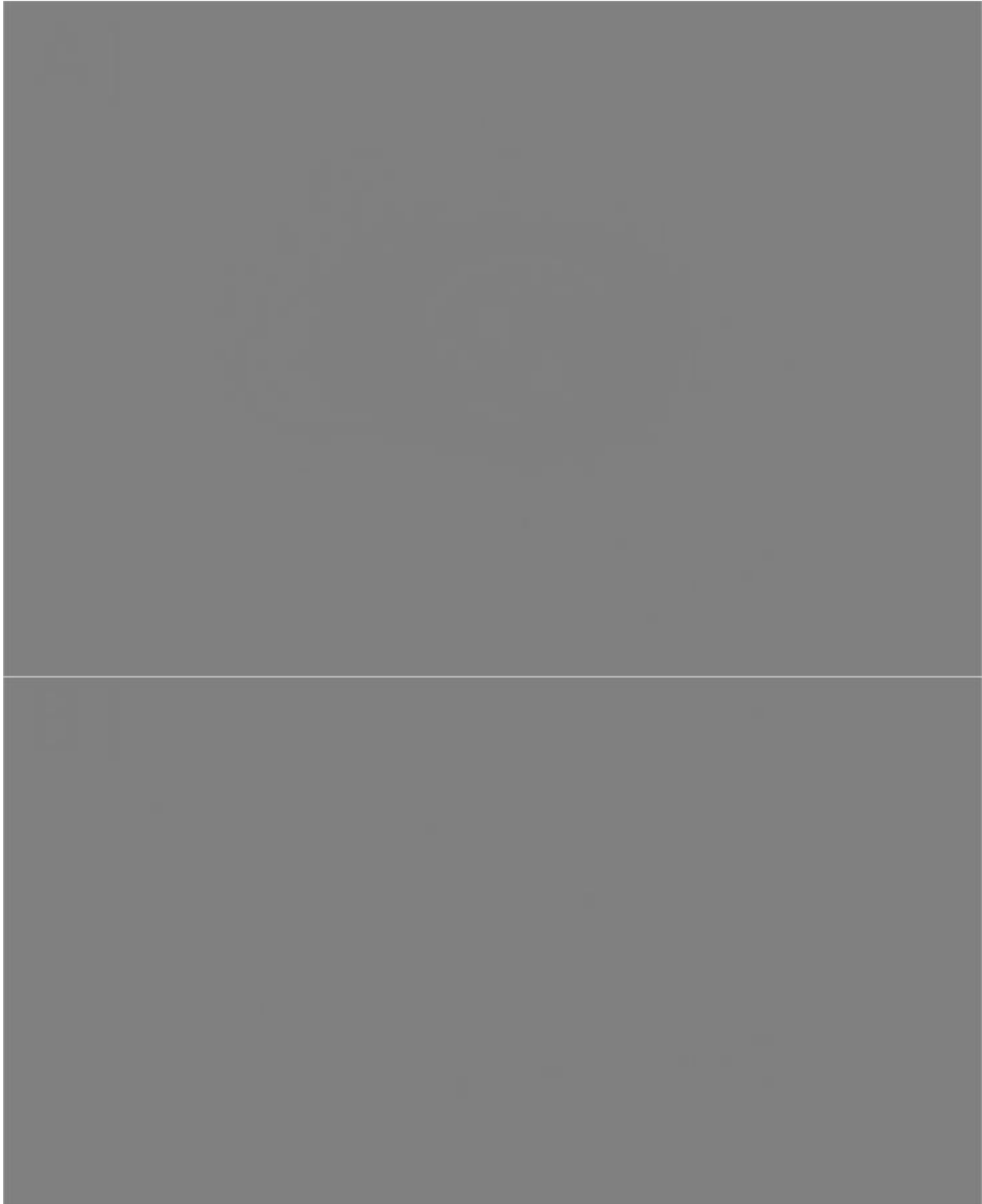
3.1.2.1, Livre III, *Éthique à Nicomaque*, Aristote (Aristote, 2004)

Dès 1954, un psychologue américain et un neuroscientifique canadien, James **Olds** et Peter **Milner**, vont faire franchir un pas à la science de la valeur et du plaisir. Là où leurs prédécesseurs avaient su identifier des lois générales gouvernant l'apprentissage pour associer une action à une récompense externe (voir *Psychologie et physiologie*), ils établiront une expérience très élégante pour voir s'il était possible d'identifier les centres cérébraux du plaisir. Ils mettent ainsi en place une expérience où des rats implantés avec des électrodes intracrâniennes peuvent délibérément choisir de stimuler électriquement les électrodes en question au moyen d'un levier. Cette méthode s'appelle **l'autostimulation**. Ils découvriront ainsi pour la première fois que stimuler certaines zones du cerveau peut être source de plaisir pour un individu puisque les rats en question vont activement chercher à appuyer sur le levier de manière répétée comme pour une récompense primaire (Olds and Milner, 1954). De plus, non seulement le siège du plaisir est donc bien dans le cerveau, mais il semble circonscrit à certaines zones, puisque les rats implantés par Olds et Milner ne sont pas sensibles de la même manière selon la zone que le levier va stimuler : ils peuvent chercher à s'autostimuler plus d'une centaine de fois par heure quand les électrodes se trouvent dans certaines zones, alors qu'ils ne vont quasiment pas le faire pour d'autres. Les décennies qui suivront permettront non seulement de mieux identifier les zones pouvant générer une sensation de plaisir quand elles sont stimulées, mais, en plus, d'identifier les bases chimiques des neurotransmetteurs sécrétés dans les synapses qui provoquent ces sensations.

Si le fait que le plaisir, comme sensation subjective, soit circonscrit à une seule zone du cerveau fait encore débat (Ikemoto, 2010), il est néanmoins plus que démontré que deux populations neuronales y sont systématiquement liées : **les neurones dopaminergiques de l'aire tegmentale ventrale et le striatum ventral** chez les Mammifères, en particulier **le noyau accumbens**. Les neurones dopaminergiques de la substance noire et de l'aire tegmentale ventrale, tout comme les neurones du striatum, existent chez tous les Vertébrés (Cisek, 2019; Grillner and Robertson, 2016). Les neurones dopaminergiques ont plusieurs cibles sur lesquelles ils vont projeter leurs axones. Le striatum est l'une de leurs cibles principales (voir **Figure 3**). La conservation très forte des propriétés fondamentales de ces structures à travers plusieurs espèces a permis de nettement faire avancer nos connaissances sur le sujet via l'étude

## Introduction

des modèles animaux depuis les années 1950. En effet, différentes expériences vont montrer que les rats apprennent très vite à s'auto-injecter des amphétamines, de la cocaïne ou de la dopamine directement dans le striatum ventral (Carlezon, and Wise, 1996; Ikemoto, 2010; Ikemoto et al., 1997). En 1978, le psychologue Anthony **Phillips** et le psychiatre Hans **Fibiger** démontreront que l'autostimulation de l'aire tegmentale ventrale par des rongeurs est drastiquement réduite si les neurones qui projettent depuis cette zone sont lésés. Ce résultat va de pair avec une réduction drastique des quantités de dopamine dans le striatum (Phillips and Fibiger, 1978). De plus, dans ce même article, ils montrent que l'autostimulation du striatum est peu affectée par une lésion de l'aire tegmentale ventrale. Ceci laisse penser que **le plaisir ressenti dépend de l'activité électrique des neurones du striatum**. En tout cas, ce qui apparaît clairement c'est qu'un animal va activement chercher à stimuler les neurones de cette zone. L'activité des neurones du striatum est normalement élicitée par la sécrétion de **dopamine** des neurones dopaminergiques de l'aire tegmentale ventrale. Ces résultats seront partiellement confirmés chez l'homme par le psychiatre américain Robert Heath. Dès les années 1950, Robert Heath implante ses patients atteints de troubles psychiatriques avec des électrodes profondes pour stimuler différentes zones du cerveau. Il espère ainsi guérir ses patients des troubles dont ils sont atteints. En 1972, en stimulant des patients près du striatum ventral, il réalise que certains patients vont chercher à autostimuler répétitivement et intensément cette zone. En parallèle, la stimulation provoque chez certains d'entre eux un fort désir sexuel (Berridge and Kringelbach, 2015; Heath, 1972).



**Figure 3 : Zones effectrices des neurones dopaminergiques.** A] Schéma simplifié des zones innervées par des neurones dopaminergiques dans le cerveau humain. Figure issue de (Glimcher, 2014). B] Schéma détaillé du circuit reliant l'aire tegmentale ventrale au noyau accumbens dans un cerveau de rongeur. La signification des sigles est donnée ci-après : VTA : aire tegmentale ventrale, LH : hypothalamus latéral, Amy : amygdale, NAc : noyau accumbens, mPFC : cortex préfrontal médian, Hipp : hippocampe, LHb : habenula latérale, LDTg : tegmentum dorsal latéral, RMTg : noyau tegmental rostromédian. Les codes couleurs indiquent la nature des projections : glutamatergiques en rouge, dopaminergiques en vert et gabaergiques en bleu. Figure issue de (Russo and Nestler, 2013).

## Introduction

Plus tard, Kent **Berridge** insistera sur la distinction entre la sensation de **plaisir** (*liking*) et l'incitation à agir en anticipation d'un plaisir, c'est-à-dire le **désir** (*wanting*) (Berridge and Valenstein, 1991). Il est d'ailleurs intéressant de constater que les mêmes concepts vont être découverts indépendamment en économie. Daniel Kahneman distinguera ainsi d'un côté **l'utilité liée à l'expérience** (*experienced utility*) qui correspond au plaisir ressenti et de l'autre **l'utilité liée à la prise de décision** (*decision utility*) qui correspond au désir (Kahneman, 2012; Kahneman and Sugden, 2005). Kent Berridge identifiera ainsi différentes sous-parties du noyau accumbens en fonction de si elles corrèlent avec le plaisir ou l'incitation à agir pour un plaisir (Berridge et al., 2009). Plus que la dopamine issue de l'aire tegmentale ventrale, Kent Berridge propose, sur la base d'expériences chez le rat principalement, que c'est les **opioïdes** et les **endocannabinoïdes** qui jouent un rôle important dans la sensation de plaisir (Smith et al., 2010b). De plus, il est à noter que l'activité du **pallidum ventral**, stimulée par des opioïdes, et celle du **noyau parabrachial** semblent elles aussi éliciter une sensation plaisante (voir **Figure 4**). Ceci irait dans le sens de l'idée que la sensation de plaisir pourrait être répartie à travers plusieurs zones dans le cerveau et non localisée à un seul endroit (Ikemoto, 2010). Kent Berridge propose de remettre en question le rôle qui était attribué à la dopamine jusque-là. La dopamine jouerait plus un rôle dans la stimulation du désir (*wanting*), alors que la sensation de plaisir elle-même (*liking*) reviendrait aux opioïdes et aux endocannabinoïdes, principalement. De ce fait, Kent Berridge remet aussi en question le rôle de l'aire tegmentale ventrale. L'aire tegmentale ventrale se transforme ainsi d'un centre de la sensation de plaisir en centre motivationnel qui pousse à l'action. Reprenant la littérature sur les stimulations profondes chez l'homme, Berridge remarquera d'ailleurs que là non plus aucun ressenti de plaisir n'est signalé en lien avec l'activité des centres dopaminergiques, mais bien plutôt une augmentation de la volonté de s'engager dans certaines activités, c'est-à-dire d'agir en vue de ressentir du plaisir (Berridge and Kringelbach, 2015; Schlaepfer et al., 2008). Cela n'explique cependant pas pourquoi les animaux comme les hommes rechercheraient à stimuler leur désir de recherche de plaisir en activant ces zones cérébrales si elles ne procurent aucun plaisir réel en tant que tel.





**Figure 4 : Centres principaux liés à la sensation de plaisir dans un cerveau de rongeur.** Les projections de l'aire tegmentale ventrale (VTA) sont en pointillés. La sensation de plaisir serait liée à des boucles entre différentes aires signalées par les flèches orange et rouge. Figure issue de (Smith et al., 2010b).

### I.2.b Circuits dopaminergiques et erreurs de prédiction

On a vu précédemment que la psychologie, la physiologie et les mathématiques ont permis d'établir des modèles mathématiques de l'apprentissage dans les années 1950-70. En parallèle, plusieurs études avaient permis de commencer à identifier des neurones dont l'activité était associée au plaisir notamment dans le striatum (voir *Histoire de la découverte des centres cérébraux du plaisir*). En 1997, Wolfram **Schultz**, Peter **Dayan** et Read **Montague** vont mener des expériences sur des primates non-humains. Ils enregistrent l'activité des neurones dopaminergiques de l'aire ventrale tegmentale et de la substance noire de singes en train d'effectuer une tâche d'apprentissage. Ils vont découvrir que ces neurones réagissent à l'**erreur de prédiction** telle que modélisée selon les modèles développés par Rescorla et Wagner (voir **Figure 5** et (Schultz et al., 1997)). C'est une évolution conséquente dans l'étude des bases physiologiques du plaisir puisque c'est un premier pas vers une liaison entre les théories du comportement humain et les découvertes de la biologie sur les centres cérébraux du plaisir. En effet, si l'on part du principe que ce que les individus optimisent, c'est l'activité dans leurs neurones dopaminergiques ou dans leur striatum par exemple, alors on peut essayer de vérifier les théories de l'utilité sur une mesure directement observable, à condition de trouver un moyen de mesurer l'activité dans le striatum. Bien évidemment, comme évoqué précédemment, l'idée qu'une seule zone du cerveau définirait la sensation globale de plaisir est contestée (Ikemoto,

## Introduction

2010; Smith et al., 2010b), mais il faut souligner premièrement que ceux qui défendent l'idée que la sensation de plaisir est distribuée en réseau ne soutiennent pas que les différents nœuds du réseau sont indépendants. Au contraire, ils proposent premièrement que ces différentes zones se renforcent les unes les autres via des boucles, comme l'activité du pallidum ventral et l'activité du striatum ventral qui se stimulent l'une l'autre via une boucle (Smith et al., 2010b). Deuxièmement, même si cela ne reflétait pas à 100% la sensation subjective, cela pourrait déjà fournir un proxy intéressant pour pouvoir tester certaines hypothèses formulées par la psychologie et l'économie.



**Figure 5 : L'erreur de prédiction dans les neurones dopaminergiques.** Figure du haut : **Erreur de prédiction positive** : les neurones dopaminergiques s'activent quand une récompense est délivrée au singe sans qu'il s'y attende. Figure du milieu : **Récompense prévue** : les neurones dopaminergiques réagissent quand un stimulus conditionné permet d'anticiper qu'une récompense va avoir lieu mais leur activité ne varie pas quand une récompense attendue est donnée au singe. Figure du bas : **Erreur de prédiction négative** : l'activité des neurones dopaminergiques baissent drastiquement quand une récompense était attendue mais n'a pas eu lieu. Figure extraite de (Schultz et al., 1997).

L'imagerie par résonance magnétique (IRM) a été inventée dans les années 1970 par Paul Lauterbur et Peter Mansfield. Elle est un grand pas en avant dans plusieurs domaines, mais en particulier au niveau médical, puisqu'elle permet d'observer directement l'intérieur du corps humain, tout comme les rayons X le permettait auparavant, mais sans nocivité pour le corps puisqu'elle est basée sur les ondes magnétiques. Cette technique non-invasive permet

## Introduction

d'observer directement l'anatomie du cerveau. En 1992, trois groupes montrent qu'en adaptant les paramètres d'acquisition de l'IRM, il est possible d'observer les variations de la concentration en hémoglobine oxygénée du sang (Bandettini et al., 1992; Kwong et al., 1992; Ogawa et al., 1992). Ces trois groupes marqueront le début de l'**IRM fonctionnelle (IRMf)** via la mesure du signal BOLD (*Blood Oxygen Level Dependent*) (Logothetis and Wandell, 2004). En effet, des expériences menées dans le cortex visuel du singe le confirmeront (Logothetis et al., 2001), l'activité accrue d'une population neuronale, en termes de potentiels d'action, induit un apport sanguin accru localement (**Figure 6**) qui a une dynamique assez typique : la réponse hémodynamique enregistrée.



**Figure 6 : Relation entre différentes mesures de l'activité électrique des neurones du cortex visuel d'un singe et le signal BOLD dans le cortex visuel.** Le signal BOLD a un temps de latence et augmente à la suite de l'activité neuronale alors même que l'activité neuronale elle-même a déjà stoppé. Le sens des abréviations est donné ci-après : LFP : *local field potentials* = potentiels de champs locaux, MUA : *multi-unit activity* = enregistrements multi-unitaires, SDF : *spike-density function* = fonction de densité des potentiels d'action. Figure adaptée de (Logothetis et al., 2001)

Quelques années à peine après l'émergence de ce nouvel outil qu'est l'IRMf, Hans Breiter *et al.* cherchent déjà à l'utiliser pour localiser les substrats neuronaux qui réagissent à une récompense monétaire dans le cerveau humain. Dès 1996, chez seulement quatre sujets humains, ils identifient déjà le **striatum** comme une zone s'activant en réponse à un son associé à l'obtention imminente d'une récompense monétaire (Breiter et al., 1996). Son équipe sera bientôt rejointe par l'équipe de Brian Knutson dans cette course pour voir si les résultats du trio Schultz, Dayan et Montague chez le primate non-humain se répliquent chez l'homme. Les deux équipes publieront finalement leurs résultats à peu près simultanément en avril et en mai 2001. Dans les deux cas, et malgré le faible nombre de sujets de chaque expérience en comparaison aux standards actuels (8 sujets chez Knutson *et al.* et 12 sujets chez Breiter *et al.*), le striatum s'active de manière proportionnelle au gain anticipé (Breiter et al., 2001; Knutson et al., 2001a).

## Introduction

Du fait de leur localisation et des contraintes techniques de l'IRM, il est difficile d'observer directement l'activité des neurones dopaminergiques identifiés par Schultz, Dayan et Montague. Cependant, étant donné l'importance de la connexion entre les neurones dopaminergiques et le striatum dans l'anticipation d'un plaisir chez l'animal (voir *Histoire de la découverte des centres cérébraux du plaisir*), les auteurs revendiquent de considérer l'activité du striatum comme un bon reflet de l'activité des neurones dopaminergiques. Il est intéressant au passage de noter que, dès cette époque-là, des ponts entre divers disciplines se forment puisque l'on retrouve dans ces deux publications des psychologues (Brian Knutson), des psychiatres (Hans Breiter) et même le fameux psychologue et économiste Daniel Kahneman (voir *L'économie comportementale*). Plus tard, plusieurs études viendront compléter ce résultat. Une méta-analyse basée sur 109 études d'IRMf chez l'homme de 2013 démontre ainsi que le striatum est bien une structure essentielle à l'encodage des erreurs de prédiction (Garrison et al., 2013). Sur le site Neurosynth, on peut de même voir que le striatum ventral est l'aire principale corrélant avec le terme « *prediction error* » au travers de 93 études en IRMf chez l'homme<sup>7</sup>. Une étude chez le rat a d'ailleurs directement confirmé qu'activer les neurones dopaminergiques de l'aire tegmentale ventrale induisait une activation, en IRMf, du striatum ventral (Ferenczi et al., 2016).

Le plus intéressant dans ces recherches est probablement les perspectives qu'elles ouvrent. En effet, si ce que les individus cherchent à maximiser est l'activité dans leur striatum ventral, alors mieux comprendre le fonctionnement et les dérégulations de cette zone et des zones qui y sont connectées pourra permettre de mieux comprendre le comportement des individus. En effet, plus récemment, les travaux de **Robb Rutledge**, **Eran Eldar** et **Yael Niv** montrent que **l'humeur** rapportée par des individus est une fonction de l'activité du striatum et de la quantité d'erreurs de prédiction positives ou négatives reçues dernièrement (Eldar et al., 2016; Rutledge et al., 2014). Ceci confirme aussi indirectement une hypothèse développée par Daniel Kahneman et Amos Tversky selon laquelle l'utilité est fonction non pas de la somme d'argent total possédé/obtenu mais d'un point de référence et de l'écart par rapport à ce point (voir *L'économie comportementale*). Cela permet aussi de faire directement le lien avec les problématiques développées déjà par les philosophes de l'Antiquité sur le rapport entre plaisirs et bonheur et sur les différentes manières d'atteindre une vie bienheureuse. Certains des philosophes que nous avons évoqués précédemment (voir *Philosophie*), comme les stoïciens,

---

<sup>7</sup> <https://neurosynth.org/analyses/terms/prediction%20error/>

## Introduction

proposaient en effet de réduire ses attentes autant que possible à ce qui peut être maîtrisé par l'individu. Adopter cette philosophie de vie revient ainsi naturellement à augmenter le taux d'erreurs de prédiction positives qui vont être expérimentées.

### I.2.c Le réseau des valeurs et ses propriétés

« *En définitive, les multiples marchandises ne sont toutes convertibles qu'en une seule valeur : l'argent : alors que l'argent, lui, est convertible dans toute leur diversité.* »

*Philosophie de l'argent*, Georg Simmel (Simmel, 2009)

Jusque-là les recherches sur le plaisir et les erreurs de prédiction positives se concentraient sur les neurones dopaminergiques et le striatum, autant dans les études sur les modèles animaux que dans les études chez l'homme. L'essor de l'IRMf va pourtant très vite permettre de mieux comprendre qu'un réseau plus vaste est impliqué dans l'évaluation et l'anticipation de récompenses. Dans la citation de Georg Simmel ci-dessus, il est rappelé que le rôle premier de l'argent est de servir de valeur d'échange entre différents biens. Dans le domaine de la prise de décision et de la neuroéconomie, il est intéressant de constater qu'une pensée très similaire a émergé à la fin des années 2000. Ainsi, va naître l'idée que certaines zones du cerveau, en particulier le cortex préfrontal ventromédian, agiraient de même dans le cerveau pour convertir tous les biens en une seule valeur : l'activité neuronale de cette zone. Une littérature très vaste en IRMf va en effet converger, au cours des deux dernières décennies, à montrer que le **cortex préfrontal ventromédian (vmPFC : ventromedial prefrontal cortex)** humain, en plus du striatum ventral, corrèle quasi-systématiquement avec la valeur subjective attribuée par les individus à un item et ce, indépendamment du type de récompense (récompense primaire comme la nourriture, ou récompense secondaire comme l'argent), du type de tâche (tâche d'évaluation ou tâche de choix par exemple) ou du moment considéré (lors de la prise de décision ou lors de l'obtention de la récompense) (Bartra et al., 2013; Boorman et al., 2009; Chib et al., 2009; Clithero and Rangel, 2014; Hare et al., 2009, 2011a; Lebreton et al., 2009, 2015; Levy and Glimcher, 2011, 2012; Lin et al., 2012; Litt et al., 2011; Liu et al., 2011; Motoki et al., 2019; O'Doherty, 2011; Philiastides et al., 2010; Sescousse et al., 2013; Shapiro and Grafton, 2020; Wu et al., 2011). L'idée que le cerveau posséderait donc une monnaie commune qui permettrait d'encoder tout type de récompense semble donc validée par les données d'imagerie. Les acteurs clés de ce réseau des valeurs, en IRMf tout du moins, sont principalement le **striatum ventral** et le **vmPFC (Figure 7)**, mais on constate aussi souvent

## Introduction

que le **cortex cingulaire postérieur (PCC : posterior cingulate cortex)** et parfois **l'hippocampe** y sont régulièrement associés (Lebreton et al., 2009).



**Figure 7 : Corrélatés neuraux de la valeur subjective.** La figure représente le résultat d'une méta-analyse et montre les zones s'activant en corrélation avec la valeur subjective lors de la période de décision et conjonction avec celles s'activant en corrélation avec la valeur subjective lors de la période où les participants obtiennent un feedback sur leur performance. Figure issue de (Bartra et al., 2013).

Depuis que ce réseau a été clairement identifié en neuroimagerie chez l'homme, plusieurs études ont permis de mieux comprendre son fonctionnement et les biais auxquels il pouvait être soumis à cause de ses propriétés intrinsèques.

La première propriété que nous avons déjà évoquée est qu'il est **générique**, c'est-à-dire qu'il encode la valeur subjective indépendamment de la nature de ce qui est considéré. En effet, il a déjà été montré à de multiples reprises que le vmPFC notamment corrélait avec la valeur subjective autant dans le domaine des récompenses primaires (alimentation, érotique, boisson) que secondaires (c'est-à-dire apprises par association aux récompenses primaires, comme l'argent par exemple) (Sescousse et al., 2013), dans le domaine des objets (Chib et al., 2009) que dans le domaine des actions (Gläscher et al., 2009) ou le domaine social (Lin et al., 2012).

Une deuxième propriété importante de ce réseau est son **automaticité**. En effet, il a été montré que, même quand l'évaluation n'a pas d'intérêt pour la tâche en cours, le réseau va automatiquement s'activer en fonction de la valeur attribuée à ce qui est observé (Grueschow et al., 2015; Lebreton et al., 2009, 2015; Levy et al., 2011; Lopez-Persem et al., 2020; Tusche et al., 2010). Plus particulièrement cela implique que même quand l'évaluation n'a pas d'intérêt immédiat, soit parce que les instructions de la tâche ne demandent pas de faire une évaluation, soit parce qu'aucun choix nécessitant d'être au courant de la valeur attribuée aux items à disposition ne doit être implémenté, l'évaluation a quand même lieu dans ce réseau. Cependant,

## Introduction

il n'est pas encore totalement clair de savoir quelles sous-parties du réseau sont dédiées à quelle tâche et lesquelles réagissent ou non automatiquement à la valeur. En effet, là où certains ont montré que la valeur subjective est encodée automatiquement dans l'ensemble du réseau et en particulier dans le vmPFC (Kim et al., 2007; Lebreton et al., 2009, 2015; Levy et al., 2011; Lopez-Persem et al., 2020; Tusche et al., 2010), l'équipe de Christian Ruff a montré que seul le cortex cingulaire postérieur réagit automatiquement à la valeur subjective quand elle n'est pas instrumentale pour effectuer un choix, probablement à cause d'une attraction attentionnelle, alors que le vmPFC ne corrèlerait avec la valeur subjective que lorsque cette valeur est instrumentale pour le choix (Grueschow et al., 2015). A ma connaissance, cette controverse n'a pour l'instant pas été résolue, mais cela suggère que les différentes parties du réseau des valeurs ne sont pas complètement redondantes et que les redondances qui y sont observées pourraient être plus liées au fait que les tâches effectuées ne permettent pas de discriminer leurs rôles respectifs, plus qu'elles aient des rôles similaires.

Un dernier effet très intéressant qui a été mis en évidence est que l'activité d'évaluation de ce réseau est sujette à des **biais d'état initial** et des **biais d'attribution** connus depuis longtemps des psychologues. En effet, il a été montré que le contexte dans lequel on se trouve pouvait avoir un impact sur l'évaluation d'un item donné dans ce contexte. En 1895, dans son fameux ouvrage *Psychologie des foules*, Gustave Le Bon rapporte l'anecdote d'un habile orateur qui cherche à convaincre son jury lors d'un procès. Cet orateur a du mal à convaincre un membre du jury en particulier qui semble garder un air renfrogné peu importe les efforts de l'orateur pour le convaincre lorsque, subitement, il s'adresse au président de la cour et dit « *Monsieur le Président, dit-il, ne pourriez-vous pas faire tirer le rideau là en face, Monsieur le septième juré est aveuglé par le soleil.* » ce qui lui permettra de mettre le juré en question dans sa poche (Le Bon, 2006). Plus récemment, des études suggèrent que les juges au tribunal rendraient des jugements biaisés selon s'ils viennent d'avoir une pause repas ou non (Danziger et al., 2011, 2015). En laboratoire, notre équipe a montré que le fait d'écouter une musique jugée agréable en même temps que l'on doit juger le plaisir procuré par l'observation d'un tableau biaisait la perception du plaisir esthétique associé au tableau en fonction du jugement porté sur la musique entendue (Abitbol et al., 2015). La valeur de la musique comme celle du tableau étaient corrélées à un cluster au sein du vmPFC suggérant que c'est l'activité de cette zone qui détermine le jugement de valeur donné par les participants et que biaiser son activité peut aussi biaiser la valeur attribuée à un évènement. Une étude encore plus récente de notre équipe avec des patients épileptiques implantés avec des électrodes intracrâniennes montre que

## Introduction

l'activité basale du vmPFC a en effet un impact sur l'activité du vmPFC en réponse à un stimulus donné (Lopez-Persem et al., 2020).

Il semblerait donc que le réseau des valeurs soit générique et que, peut-être de ce fait, il soit influençable si l'on peut dire puisqu'élever son activité via un stimulus donné peut le biaiser dans son estimation d'un autre stimulus. La question de son automaticité, c'est-à-dire qu'il encoderait la valeur subjective d'items non-pertinents pour ce que l'individu est en train de faire, reste à résoudre, bien que de nombreuses études pointent en ce sens.

### I.2.d L'anatomie du réseau des valeurs

Un point important à noter est que l'IRMf révèle des clusters d'activation qui, selon les paramètres utilisés dans le prétraitement (*preprocessing*), en particulier pour le lissage des images, et la qualité des données, correspondent, après transformation, plus ou moins aux structures anatomiques d'un modèle de référence, généralement le modèle MNI (*Montreal Neurological Institute*) basé sur 152 cerveaux scannés. Ces clusters d'activation ne correspondent pas toujours aux régions anatomiques définies par les neurophysiologistes qui identifient des régions sur la base du type de neurones ou de leur connectivité. Dans le cas de ce réseau des valeurs que nous venons d'évoquer, il est donc particulièrement intéressant de constater que les clusters fonctionnels observés en lien avec la valeur subjective correspondent pour ce qui est du striatum ventral, mais aussi pour le cortex postérieur cingulaire et l'hippocampe, à des régions anatomiques prédéfinies. Le cluster identifié dans le striatum ventral (**Figure 7**) correspond en effet parfaitement à la région du **noyau accumbens** (**Figure 8A**) telle qu'identifiée par les anatomistes, ce qui permet de faire un lien direct avec la littérature sur les modèles animaux où c'est lui qui avait en effet été identifié comme le siège des réactions de plaisir (voir *Histoire de la découverte des centres cérébraux du plaisir*) et des erreurs de prédiction (voir *Circuits dopaminergiques et erreurs de prédiction*).

Pour ce qui est du cluster dans le **cortex préfrontal ventromédian (vmPFC)**, les choses semblent un peu moins claires car, comme on peut le voir en comparant la **Figure 7** et la **Figure 8B**, le cluster fonctionnel semble se superposer à plusieurs régions anatomiques. Le lissage (*smoothing*) et les autres transformations effectuées lors du traitement des images en IRMf peuvent créer des biais en élargissant, décalant ou en réduisant la taille de certains clusters fonctionnels. Il faut cependant aussi prendre en compte que, alors que le striatum est une structure relativement homogène à travers les individus, probablement en partie du fait de son ancienneté dans l'arbre évolutif (Cisek, 2019; Grillner and Robertson, 2016), le cortex



## Introduction

préfrontal, bien plus récent dans notre évolution, présente de fortes hétérogénéités à travers les individus (Chiavaras and Petrides, 2000; Chiavaras et al., 2001). Quelques études pointent d'ailleurs vers le fait que ces hétérogénéités créent des décalages dans la localisation des clusters d'activation au sein du vmPFC (Li et al., 2015; Lopez-Persem et al., 2018). Dans tous les cas, s'il fallait identifier une seule région qui correspond au vmPFC fonctionnel, il semblerait que **l'aire 14m (Figure 8B)** soit la meilleure candidate pour ce rôle, mais d'autres études seraient nécessaires pour valider ce résultat. A l'heure actuelle, les chercheurs de la prise de décision tendent à ignorer provisoirement les bases anatomiques de cette région, mais cela va probablement évoluer dans les années à venir.



**Figure 8 : Correspondances anatomiques des structures fonctionnelles du réseau de la valeur.** A] Identification probabiliste des différentes sous-régions du striatum basée sur des scans IRM. En jaune figure le noyau caudé, en rouge le putamen, en vert clair le **noyau accumbens** et en violet le globus pallidus externus d'après (Pauli et al., 2017). B] Labels anatomiques de Brodmann correspondants aux différentes sous-parties du cortex préfrontal ventromédian. Les zones en question ont été identifiées sur des cerveaux d'individus décédés sur la base de critères histologiques. A gauche, figure issue de (Mackey and Petrides, 2010) représentant les limites théoriques des différentes zones après parcellisation et à droite figure issue de (Mackey and Petrides, 2014) montrant le nombre d'hémisphères d'individus étudiés post-mortem où la zone a été identifiée pour un voxel donné.

De plus, une controverse qui n'est toujours pas complètement résolue existe sur le rôle respectif du cortex orbitofrontal médian, c'est-à-dire le vmPFC, et du **cortex orbitofrontal latéral**. Une ancienne théorie proposait que la partie médiane soit dévolue aux récompenses de

## Introduction

valence positive et la partie latérale aux punitions (Kringelbach, 2004, 2005). Si cette théorie a depuis été abandonnée, la controverse existe toujours entre les chercheurs sur les modèles primates non-humains et les études en neuroimagerie chez l'homme. Les chercheurs sur les primates non-humains trouvent souvent qu'une proportion conséquente des neurones du cortex orbitofrontal latéral est corrélée avec l'estimation des valeurs subjectives, alors que, dans la partie médiane du cortex orbitofrontal, la proportion de neurones corrélant avec la valeur subjective est très faible voire nulle en comparaison (Padoa-Schioppa and Conen, 2017; Wallis, 2011). A l'inverse, la plupart des études en neuroimagerie chez l'homme convergent vers un rôle prépondérant du vmPFC en lien avec la valeur subjective (Bartra et al., 2013; Levy and Glimcher, 2012), c'est-à-dire la partie médiane du cortex orbitofrontal justement. Si certains ignorent ces distinctions et citent les études montrant des effets dans le cortex orbitofrontal latéral et celles montrant un effet dans le vmPFC indifféremment comme s'il s'agissait de la même unité fonctionnelle, il est difficile d'imaginer que ces deux portions conséquentes du cortex préfrontal aient des rôles totalement équivalents étant donné leur connectivité très différente avec le reste du cerveau (Briggs et al., 2018a, 2018b; Mars et al., 2016; Zald et al., 2014) et leur composition différente en termes de populations neuronales (Bludau et al., 2014; Mackey and Petrides, 2009). Une explication de ces résultats divergents pourrait reposer sur des différences entre espèces. Cependant, l'organisation du cortex orbitofrontal paraît relativement homologue entre les primates non-humains et l'homme (Wallis, 2011). La différence pourrait peut-être alors s'expliquer bêtement par des **différences de méthode et d'échelle d'étude** ? En effet, la plupart des études chez l'homme utilisent l'IRMf, alors que les études chez le primate non-humain utilisent des enregistrements électrophysiologiques au niveau d'un seul neurone (enregistrements unitaires) ou d'un groupe de neurones (mesure des potentiels de champs locaux). Les contraintes techniques de la plupart des méthodes d'acquisition des images en IRMf font qu'une partie du cortex orbitofrontal n'ait pas toujours acquises lors des scans, à moins d'adapter le protocole d'acquisition pour optimiser cela (Deichmann et al., 2003; Domsch et al., 2013; Weiskopf et al., 2007). De plus, au moins une étude sur des primates non-humains en IRMf a pu identifier le vmPFC comme étant corrélé à la valeur subjective (Papageorgiou et al., 2017), bien que le signe de la corrélation était inversé par rapport aux études classiques chez l'homme. Le vmPFC encodait la valeur non-choisie moins la valeur choisie alors qu'on trouve généralement la corrélation inverse chez l'homme en IRMf (voir *Bases neurales du choix*). En électrophysiologie chez le primate non-humain, une corrélation entre l'activité d'une population de neurones du vmPFC, mais pas du cortex

## Introduction

orbitofrontal latéral, et l'aspect appétitif d'une récompense a aussi déjà été observé (Abitbol et al., 2015). Réciproquement, une étude récente a montré que l'activité du cortex orbitofrontal latéral corrélait tout autant que le vmPFC avec la valeur subjective chez des patients épileptiques implantés avec des électrodes intracrâniennes (Lopez-Persem et al., 2020). Enfin, dans le cortex orbitofrontal latéral des primates non-humains, environ la moitié des neurones émet plus de potentiels d'action lorsque la valeur subjective augmente, alors que l'autre moitié environ montre une activité décroissante à mesure que la valeur subjective augmente, ce qui pourrait expliquer qu'en IRMf, ces deux effets s'annulent lorsque l'on regarde l'activité globale de la zone en signal BOLD (Wallis, 2011). Dans le même ordre d'idée, tout **encodage non-linéaire** de la valeur dans le cortex orbitofrontal latéral empêcherait d'y observer un lien avec la valeur subjective dans les études classiques d'IRMf qui recherchent des zones dont le signal BOLD corrèle linéairement avec la valeur subjective. Il se pourrait donc aussi que, bien que les deux zones encodent la valeur subjective, l'encodage soit de nature différente entre ces deux zones du cortex orbitofrontal. Toutes ces données suggèrent que des biais liés aux différentes techniques utilisés dans les mesures de l'activité du cortex orbitofrontal seraient les principaux responsables des différences entre les études qui trouvent plutôt le cortex orbitofrontal médian ou plutôt le latéral en lien avec la valeur subjective. Cependant, les mécanismes précis et le rôle de chacune de ces portions du cortex orbitofrontal restent à préciser, notamment du fait de leur connectivité très différente au reste du cerveau suggérant qu'elles ne sont pas non plus fonctionnellement complètement équivalentes.

En plus de cette distinction entre le cortex orbitofrontal latéral et le cortex orbitofrontal médian, il a été montré qu'une distinction **antéropostérieure** pourrait aussi avoir lieu au sein du cortex orbitofrontal en fonction du type de récompense considérée. Ces travaux pointent vers le fait que la partie **postérieure** du cortex orbitofrontal qui serait plus ancienne évolutivement parlant serait plus fortement activée par des **récompenses primaires** (sucre, eau, etc.) alors que la partie **antérieure**, plus récente dans l'évolution, s'activerait plus en réponse à des **récompenses secondaires**, c'est-à-dire des récompenses apprises au cours de la vie par association avec des récompenses primaires, comme c'est le cas pour l'argent (Kringelbach, 2004, 2005; Li et al., 2015; Sescousse et al., 2013). Il est à noter que ces différences fonctionnelles semblent aller de pair avec des différences dans la nature des populations neuronales réparties sur l'axe antéro-postérieur avec notamment des différences de densités dans la couche IV du cortex (Mackey and Petrides, 2010). Cette distinction ne nie cependant pas que ce soit le vmPFC qui encode la valeur subjective mais affirme plutôt qu'il se décompose

## Introduction

en un continuum avec des neurones spécialisés dans les récompenses primaires au pôle postérieur et des neurones s'associant de manière de plus en plus flexible à différents types de stimuli indiquant une récompense au pôle antérieur.

En résumé, si l'anatomie du striatum ventral et son rôle dans l'estimation d'une valeur subjective est de mieux en mieux caractérisé, il reste encore beaucoup d'études transversales à effectuer pour mieux comprendre le rôle du cortex orbitofrontal et réconcilier les différentes théories qui essaient d'expliquer son organisation fonctionnelle.

Un dernier point intéressant à évoquer est que le striatum ventral, qui correspond donc au noyau accumbens, est souvent identifié comme étant fortement connecté au vmPFC autant chez l'animal (**Figure 3, Figure 4**) que chez l'homme (Choi et al., 2012; Morris et al., 2016) anatomiquement et fonctionnellement. Cela implique donc que ces deux régions interagissent et ne sont donc pas totalement indépendantes l'une de l'autre, même si cela ne veut pas dire pour autant que leurs rôles respectifs soient complètement redondants.

Pour ce qui est des aires moins canoniques du réseau des valeurs que sont l'hippocampe et le PCC, leurs rôles précis dans le réseau sont moins clairs. **L'hippocampe** serait particulièrement recruté dans des cas requérant d'évaluer des situations futures imaginées (Lebreton et al., 2013) ou quand les choix à effectuer sont basées sur la mémoire (Gluth et al., 2015; Weilbacher and Gluth, 2016). De plus, l'hippocampe est connecté au striatum (Pennartz et al., 2004) et au vmPFC (Weilbacher and Gluth, 2016). Il pourrait, de ce fait, moduler leur activité et interagir avec eux. Le **PCC**, lui, fait partie, avec le vmPFC, du réseau dit du « mode par défaut » du cerveau. Il présente de ce fait une connectivité fonctionnelle forte avec le vmPFC (Buckner and DiNicola, 2019; Greicius et al., 2009; Raichle, 2015). Différentes hypothèses existent sur son rôle : il serait actif lorsqu'un individu est concentré sur ses pensées internes, comme c'est le cas dans le « mode par défaut ». Il pourrait aussi jouer un rôle dans l'attention (Leech and Sharp, 2014), en particulier dans le cas de stimuli à valeur haute, peut-être pour orienter l'attention vers eux (Grueschow et al., 2015). En résumé, l'hippocampe et le PCC sont en interaction directe avec le vmPFC et le striatum ventral et, selon les situations, leur activité pourrait aussi participer à la formation de la valeur subjective.

### I.2.e Impact des lésions du réseau des valeurs

Dès le XIX<sup>ème</sup> siècle, grâce à des médecins comme Paul Broca et Carl Wernicke qui vont étudier le cerveau de patients aphasiques, l'étude du cerveau franchit un cap puisque des zones précises du cerveau semblent impliquées dans des fonctions spécifiques. À leur époque,

## Introduction

ceci se fait principalement par l'étude post-mortem du cerveau de patients ayant eu des troubles comportementaux au cours de leur vie. Au XIX<sup>ème</sup> siècle toujours, Phineas Gage, après avoir reçu une barre de fer dans le cerveau suite à l'explosion d'un bâton de dynamite, voit son comportement complètement modifié du jour au lendemain : alors qu'il recouvre très vite la plupart de ses facultés, son comportement diffère tellement de sa personnalité d'avant l'accident que ses proches ne le reconnaissent plus. Il est plus grossier, a du mal à se conformer aux normes sociales et éprouve des difficultés à prendre des décisions à long terme et à s'y tenir. Un siècle plus tard, Hanna Damasio, sur la base de photos prises sur le crâne de **Phineas Gage**, conservé au musée médical Warren, va reconstituer en 3D la trajectoire hypothétique de la barre de fer dans le cerveau de Phineas Gage. Elle identifie le cortex préfrontal comme l'une des zones principales ayant été impactée par l'incident. Son mari, le neurologue Antonio Damasio, identifie un de ses patients, Elliot, qui, depuis une résection de tumeur cérébrale, présente lui aussi des différences de comportement similaires à celles décrites par le médecin de Phineas Gage, le docteur Harlow (Damasio, 2010). Grâce aux techniques d'imagerie cérébrale modernes évoquées précédemment, il peut étudier le cerveau d'Elliot et voir que le cortex préfrontal, et en particulier sa partie ventromédiane, semblent être fortement lésées chez ce patient. Il va développer par la suite la tâche du « jeu de casino » (*Iowa Gambling Task*) afin de mieux caractériser les patients atteints de lésions au cortex préfrontal. Le principe de la tâche est plutôt simple : il faut choisir plusieurs fois à la suite entre deux paquets de cartes duquel on veut piocher. L'un des deux paquets permet d'obtenir des récompenses monétaires plus élevées quand c'est le cas, mais il est en moyenne moins avantageux que l'autre paquet car on y obtient plus souvent des pertes que des gains. Les sujets sains ont tendance à commencer par choisir le « mauvais » paquet, qui procure des gains plus importants, puis à se tourner vers l'autre paquet une fois qu'ils ont compris la structure du jeu. Antonio Damasio et son équipe démontrent que les patients avec des lésions dans le cortex préfrontal ont du mal à s'adapter à cette tâche et restent généralement fixés sur le mauvais paquet (Bechara et al., 1994). Antonio Damasio fait ainsi naître les études du **lien entre la prise de décision et le rôle du vmPFC**.

Cependant, bien que les premiers travaux des Damasio sur ce sujet avec des patients atteints de lésions au vmPFC datent maintenant d'il y a presque 30 ans, le rôle précis du vmPFC dans la prise de décision n'est pas encore complètement élucidé, en particulier pour mieux comprendre les conséquences fonctionnelles de son absence chez les patients atteints de lésions, comme les patients atteints de lésions fronto-temporales. Chez le primate non-humain, où il est possible de générer une lésion spécifique du vmPFC, les résultats montrent qu'une lésion du

## Introduction

vmPFC, l'aire 14 de Brodmann, ne semble pas changer significativement le comportement des primates dans une tâche où ils doivent apprendre à choisir entre différents stimuli, associés soit à un jus de haute valeur, soit à un jus de valeur faible (Papageorgiou et al., 2017). Chez l'homme, les lésions du vmPFC ne semblent pas non plus avoir impacté le comportement dans une tâche basée sur des préférences esthétiques (Aridan et al., 2019). Néanmoins, vu l'implication forte du striatum ventral dans ce type de tâche d'apprentissage par renforcement ou dans des tâches de préférence et ce, d'autant plus que la tâche a été répétée plusieurs fois, pendant des jours et des jours, il est à croire que le striatum ventral a pu assurer un comportement optimal dans ces tâches, même en l'absence du vmPFC. Cette critique s'applique d'ailleurs probablement à un certain nombre de tâches effectuées chez des patients atteints de lésions du vmPFC, où le striatum ventral suffit à effectuer la tâche à lui seul. Cependant, une expérience conduite chez 91 patients atteints de lésions du vmPFC a montré un lien fort entre une atteinte de cette région corticale, une réduction de la cohérence dans l'évaluation d'items alimentaires et le score d'apathie obtenu chez ces patients (Hogeveen et al., 2016). De plus, d'autres études ont déjà montré que les patients atteints de lésions du vmPFC ont tendance à violer le GARP (voir *Économie*) plus souvent que des participants contrôles, c'est-à-dire que **leurs préférences ne sont pas cohérentes** (Camille et al., 2011; Henri-Bhargava et al., 2012). Ceci est en accord avec le fait que les patients atteints de lésions au vmPFC paraissent très **instables dans leurs préférences**, notamment au niveau alimentaire (Ikeda et al., 2002). Le rôle exact du vmPFC reste encore à mieux définir mais il semblerait qu'il soit impliqué dans la stabilité des préférences et dans l'humeur.

Une **lésion des ganglions de la base**, un ensemble de zones fonctionnant en circuit qui inclue le striatum ventral, comme on peut en observer dans le syndrome de perte d'auto-activation psychique (AAD : auto-activation deficit syndrome), a un impact nettement plus impressionnant sur la motivation. Les patients atteints de ce syndrome sont en effet incapables d'engager un effort volontairement comme le nom du syndrome l'indique (Laplane and Dubois, 2001; Laplane et al., 1984). Dans le cas où ils font un effort sous les instructions de l'expérimentateur, ils semblent incapables de jauger leur effort en fonction du niveau de récompense en jeu (Schmidt et al., 2008). Ces résultats vont dans le sens de la théorie de Kent Berridge qui donnait un rôle important au striatum ventral dans les mécanismes incitatifs (*wanting*) (voir *Histoire de la découverte des centres cérébraux du plaisir* et (Berridge et al., 2009)). Il semblerait en effet que les patients atteints de lésions à ce niveau sont incapables de

## Introduction

désirer obtenir quelque chose par eux-mêmes alors qu'ils sont toujours capables de générer des mouvements.

Comme on l'a vu précédemment, le réseau du plaisir est redondant (voir *Histoire de la découverte des centres cérébraux du plaisir*). Il est donc possible que, là où le striatum ventral est nécessaire à son bon fonctionnement, le vmPFC, apparu plus récemment dans l'évolution, ait un rôle moins important, dans des circonstances précises et/ou qu'il permette juste d'amplifier l'activité du striatum ventral, auquel il est connecté, en fonction du contexte et d'autres attributs. Une autre possibilité bien sûr est que le vmPFC corrèle avec la valeur subjective sans pour autant que cela ait un impact direct et nécessaire sur la prise de décision, mais que cela soit lié au fait qu'il est impliqué dans l'encodage général de l'humeur positive (Vinckier et al., 2018). Cela impliquerait qu'il ne joue un rôle dans la prise de décision que lorsque l'humeur est un facteur important pour la prise de décision. Cependant, l'humeur ayant aussi été corrélée au striatum ventral (Knutson et al., 2014; Rutledge et al., 2014), l'idée d'une dissociation nette entre une zone qui serait purement dédiée à l'humeur et une zone purement dédié à la prise de décision ne peut pas non plus être tenue pour valable et ne tiendrait, de toute façon, que si on voulait croire qu'une zone ne peut être impliquée que dans une seule fonction ce qui est probablement loin d'être vrai.

## Résumé

Pour résumer ce qui a été vu jusqu'à présent, on peut dire que :

- Un réseau chez l'animal a été clairement identifié en lien avec la sensation de plaisir. Il comprend le noyau accumbens, le pallidum ventral et aussi le noyau parabrachial ;
- La sensation de plaisir (*liking*) correspond à des corrélats neuraux différents du désir (*wanting*) ;
- Un réseau fonctionnel chez l'homme a été identifié en corrélation avec la valeur subjective. Il comprend le striatum ventral, le cortex préfrontal ventromédian et peut-être aussi le cortex cingulaire postérieur et l'hippocampe ;
- Ce réseau a plusieurs propriétés : il est générique, automatique et son niveau d'activité peut biaiser l'évaluation qu'il est présentement en train de faire ;
- L'organisation fonctionnelle du cortex orbitofrontal reste à élucider, en particulier en ce qui concerne le rôle fonctionnel du cortex orbitofrontal médian par rapport au cortex orbitofrontal latéral ;



## II Les dévaluateurs de la valeur subjective

Lors d'un de mes premiers entretiens avec Mathias Pessiglione, il m'a présenté la motivation comme pouvant se décomposer en quatre grandes dimensions : la récompense, le délai, la punition et l'effort. J'ai eu beau y songer pendant plusieurs jours, ayant l'impression que cette proposition paraissait beaucoup trop simple pour être vraie, j'ai dû finir par m'y résoudre. Ces quatre dimensions, avec les probabilités qui modulent chacune d'entre elles, semblaient bien résumer l'ensemble des éléments qui déterminent la motivation d'un individu. Après avoir amplement étudié la question des circuits cérébraux associés aux récompenses, nous allons faire un tour d'horizon des différents dévaluateurs des récompenses que nous venons d'évoquer. Nous allons ainsi chercher à comprendre comment le délai, les punitions, l'effort, physique ou mental, ainsi que les probabilités, modulent la valeur subjective associée à des récompenses. Nous chercherons ainsi à mieux comprendre comment ils impactent la valeur subjective et par quelles voies neurales cette dévaluation des récompenses a lieu.

### II.1 L'incertitude

« *'I should like to buy an egg, please,' she said timidly. How do you sell them ?* »

*'Fivepence farthing for one-twopence for two,' the Sheep replied.*

*'Then two are cheaper than one ?' Alice said in a surprised tone, taking out her purse.*

*'Only you must eat them both, if you buy two,' said the Sheep.*

*'Then I'll have one, please,' said Alice, as she put the money down on the counter. For she thought to herself, 'They mightn't be at all nice, you know. »<sup>8</sup>*

*Through the Looking-Glass, Lewis Carroll (Carroll, 2009)*

---

<sup>8</sup> « *J'aimerais acheter un œuf, s'il-vous-plaît, dit-elle, intimidée : Combien les vendez-vous ?*

- *Dix sous l'un ; quatre sous les deux, répondit la Brebis.*
- *Deux œufs coûtent donc moins cher qu'un seul ? demande Alice, surprise, en tirant de sa poche son porte-monnaie.*
- *Oui, mais si vous en achetez deux, vous êtes absolument obligée de les manger tous deux, répliqua la Brebis.*
- *Alors je n'en prendrai qu'un seul, s'il-vous-plaît », dit Alice en déposant l'argent sur le comptoir. Car, en son for intérieur, elle pensait : « Il est possible qu'ils ne soient pas fameux, voyez-vous bien. » » (Carroll, 2017)*

## Introduction

Dans ce conte célèbre de Lewis Carroll, Alice est confrontée à un choix qui a l'air absurde au prime abord : payer moins cher pour obtenir deux œufs, plutôt qu'un seul. Cependant, on lui dit que si elle achète les deux œufs, elle sera obligée de les manger tous les deux. Elle a peur de ne pas aimer le goût de ces œufs. L'idée de devoir peut-être avoir à manger deux œufs de goût désagréable la convint de n'en prendre qu'un quitte à payer plus cher. Ce passage, qui peut d'abord surprendre, finit ainsi par faire sens quand on lit l'argument d'Alice. Il se trouve que, comme Alice dans le conte, la plupart des individus semblent présenter une aversion pour l'incertitude. Dans cette partie, nous allons étudier les différentes formes d'incertitude, voir comment elles impactent la valeur subjective et comment elles sont encodées dans le cerveau.

### II.1.a L'incertitude est subjective

*« On ne saurait considérer comme irrationnel un homme prudent qui préfère avoir une moindre espérance mathématique psychologique s'il peut bénéficier d'une dispersion plus faible. On ne saurait non plus considérer comme irrationnel un individu qui aime le risque en tant que tel, c'est-à-dire un individu qui préfère avoir une moindre espérance mathématique psychologique à condition de pouvoir disposer de quelques possibilités d'un gain psychologique très élevé. On pourra dire, si on veut, qu'un tel individu est imprudent, ce qu'il peut être en effet, mais en aucun cas il ne nous paraît possible de dire qu'il est irrationnel. On ne saurait trop le souligner, en dehors de la condition de cohérence, il n'y a pas de critère de la rationalité des fins considérées en elles-mêmes. Ces fins sont absolument arbitraires. Aimer les perspectives aléatoires présentant une très grande dispersion pourra sembler irrationnel à un homme prudent, mais pour celui qui a cette inclination elle n'a rien en elle-même qui soit irrationnel. Il en est ici comme en matière de goûts. Ils sont ce qu'ils sont. Ce sont des données qui diffèrent d'un individu à l'autre. »*

*Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine, Maurice Allais (Allais, 1953)*

Nous avons déjà vu comment les mathématiciens ont régulièrement réintégré la notion de probabilité dans les théories des économistes (voir *Mathématiques*). Les mathématiciens avaient en effet réussi à adapter les théories économiques à **deux cas d'incertitude** identifiés par Frank Knight :

- 1) le cas où l'obtention d'un gain dépend d'une **probabilité connue** notamment via les travaux de Von Neumann et Morgenstern. On parle alors de **risque** en économie ;

## Introduction

- 2) le cas où l'obtention d'un gain dépend d'une **probabilité inconnue** via les apports de Savage (voir *Mathématiques*). On parle alors d'incertitude pour Frank Knight (Knight, 1921) et d'**ambiguïté** en économie depuis (Ellsberg, 1961).

En 1953, l'économiste et physicien français **Maurice Allais** va énoncer un paradoxe qui porte son nom, le **paradoxe d'Allais**. Ce paradoxe va mettre à mal la théorie de l'utilité espérée subjective (Allais, 1953). Pour ce faire, Maurice Allais énoncera un exemple au cours d'une conférence en France pour lequel de nombreux participants, dont Savage lui-même, feront des choix qui ne respectent pas certains des axiomes des théories économiques de l'époque. Savage aurait accusé le vin du midi d'être responsable des choix qu'il a effectué qui bafouent sa propre théorie (Glimcher and Fehr, 2014). Considérons une situation où les participants doivent d'abord choisir entre 2 loteries qui donnent des gains avec des probabilités différentes :

Loterie A : 100€ avec 100% de chance de gagner

Loterie B : 150€ avec 90% de chance de gagner ou 10% de ne rien gagner

La plupart des participants préfèrent l'option sûre, c'est-à-dire la loterie A. Pourtant les gains qui y sont associés ( $100€ < 150€$ ) et son espérance, selon le théorème de Pascal, est inférieure :  $10\,000€$  pour la loterie A  $< 13\,500€$  pour la loterie B. Ceci s'explique par l'aversion au risque dans le domaine des gains que nous avons déjà évoqué précédemment (voir *L'économie comportementale*). Ensuite, imaginons que les participants ont le choix entre deux nouvelles loteries :

Loterie C : 100€ avec 10% de chance de gagner ou 0€ avec 90%

Loterie D : 150€ avec 9% de chance de gagner ou 0€ avec 91%

La théorie économique classique de l'agent rationnel prédit que dans ce cas les participants devraient préférer C à D, tout comme ils ont préféré A à B. En effet, les gains potentiels associés à chaque option n'ont pas bougé (100€ pour la loterie A et la loterie C, 150€ pour la loterie B et D). Seules les probabilités sont différentes pour chaque option, mais la deuxième loterie est toujours associée à une probabilité plus faible (D :  $9\% < C : 10\%$  ; B :  $90\% < A : 100\%$ ). Les participants tendent pourtant à préférer D à C car le gain potentiel est supérieur ( $150€ > 100€$ ) pour une différence de probabilité minimale (1%). Ce résultat viole l'axiome d'indépendance qui postulait que si les individus préfèrent A à B, ils auraient dû aussi préférer C à D puisque les montants en jeu sont les mêmes et seules les probabilités qui y sont associées varient. Une

## Introduction

possibilité pour expliquer ce résultat est que la fonction de probabilité interne subjective des individus n'est **pas linéaire**.

En 1961, Daniel Ellsberg énoncera un deuxième paradoxe concernant la théorie de l'utilité espérée qui porte depuis son nom : le **paradoxe d'Ellsberg**. Le paradoxe consiste encore simplement à montrer, sur la base du comportement des individus, que certains axiomes de Savage ne sont pas vérifiés. On considère une urne remplie de 90 boules de différentes couleurs (rouges, noires et jaunes) mais dont on connaît uniquement la proportion de boules rouges (30/90). Il est alors proposé de choisir entre ces deux paris

- A) Parier sur piocher une boule rouge pour gagner 100\$ (les deux autres couleurs ne font rien gagner)
- B) Parier sur piocher une boule noire pour gagner 100\$ (les deux autres couleurs ne font rien gagner)

La plupart des gens préfèrent l'option A, c'est-à-dire celle où la probabilité de gain est connue (on sait que 1/3 des boules sont rouges). Maintenant, Ellsberg propose de corser les choses avec un nouveau set de choix :

- C) Parier sur piocher une boule rouge ou une boule jaune pour gagner 100\$ (les boules noires ne font rien gagner)
- D) Parier sur piocher une boule noire ou une boule jaune pour gagner 100\$ (les boules rouges ne font rien gagner)

Daniel Ellsberg constate que la plupart des individus préfèrent l'option D à l'option C. Or, d'après le principe de transitivité, aussi appelée l'« ordre complet des actions » et le « principe de la chose sûre » qui font partie des axiomes de Savage, les individus devraient opter pour l'option sûre dans les deux cas. En d'autres termes, ils devraient parier sur les options A et C qui contiennent la boule rouge de probabilité connue. Ces décisions s'expliquent mieux par une **aversion pour l'ambiguïté**. En effet, les individus connaissent la probabilité de piocher une boule noire ou jaune ( $3/3 - 1/3$  de piocher une boule rouge =  $2/3$ ) mais pas la probabilité de piocher une boule rouge ou une boule jaune ( $1/3 + ? = ?$ ) (Ellsberg, 1961). Il semblerait ainsi que les individus préfèrent opter pour l'option de probabilité connue que pour l'option de probabilité inconnue

Daniel Kahneman et Amos Tversky, en développant la théorie des perspectives (voir *L'économie comportementale*) vont permettre de résoudre partiellement ces différents paradoxes. En lien avec le risque, ils vont faire deux découvertes principales. Premièrement, ils

## Introduction

découvrent que les individus distordent les probabilités objectives. Ils ont en effet tendance à **surestimer les probabilités faibles** (typiquement lorsque l'on joue au casino ou aux jeux à gratter par exemple) et ils **sous-estiment les événements à probabilité forte** (voir **Figure 9**). Cette transformation permet d'expliquer le paradoxe d'Allais puisque cela montre bien que la fonction de transformation des probabilités objectives en probabilités subjectives n'est pas linéaire comme son paradoxe l'avait déjà suggéré.



**Figure 9 : La transformation subjective des probabilités vues par la théorie des perspectives.**  $P$  est la probabilité objective d'un événement,  $w(p)$  la transformation qu'en font les individus, représentée en pointillés sur la figure. Le trait continu permet juste de comparer au cas où  $w(p)$  serait égal à  $p$ . Cette figure en S-inversé permet de voir que dans le cas des probabilités faibles, la transformation surcote la probabilité d'un événement alors que dans le cas des probabilités modérées ou élevées, l'individu sous-estime la probabilité d'un événement. Figure issue de (Fox and Poldrack, 2014).

Deuxièmement, Daniel Kahneman et Amos Tversky vont aussi souligner que l'aversion au risque dans le domaine des gains s'inverse dans le domaine des pertes (dans la majorité de la population). Ce qui se traduit par une tendance générale à valoriser le risque dans le domaine des pertes, dit autrement : préférer une perte qui a une certaine probabilité de ne pas survenir à une perte sûre.

Si l'incertitude se distingue au niveau interne entre le risque et l'ambiguïté comme on l'a déjà signalé, on peut aussi distinguer, d'un point de vue objectif deux autres types d'incertitude : l'incertitude aléatoire et l'incertitude épistémique. L'incertitude aléatoire serait liée à la stochasticité du monde et elle est, de ce fait, considérée comme non-réductible. L'incertitude épistémique, au contraire, serait simplement liée à l'ignorance des contingences dans le monde. Apprendre à connaître ces contingences permettrait ainsi de réduire ce niveau

## Introduction

d'incertitude (Tobler and Weber, 2014). La plupart des animaux, comme l'homme, réduisent l'incertitude épistémique par **l'expérience**, c'est-à-dire via un apprentissage. Cela revient par exemple à identifier que telle zone géographique est associée à une probabilité élevée de contenir de la nourriture à telle période de l'année. Cependant, grâce à la **communication** orale et écrite, l'homme dispose d'une capacité supplémentaire de réduire cette incertitude épistémique en changeant ses *a priori* sur le monde sans avoir besoin de passer par l'expérience. Si l'on me dit oralement, ou si j'apprends sur internet, que manger une salamandre peut être toxique pour ma santé, je n'ai pas besoin d'en faire l'expérience pour me former une idée par exemple.

Cependant, l'attitude des individus par rapport au risque peut varier significativement selon qu'elle est basée sur l'expérience ou sur une indication à propos des probabilités objectives. En anglais cette différence est résumée par l'expression « ***description-experience gap*** » (**le fossé entre la description et l'expérience**) (Barron and Erev, 2003; Camilleri and Newell, 2013; Hertwig and Erev, 2009; Hertwig et al.; Lejarraga et al., 2016). Ceci pourrait expliquer partiellement pourquoi, selon les situations, les individus ne sont **pas constants dans leur perception des risques** (Bromiley and Curley, 1992; Tobler and Weber, 2014). L'alternative étant que la perception du risque n'est pas un trait individuel. Il n'y aurait ainsi pas soit des individus prudents, soit des individus qui aiment prendre des risques en général, mais une perception du risque qui dépend de la situation. Seul le futur et les expériences à venir nous diront dans quelle mesure la perception des risques peut être vue comme un trait de personnalité et/ou si c'est juste un concept qui s'adapte pour chaque situation donnée pour un individu donné.

À la formule simple de Pascal  $SV = R * P$  avec  $SV$  la valeur subjective,  $R$  le montant en jeu et  $P$  la probabilité, Savage avait déjà substitué l'idée que  $P$  doit être une fonction subjective interne et non pas une représentation fidèle des probabilités objectives. La théorie des perspectives avait montré que cette fonction n'était pas linéaire en fonction des probabilités. En 1991, **Rachlin** va tester plusieurs fonctions et il va proposer que la valeur subjective est dévaluée soit de manière hyperbolique par le risque sous la forme  $SV = \frac{R}{1+k*(\frac{1}{p}-1)}$  où  $p$  est la probabilité d'obtenir le montant  $R$ , soit de manière exponentielle par une équation de type  $SV = R * e^{-k*(\frac{1}{p}-1)}$  (Rachlin et al., 1991). Depuis, plusieurs études ont montré que la **forme hyperbolique** expliquait mieux les données et le modèle exponentiel a été abandonné

## Introduction

(Takahashi et al., 2007). Plus tard, Levy *et al.* proposent de modéliser l'impact des différents types d'incertitude sur la valeur subjective par la formule :

$$SV = \left[ p - \beta * \left( \frac{A}{2} \right) \right] * V^\alpha$$

où  $SV$  est la valeur subjective,  $p$  la probabilité objective,  $A$  le niveau d'ambiguïté,  $V$  le montant en jeu et  $\alpha$  et  $\beta$  sont des paramètres subjectifs permettant de définir l'aversion pour le risque et l'ambiguïté, respectivement, d'un sujet donné (Levy et al., 2010). Un sujet qui n'est pas impacté par l'ambiguïté aura un  $\beta$  nul et un sujet non-affecté par le risque un  $\alpha$  égal à un. Si le sujet est « prudent » comme dirait Allais, c'est-à-dire a une aversion pour le risque,  $\alpha < 1$ , et dans le cas où il recherche la prise de risque, on aura  $\alpha > 1$ .

En résumé, on peut identifier différents types d'incertitude : celle issue de probabilités connues, le **risque**, et celle issue de probabilités inconnues, **l'ambiguïté**. Quand un environnement est fait d'événements qui ont des probabilités fréquemment changeantes, on parle d'un environnement avec une **volatilité** élevée. La notion de volatilité prise dans ce sens semble venir de la finance pour parler de l'ampleur des fluctuations des cours en bourse<sup>9</sup>. Ce terme a été adopté par les neuroéconomistes dans un sens légèrement différent pour désigner la propriété d'un environnement difficilement prédictible par un observateur du fait des fluctuations fréquentes des probabilités associées. Dans tous les cas, on constate que les individus présentent en majorité **une aversion pour le risque dans le domaine des gains**, mais qu'ils **recherchent le risque dans le domaine des pertes**. Ils ont tendance à **surestimer les probabilités faibles** et à **sous-estimer les probabilités fortes**. On peut aussi dire que les individus ont une **aversion pour l'ambiguïté** comme cela a été montré, notamment par le paradoxe d'Ellsberg. Enfin, l'impact du risque et de l'ambiguïté sur la valeur subjective est lui aussi **subjectif** puisque les individus diffèrent dans leur sensibilité à ces deux composantes, on parle d'individus recherchant le risque (*risk-seekers*) et d'autres qui ont plutôt tendance à l'éviter (*risk-averse*). Nous allons maintenant étudier comment ces différents types d'incertitude sont évalués dans le cerveau.

### II.1.b Les bases neurales de l'incertitude

Nous l'avons dit précédemment, il faut distinguer plusieurs types d'incertitude :

---

<sup>9</sup> [https://fr.wikipedia.org/wiki/Volatilit%C3%A9\\_\(finance\)](https://fr.wikipedia.org/wiki/Volatilit%C3%A9_(finance))

## Introduction

- **le risque** : un évènement souhaité (ou à éviter) est associé à des probabilités connues du sujet ;
- **l'ambiguïté** : quand un évènement souhaité (ou à éviter) est associé à une probabilité complètement inconnue de l'individu ;
- **la volatilité** : le fait que les probabilités dans l'environnement sont plus ou moins instables.

Il faudrait d'ailleurs séparer, dans l'ambiguïté, celle qui est liée à de l'incertitude aléatoire non-réductible par l'individu et celle qui est liée à de l'incertitude épistémique, c'est-à-dire qu'il est possible de réduire empiriquement ou en obtenant de l'information (Tobler and Weber, 2014). A ces différents types d'incertitude, on va pouvoir retrouver différents corrélats neuraux.

En 2007, Timothy **Behrens** va associer l'estimation de la **volatilité** de l'environnement à l'activité d'une zone qu'il appellera le cortex cingulaire antérieur (**ACC**) (Behrens et al., 2007). Cette zone cérébrale a déjà reçu de nombreuses appellations. Dans cette thèse, nous nous tiendrons à appeler cette zone le **cortex préfrontal dorso-médian (dmPFC)**. Cette zone se situe généralement un peu au-dessus et postérieure au genou du corps calleux et à la frontière entre le cortex cingulaire et la zone adjacente au niveau dorsal (voir **Figure 20**). Behrens affirme, sur la base de ses résultats, que cette estimation sert à moduler le taux d'apprentissage de manière dynamique. Le dmPFC servirait à identifier les situations où la volatilité est grande, c'est-à-dire que les probabilités associées à chaque évènement changent très vite, et à faire augmenter le taux d'apprentissage le cas échéant. Un taux d'apprentissage plus élevé revient à donner plus de poids aux évènements récents qui devraient refléter plus fidèlement l'état actuel des probabilités dans l'environnement. Cette proposition a cependant été critiquée depuis par l'équipe de Massimo Silvetti qui montre que l'activité du dmPFC est mieux expliquée par un **codage général de l'incertitude** lié aux erreurs de prédiction, plus qu'à la volatilité en particulier (Silvetti et al., 2013). Il a aussi été proposé que le dmPFC encode le bruit dans l'apprentissage lors de la mise à jour des valeurs associées aux différentes options d'une tâche d'apprentissage par renforcement (Findling et al., 2019). Son activité dans des environnements plus volatils ne serait ainsi pas directement liée à l'estimation de la volatilité en soi, mais au fait que le bruit dans l'estimation associée aux différentes options à disposition est accru. Cela pose cependant le problème de savoir comment ce niveau de bruit dans l'apprentissage est mis à jour en fonction de la volatilité de l'environnement.





**Figure 10 : Cluster au sein du cortex préfrontal dorso-médian corrélant à la volatilité.** Figure issue de (Behrens et al., 2007).

Une méta-analyse récente sur le sujet suggère que différentes zones pourraient être impliquées selon qu'on étudie les prises de décisions dans une condition avec du risque ou avec de l'ambiguïté. Le **dmPFC** en particulier serait plus spécifiquement activé dans une condition de **risque** que dans une condition ambiguë (Poudel et al., 2020), alors que c'est le **cortex préfrontal dorsolatéral** (*dorsolateral prefrontal cortex : dlPFC*) qui réagirait dans le cas d'une situation **ambigüe** (Huettel et al., 2006; Poudel et al., 2020). Ce résultat reste cependant à confirmer par de futures analyses puisque des différences entre les tâches utilisées pour représenter chaque concept dans la méta-analyse pourraient aussi expliquer cette division.

### **II.1.c Le réseau des valeurs est impacté par l'incertitude**

En plus du fait qu'il semblerait qu'une (ou plusieurs) zones du cerveau semble(nt) corrélées avec les différents types d'incertitude dans l'environnement, il est intéressant aussi de constater que l'incertitude va directement impacter l'activité du réseau des valeurs. Il a notamment été montré que l'activité du striatum ventral respectait l'une des propositions de la théorie des perspectives en surestimant les probabilités faibles et en sous-estimant les probabilités hautes (voir **Figure 11** et (Hsu et al., 2009)). Plus généralement, il semblerait que le **réseau des valeurs** présenté précédemment (voir *L'estimation des valeurs positives : le réseau des valeurs*) voit son **activité dévaluée par l'incertitude** qu'il s'agisse d'une incertitude

## Introduction

liée à des probabilités connues comme dans le risque ou des probabilités inconnues comme dans l'ambiguïté (Levy et al., 2010).



**Figure 11 : Activité du striatum gauche (à gauche) et du striatum droit (à droite) en fonction de la probabilité.** Figure issue de (Hsu et al., 2009). La corrélation respecte l'une des prédictions de la théorie des perspectives (voir **Figure 9**).

En somme, on voit que l'activité du dmPFC est liée à l'estimation des différentes formes d'incertitude. En particulier, son activité paraît corrélée avec la volatilité, le risque et l'ambiguïté. Les années à venir permettront sans doute de mieux détailler l'interaction entre l'activité du dmPFC et ces différentes formes d'incertitude. De plus, il est relativement clair qu'en parallèle de ce codage direct des différentes formes d'incertitude dans le cerveau, le réseau des valeurs voit aussi son activité directement dévaluée par l'incertitude et de manière cohérente avec le comportement des individus.

## II.2 Le délai

*« Toi, tu n'aimes pas les gens lents*

*Qui mettent toujours des heures*

*À passer leur commande au restaurant*

*Toi, tu n'aimes pas les gens lents*

*Ceux qui bloquent les gens*

*Dans les escalators –ors – ors*

*Escalators – ors*

## Introduction

*Ça c'est vraiment énervant*

*Parce que toi...*

*Tu cours vite, tu roules vite,*

*Tu parles vite, tu penses vite,*

*Tu manges vite, tu bois vite »*

*Vite, vite, vite, La Chanson du Dimanche*<sup>10,11</sup>

Comme cette chanson parodique le souligne, il semblerait que, dans nos sociétés du moins, l'attente est perçue comme quelque chose d'avérsif. En français, l'expression « *temps perdu* » l'illustre bien en montrant que le temps y est vu comme une ressource qu'il faudrait rentabiliser d'une certaine manière. Chez beaucoup d'animaux également, le délai dans l'attente d'une récompense peut être perçu comme très avérsif. Dans cette partie, nous allons essayer de comprendre pourquoi et quelles sont les bases neurales de cette aversion.

### II.2.a L'aversion pour le délai

*« souviens-toi aussi que chacun ne vit que dans le moment présent, et que ce moment ne dure qu'un instant ; le reste, il a été vécu ou est dans l'incertain. »*

*X, Livre III, Pensées pour moi-même, Marc-Aurèle (Marc-Aurèle, 2007)*

Les humains et les animaux en général tendent à avoir une aversion pour le délai, en particulier dans nos sociétés occidentales modernes où la vitesse est souvent présentée comme une valeur<sup>12,13</sup>. Afin de minimiser la pénibilité du délai, nous avons inventé des véhicules allant toujours plus loin à des vitesses toujours plus impressionnantes : nos trains vont à 300km/h pour relier différentes villes entre elles le plus vite possible, le concorde, maintenant abandonné, pouvait aller jusqu'à plus de 2 000km/h<sup>14</sup>, etc. Dans les années 1970, le psychologue américain Walter **Mischel** crée une expérience, toujours très en vogue, où des enfants sont dans une salle avec une récompense alimentaire devant eux. Il peut s'agir de 5 bretzels ou de 2 cookies. Walter Mischel s'assure que la récompense qui est présentée devant eux est l'aliment que les enfants

---

<sup>10</sup> <https://www.paroles.net/la-chanson-du-dimanche/paroles-vite-vite-vite>

<sup>11</sup> <https://www.youtube.com/watch?v=pUATJE58Y-Q>

<sup>12</sup> <https://nypost.com/2014/01/04/societys-addiction-to-faster-living-is-destroying-us-doctor/>

<sup>13</sup> [https://www.scienceshumaines.com/la-tyrannie-de-la-vitesse\\_fr\\_29044.html](https://www.scienceshumaines.com/la-tyrannie-de-la-vitesse_fr_29044.html)

<sup>14</sup> [https://fr.wikipedia.org/wiki/Concorde\\_\(avion\)](https://fr.wikipedia.org/wiki/Concorde_(avion))

## Introduction

aiment le moins entre les bretzels et les cookies. Il indique ensuite aux enfants que, s'ils parviennent à attendre que le psychologue revienne dans la salle sans manger la récompense qui est devant eux, ils obtiendront l'autre récompense, c'est-à-dire les 5 bretzels ou les 2 cookies, qui est plus appétissante pour eux. Cependant, s'ils mangent la récompense avant que l'expérimentateur revienne spontanément dans la salle, l'expérience prendra fin mais ils n'auront pas accès à la récompense plus appétissante (Mischel and Ebbesen, 1970). Ce test est depuis connu depuis sous le nom du **test du marshmallow**. Les bretzels et les cookies ont en effet été remplacés par des marshmallows par la suite : si les enfants arrivent à se retenir de manger un marshmallow qui est présenté devant eux, alors ils en obtiendront deux quand l'expérimentateur reviendra dans la salle. Cette attente semble réellement pénible pour les enfants et ils diffèrent par leur capacité à se retenir de manger la récompense devant eux pour obtenir quelque chose de mieux quelques minutes plus tard.

En 1937, l'économiste Paul Samuelson que nous avons évoqué précédemment (voir *Économie*) a proposé un modèle permettant de modéliser les choix des individus dans le cadre de choix intertemporels (Samuelson, 1937). C'est ce qu'on appelle le **modèle d'utilité décomptée** (*discounted utility model* : **DUT**). Pour savoir quelle est l'utilité future d'une récompense, c'est-à-dire sa valeur subjective attendue *SV*, il propose la formule :

$$SV = \sum_{t=0}^D e^{-k*t} * u(t)$$

*SV* est l'utilité de la récompense au bout du délai *D*, *k* est une constante permettant d'exprimer la sensibilité individuelle au délai, autrement dit le paramètre de décompte temporel et *u(t)* est l'utilité instantanée de consommer la récompense au moment *t*. Depuis, un nombre conséquent d'études a montré qu'une fonction **hyperbolique** correspondait bien mieux aux données observées chez les individus (Frederick et al., 2002; Kalenscher and Pennartz, 2008; Takahashi et al., 2007). Une des raisons justifiant que la fonction hyperbolique corresponde mieux aux données est que le modèle d'utilité décomptée imposait que les préférences intertemporelles soient fixes indépendamment du délai considéré. Par exemple si l'on considère un choix entre les options suivantes :

- A) Obtenir 5€ aujourd'hui
- B) Obtenir 7€ dans 7 jours

Maintenant considérons un autre choix avec les options suivantes :

## Introduction

C) Obtenir 5€ dans 363 jours

D) Obtenir 7€ dans 365 jours

Un participant qui préfère l'option (A) à l'option (B) devrait, d'après le DUT aussi préférer l'option (C) à l'option (D) puisqu'on a juste translaté le set d'option de 363 jours. Cependant, le comportement des individus montre que la plupart des participants présentent une inversion des préférences selon la distance temporelle considérée. Il ne paraît ainsi pas choquant de préférer (A) à (B) dans le premier set d'options et (D) à (C) dans le deuxième set. Désormais, la plupart des études emploient ainsi des fonctions hyperboliques pour représenter l'impact du délai sur la valeur subjective. La formule correspondante est :

$$SV = \frac{V}{1 + k * D}$$

$SV$  est la valeur subjective,  $V$  la valeur monétaire à gagner,  $D$  le délai et  $k$  le taux de décompte temporel subjectif.

En neuroéconomie, la **tâche de décompte intertemporel** a été développée et grandement utilisée pour évaluer la sensibilité subjective des individus au délai. Les individus doivent y choisir entre deux options :

A)  $X1€$  dans un délai court  $T1$  ( $T1 = 0$  systématiquement dans certaines versions)

B)  $X2€$  ( $X2 > X1$ ) dans un délai long  $T2$  ( $T2 > T1$ )

Les individus diffèrent par leur point d'indifférence. Le point d'indifférence correspond au montant  $X2€$  nécessaire pour rendre les offres (A) et (B) équivalentes subjectivement pour un montant  $X1€$ , un délai  $T1$  et un délai  $T2$  donnés. Dans un cas où  $T1 = 0$ , il suffit donc de trouver le montant  $X2$  pour lequel  $SV_A = X1€$  est équivalent à  $SV_B = \frac{X2€}{1+k*T2}$ . Cela permet d'identifier quel est le paramètre  $k$  qui correspond le mieux à chaque individu dans une situation donnée.

Là aussi, comme pour l'incertitude, il semblerait qu'il y ait une **asymétrie** entre le taux de décompte temporel dans le domaine des **gains** et celui des **pertes** (Frederick et al., 2002; Kalenscher and Pennartz, 2008; Zhang et al., 2018). Le paramètre de décompte temporel dans le domaine des pertes serait encore plus important par rapport à celui des gains. En d'autres termes, les individus préfèrent se débarrasser d'une punition au plus vite, plutôt que de vivre dans l'attente de subir l'événement aversif que peut représenter une perte.

## II.2.b Pourquoi le délai est aversif ?

« *La denrée la plus rare, c'est le temps.* »

*Sur la télévision*, Pierre Bourdieu<sup>15</sup>

Les raisons avancées pour l'aversion au délai sont multiples :

- 1) L'une des hypothèses est liée au **coût d'opportunité** (*opportunity cost*) (Fawcett et al., 2012), aussi appelé en français coût de renoncement ou coût de substitution. Ce principe a été introduit par l'économiste David L. **Green** au XIX<sup>ème</sup> siècle (Green, 1894) qui propose que ce qui est réellement aversif dans le travail n'est pas tant la douleur expérimentée par le travail en lui-même, mais plutôt le fait que le travail oblige à ne pas profiter de différentes opportunités de satisfaire ses désirs et de prendre du plaisir. Le salaire servirait alors de compensation pour ces opportunités dont le travailleur n'a pas pu profiter à cause de son travail. L'hypothèse du coût d'opportunité se résume bien par le proverbe français « *Le temps c'est de l'argent* », traduit d'un passage d'une lettre de Benjamin Franklin (« *time is money* ») (Franklin, 1748). Attendre pour une récompense donnée, ce serait autant de temps « gâché » pendant lequel on n'est pas en train de poursuivre d'autres récompenses potentielles dans l'environnement.
- 2) Une autre hypothèse bien sûr pourrait simplement être liée aux **probabilités**. En effet, comme suggéré par la citation de Marc-Aurèle évoquée précédemment, le futur est incertain et donc plus le bien à obtenir paraît lointain, moins la probabilité de pouvoir en profiter paraît subjectivement plausible. L'aversion pour le délai serait donc équivalente à l'aversion pour l'incertitude. Dans ce cas, le délai serait retraduit psychologiquement comme une forme de probabilité (Frederick et al., 2002; Kagel et al., 1986; Rachlin et al., 1991; Yi et al., 2006). Cette proposition semble néanmoins avoir été abandonnée au vu des évidences comportementales et neurales qui semblent montrer qu'il s'agit bien de deux mécanismes différents (Kalenscher and Pennartz, 2008).
- 3) Enfin, la **disponibilité** du bien et le **besoin interne du bien** peuvent fortement contribuer à modifier la perception du délai. Quelqu'un qui a très faim préférera obtenir de la nourriture maintenant, d'autant plus si ses sens y sont directement exposés, même

---

<sup>15</sup> <https://www.youtube.com/watch?v=vcc6AEpjdY>

## Introduction

si elle est moins bonne ou en moins grande quantité, plutôt qu'attendre pour avoir un meilleur repas par exemple (Frederick et al., 2002).

Ce dernier point permet probablement de comprendre pourquoi les mêmes individus peuvent différer dans l'impact du délai sur la valeur subjective en fonction du bien considéré. Il a notamment été montré chez 2.400 ougandais que le délai n'était pas perçu de la même manière selon les biens considérés, sauf chez des traders et des individus possédant déjà les différents biens proposés dans l'étude en quantité chez eux (Ubfal, 2016).

### II.2.c Bases neurales du délai

*« les désirs naissent en lutte les uns contre les autres, et cela se produit quand raison et appétits militent en sens contraires : c'est le propre des êtres qui ont la perception du temps (l'intellect nous pousse à résister en considération de l'avenir, l'appétit nous entraîne dans la seule vue de l'immédiat : car le plaisir du moment paraît être agréable absolument et bon absolument, du fait qu'on ne voit pas l'avenir »*

*Livre troisième, chapitre 10, De l'âme, Aristote (Aristote, 1969)*

Dès 2004, Samuel McClure *et al.* vont étudier les bases neurales de la valeur subjective décomptée par le délai. Ils montrent que le réseau des valeurs, de manière cohérente avec le comportement des individus, est plus actif dans les choix où une des récompenses est disponible maintenant par rapport à ceux où les récompenses sont toutes décalées (voir **Figure 12** et (McClure, 2004)). A l'époque, et sur la base d'une proposition théorique de l'économiste David Laibson (Laibson, 1997), McClure propose qu'il existe deux systèmes de valeurs dans le cerveau. Un de ces systèmes s'occuperait des récompenses à court terme. Il appelle ce système la composante  $\beta$ . Ce système correspond anatomiquement au réseau des valeurs. Le deuxième système qu'il appelle la composante  $\delta$  serait dédié à une planification à plus long terme. Ce deuxième système se baserait sur des aires du réseau exécutif (cortex préfrontal dorsolatéral, dmPFC, etc.). Cette théorie a depuis été largement invalidée car il est maintenant relativement clair que **l'activité du réseau des valeurs est dévaluée par le délai de manière similaire à la valeur subjective déduite du comportement** (Kable, 2014; Kable and Glimcher, 2007; Peters and Büchel, 2011).



**Figure 12 : Le réseau des valeurs est plus actif pour les récompenses immédiates.** A] Cartes d'activation montrant les clusters avec une p.value < 0.001 non-corrigée des zones qui s'activent plus pour des récompenses immédiates que pour des récompenses lointaines. B] Décours temporel de l'activité dans chaque zone montrant l'activité de chacune des zones en fonction du type de choix auxquels les sujets doivent répondre. N= 14 participants. Figure issue de (McClure, 2004)

Un résultat curieux des études lésionnelles, soulevé déjà par l'équipe de Damasio, est que les patients atteints de **lésions au vmPFC semblent dévaluer encore plus fortement les récompenses associées à un délai élevé** par rapport aux sujets sains (Sellitto et al., 2010). Or, ces patients présentent une performance diminuée dans la tâche de pari d'Iowa (voir *Impact des lésions du réseau des valeurs*). Le fait que ces patients aient plus de difficulté à se projeter dans l'avenir est l'une des hypothèses avancées pour expliquer cette performance réduite (Bechara et al., 1994). Si je dis que ce résultat est curieux c'est que, si le vmPFC encode la valeur subjective, on comprend mal pourquoi une lésion de celui-ci implique que seules les récompenses lointaines soient ignorées. On pourrait en effet s'attendre à ce que l'évaluation des récompenses en général en soit impactée et pas seulement celles associées à un délai.

La perception subjective du temps semble corrélée avec l'activité du dmPFC (Coull et al., 2011; Kolling and O'Reilly, 2018; Protopapa et al., 2019; Radua et al., 2014). Cependant, il



## Introduction

paraît très spéculatif, à l'heure actuelle, de dire s'il existe des zones du cerveau qui corréleront spécifiquement avec l'estimation du délai dans les tâches de choix intertemporels. Néanmoins, l'idée initiale de McClure qu'il existe un système des valeurs pour la planification à long terme (McClure, 2004) a évolué dans une nouvelle direction. On aurait ainsi d'un côté le système des valeurs qui évalue la valeur associée à chaque item, équivalent en quelque sorte à la composante  $\beta$ . De l'autre côté, il y aurait un **système de contrôle cognitif qui permettrait de moduler l'activité du réseau des valeurs afin de choisir les récompenses lointaines** lorsque cela est pertinent. Un peu comme dans la citation d'Aristote ci-dessus, on aurait d'un côté le vmPFC dont l'activité serait principalement dirigée par les « appétits », et, de l'autre, l'« intellect », représenté par le **cortex préfrontal dorsolatéral (dIPFC)**, qui modulerait l'activité du vmPFC en vue des buts à long terme (Hare et al., 2014). Peut-être que c'est là l'explication de l'impact des lésions du vmPFC sur le décompte temporel puisque, sans le vmPFC, le cortex préfrontal dorsolatéral pourrait se retrouver dans l'incapacité de moduler l'activité du réseau des valeurs et ce, alors même que le striatum ventral est intact. Dans ce cas, il est facile d'imaginer que le striatum ventral permet toujours d'agir et de faire des choix, mais, sans le circuit allant du dIPFC au vmPFC, il n'arriverait plus à considérer les options à long terme. En effet, il semblerait que le cortex préfrontal dorsolatéral soit fortement connecté au striatum dorsal et au vmPFC, mais pas ou peu au striatum ventral qui est aussi impliqué dans le réseau des valeurs (Choi et al., 2012; Morris et al., 2016; Shen et al., 2019). De plus, il a été montré que lorsque le dIPFC est moins actif, notamment à cause de la fatigue, les individus ont plus tendance à choisir des récompenses immédiates (Blain et al., 2016). Les études à venir permettront peut-être d'éclaircir le rôle spécifique du vmPFC dans le cas de récompenses associées à un délai plus ou moins élevé.

### II.3 L'effort physique

#### II.3.a Le mouvement

*« L'animé semble se distinguer de l'inanimé par deux caractères principaux : le mouvement et la sensation. »*

*Livre premier, chapitre 2, De l'âme, Aristote (Aristote, 1969)*

Contrairement aux plantes qui n'ont pas besoin de se mouvoir pour obtenir des ressources, la plupart des Eumétazoaires passent par la locomotion pour atteindre les ressources nécessaires à leur survie. Une vision très caricaturale du cerveau peut à vrai dire se résumer à un système somatosensoriel permettant d'identifier les ressources disponibles et, par la locomotion, à se diriger vers elles afin de se les procurer. De la même manière que des

## Introduction

organismes unicellulaires simples, comme les Eubactéries ou les Archéobactéries, sont attirés via leurs chémorécepteurs vers les zones les plus concentrées dans les ressources dont elles ont besoin par un système de gradients chimiques, on constate chez les Eumétazoaires des mécanismes similaires. Pour savoir où aller, le « crabe bleu » (*Callinectes sapidus*), par exemple, se fie aux chémorécepteurs de ses antennes<sup>16,17</sup> qui lui permettent de détecter les sécrétions de ses proies et donc leur localisation. Spontanément, il va donc entamer un mouvement vers l'origine de ces sécrétions afin de dévorer la proie qui l'aura sécrétée. La plupart des Mammifères se fient eux aussi à différents inputs sensoriels (odorat, visuel, auditif, etc.) qui vont leur indiquer la localisation spatiale de ressources dans leur environnement immédiat. C'est d'ailleurs sans surprise qu'on peut constater qu'une **partie conséquente du système nerveux humain est dévolue soit à la représentation d'inputs sensoriels** (le cortex visuel occupe environ 20% du cortex humain (Wandell et al., 2007) et le cortex auditif jusqu'à 8% de la surface du cortex<sup>18</sup>, pour ne citer que ces deux systèmes sensoriels), **soit à la motricité** (moelle épinière, cervelet, ganglions de la base, cortex moteur primaire et secondaire, aire motrice supplémentaire ; notez que le cervelet à lui seul représente 10% du volume total de la masse encéphalique contenue dans le crâne et qu'il contiendrait la moitié de la totalité des neurones (Vibert et al., 2005)). De plus, si on regarde l'arbre évolutif qui a mené jusqu'à *Homo Sapiens Sapiens*, on peut constater qu'une partie majeure des différentes « innovations » évolutives qui ont mené jusqu'à nous ont un rapport avec le développement du système sensoriel et surtout du système moteur (voir **Figure 13** et (Cisek, 2019)).

On peut identifier au moins trois types de mouvements (Vibert et al., 2005):

- Les réponses réflexes qui règlent la posture et une réaction spontanée et automatique au niveau des muscles et ne passent pas par la volonté ;
- Les activités motrices rythmiques, principalement automatiques mais qui peuvent être légèrement orientées par la volonté, comme la respiration ou le pouls cardiaque ;
- Les mouvements volontaires.

Ici, nous allons nous concentrer sur ce troisième type de mouvement qui semble être le seul à être lié à une sensation aversive d'effort. Il faut d'ailleurs préciser que la définition d'**effort** que

---

<sup>16</sup> <https://www.thoughtco.com/how-do-crabs-find-food-2291888>

<sup>17</sup> <https://www.bluecrab.info/faq.htm>

<sup>18</sup> <http://www.cochlea.eu/cerveau-auditif/thalamo-cortex/cortex-auditif-organisation>

## Introduction

nous allons prendre et garder pour la suite est la **sensation subjective et aversive liée à l'exécution d'une activité physique ou mentale.**



**Figure 13 : Arbre phylogénétique des animaux avec un focus sur les différentes innovations évolutives qui ont mené jusqu'à l'homme.** Certaines de ces innovations ont un rapport avec les systèmes somatosensoriels : yeux appariés chez les Vertébrés, olfaction avancée chez les Mammifères, mais la plupart ont un lien avec des fonctions motrices : réflexes et locomotion orientée chez les Eumétazoaires, moelle épinière chez les Chordés, ganglions de la base chez les Vertébrés, cervelet chez les Gnathostomes, locomotion terrestre chez les Tétrapodes, etc. Figure issue de (Cisek, 2019).

Bien que toutes ne soient pas recrutées pour tout type de mouvement, il reste impressionnant de constater, comme énoncé précédemment, qu'un nombre considérable de zones du cerveau est impliqué dans la génération du mouvement. Les zones impliquées sont, au moins : les **ganglions de la base** et le **cortex moteur** pour la génération des mouvements volontaires, le **cervelet** pour la coordination et la précision des mouvements, **l'aire prémotrice** qui prépare les systèmes moteurs aux mouvements, **l'aire supplémentaire motrice** qui est

## Introduction

notamment impliquée dans la préparation et la coordination des mouvements complexes, et même certaines parties du **cortex pariétal** pour la coordination des mouvements (Vibert et al., 2005). A la fin, tout va se conclure par un signal sous la forme de potentiels d'actions dans les motoneurons de la moelle épinière qui vont induire la contraction des muscles du corps.

Quand nous parlons d'effort, nous avons en tête une sensation aversive qui accompagne ces mouvements volontaires. D'où vient cette sensation ? Est-elle liée à une lecture directe de la quantité d'énergie consommée pour contracter les muscles ou à l'activité des neurones impliqués dans la contraction des muscles ? Si elle vient des neurones, quelle est la population de neurones à l'origine de cette sensation ?

### II.3.b L'effort physique est avant tout subjectif

En 1894, le philosophe français Guillaume Ferrero, dans son ouvrage *L'inertie mentale et la loi du moindre effort*, dira « *L'homme n'aime pas le travail, ni le travail des muscles, ni le travail du cerveau.* » (Ferrero, 1894). Dans cet article, il développe l'idée d'une **loi du moindre effort**, en particulier dans le domaine mental. C'est d'ailleurs selon cette logique que David Ricardo et Adam Smith proposaient quelques décennies plus tôt que les prix des objets soient déterminés par l'effort requis pour leur production (voir *Économie*). Au début du XX<sup>ème</sup> siècle, l'idée va se développer avec des expérimentalistes et des psychologues qui vont la reprendre sous différentes appellations « *law of less work* » (loi du moindre travail) chez Clark Hull (Hull, 1943), « *principle of least effort* » (principe du moindre effort) chez Waters (Waters, 1937) ou Gengerelli (Gengerelli, 1930), « *law of minimum effort* » (loi de l'effort minimum) chez Tsai (Tsai, 1932) ou encore « *law of least action* » (loi de la moindre action) chez Wheeler (Wheeler, 1940). En somme, l'idée simple de tous ces penseurs est que si deux chemins mènent à une récompense, les animaux tendent à préférer le chemin qui va demander le moins de force physique en intensité (chemin le plus plat ou en descente) et en durée (chemin le plus court), comme s'ils cherchaient à minimiser leurs dépenses en ressources énergétiques. **L'activité physique est aversive** pour les individus qui cherchent autant que possible à la minimiser.

Quelques années plus tard, le chercheur suédois Gunnar Borg va mettre en place une échelle connue depuis sous le nom d'échelle de Borg, ou mesure de perception de l'effort (*rating of perceived exertion, RPE*), pour que les individus puissent quantifier subjectivement l'activité physique effectuée (Borg, 1970; Borg and Linderholm, 1970). Il montre que les évaluations subjectives des sujets sur cette échelle sont en corrélation très forte avec leur pouls cardiaque (Borg, 1970) ou avec la force exercée par les muscles, mais qu'elles semblent aussi dépendre de facteurs motivationnels et de l'aversion à faire la tâche (Borg, 1978). Il est depuis clair que

## Introduction

les évaluations rapportées par les sujets sur l'échelle de Borg sont une construction qui varie autant avec des variables physiologiques comme le pouls cardiaque ou la capacité musculaire qu'avec des facteurs psychologiques subjectifs (Marcora and Staiano, 2010; Morgan, 1994). En somme, **l'effort physique est avant tout subjectif.**

Dès les années 1960, Gunnar Borg avait déjà cherché à identifier une fonction pour estimer comment une intensité d'exercice physique était traduite en une évaluation subjective d'effort. En 1961, il avait proposé la fonction  $SV = a + k * (S - b)^n$  pour comprendre comment les sujets traduisaient l'exercice physique effectué en pédalant sur un vélo d'intérieur à une évaluation d'effort physique, où  $a$  et  $b$  sont des constantes,  $S$  la vitesse à laquelle pédale le sujet et  $n$  l'exposant qui dépend de la tâche en cours. Il identifie en effet que  $n = 1.6$  quand les sujets doivent pédaler sur un vélo d'intérieur et  $n = 2$  quand les sujets marchent dans une expérience similaire où les sujets doivent marcher au lieu de pédaler et il met ces résultats en parallèle avec les résultats d'autres chercheurs qui ont identifié  $n = 1.7$  dans un cas où des sujets serrent une poignée de force avec leur main (Borg, 1978, 1982; Stevens and Mack, 1959). Depuis, plusieurs autres fonctions ont été testées afin de voir comment les récompenses sont décomptées par l'effort physique. La fonction la plus simple étant une fonction linéaire  $SV = R - k * E$  (Burke et al., 2013) où  $SV$  est la valeur subjective,  $R$  le montant en jeu,  $E$  l'effort et  $k$  un paramètre individuel de sensibilité subjective à l'effort. Cette fonction ne semble cependant pas faire l'unanimité. Plusieurs chercheurs, sur la base des résultats des tâches de décompte par le délai, ont utilisé une fonction de décompte hyperbolique sous la forme  $SV = \frac{R}{1+k*E}$  (Prevost et al., 2010; Seaman et al., 2018; Skvortsova et al., 2014; Sugiwaka and Okouchi, 2004). Une des limites de cette fonction hyperbolique est qu'elle implique une valeur subjective toujours positive. Faire un effort, même s'il est élevé et que la récompense est faible vaudrait mieux que de ne rien faire puisque  $SV$  est contraint à être positif au vu de la fonction. Il a ainsi été montré que des fonctions de forme **concave**, c'est-à-dire où le coût augmente peu pour des efforts faibles mais augmente très vite dès que les efforts sont plus conséquents (voir **Figure 14**), représentaient bien mieux la manière dont les récompenses sont décomptées par l'effort physique (Arulpragasam et al., 2018; Białaszek et al., 2017; Chong et al., 2017; Hartmann et al., 2013; Klein-Flügge et al., 2015). La formule  $SV = R - k * E^2$  semble particulièrement bien rendre compte des données (Chong et al., 2017; Hartmann et al., 2013). La fonction qui reflète le mieux l'impact de l'effort physique sur l'évaluation subjective semble être celle qu'avait déjà identifié Borg dans les années 1960.



**Figure 14 : Différents types de décompte de la valeur subjective par l'effort physique sur la base de données réelles.** A gauche, modèle linéaire où  $SV = R - k * E$  ; au milieu modèle hyperbolique où  $SV = \frac{R}{1+k*E}$  ; à droite le modèle qui explique le mieux les données, un modèle de la forme  $SV = R - k * E^2$ . Figure issue de (Hartmann et al., 2013).

Une spécificité non-triviale de l'activité physique est qu'elle peut être associée à une **valeur positive**, bien qu'elle s'accompagne généralement de la sensation aversive d'effort. Dans nos sociétés modernes, il est facile d'obtenir des quantités importantes de nourriture. La prise de poids, et en particulier le diabète et l'obésité, sont devenus des vrais problèmes sociétaux ces dernières années<sup>19,20</sup>. L'activité physique, dans nos sociétés, est ainsi présentée de plus en plus comme étant associée à une valeur positive puisqu'elle permet de limiter le diabète de type 2 et l'obésité. La société associe ainsi une valeur positive à l'activité physique qui va cependant à l'encontre de la sensation d'effort associée à l'activité physique que l'on retrouve chez la plupart des animaux.

Un autre phénomène encore plus intéressant car il est lié à la physiologie du corps humain et non pas à des valeurs sociétales est qu'à partir d'un certain seuil d'activité physique en durée et en intensité, l'exercice physique provoque la sécrétion d'endorphines. Or, ces endorphines génèrent une sensation agréable chez l'individu en train de faire l'exercice physique (Harber and Sutton, 1984). Il y a d'ailleurs des individus qui développent une addiction au sport, comme on pourrait l'être d'une drogue, car ils veulent retrouver cette sensation liée aux endorphines (Hausenblas et al., 2017; Weinstein, 2014). En moyenne, et à travers les individus, l'activité physique est probablement majoritairement associée à la sensation aversive d'effort. Cependant, il se peut ainsi que chez certains individus, selon qu'ils ont ou non régulièrement expérimenté ce lien entre l'activité physique et les endorphines ou selon l'importance qu'ils

---

<sup>19</sup> <https://solidarites-sante.gouv.fr/soins-et-maladies/maladies/article/diabete>

<sup>20</sup> <https://www.who.int/bulletin/volumes/90/8/12-040812/fr/>

## Introduction

donnent à l'idée de garder un corps sain, la valeur subjective d'une activité physique donnée soit positive.

### II.3.c Bases neurales de l'effort physique

« *Unique est donc le principe moteur premier : l'objet désirable.* »

*Livre troisième, chapitre 10, De l'âme, Aristote (Aristote, 1969),*

L'effort, c'est-à-dire la sensation aversive liée à de l'activité musculaire, en donnant une idée des ressources mobilisées par les individus peut leur permettre d'avoir un proxy des ressources engagées dans une action donnée. Sous ce point de vue, on comprend déjà mieux pourquoi il est subjectif (voir *L'effort physique est avant tout subjectif*), c'est-à-dire potentiellement relatif aux ressources internes de l'individu, et non juste une mesure absolue de l'énergie dépensée dans l'activité physique ou des ressources consommées pour ce faire. Une autre question reste cependant de savoir quelle est la **source de cette sensation subjective**. Elle pourrait en effet dépendre soit de l'activité corticale qui commande l'activité musculaire soit de l'activité musculaire en elle-même. Les deux étant très corrélées, il n'est bien sûr pas évident de distinguer la part de l'une de la part de l'autre. En 2001, Gallagher *et al.* vont montrer que des individus à qui on a injecté du curare expriment une sensation subjective plus forte d'exercer un effort pour le même niveau de contraction musculaire que lorsqu'ils le faisaient sans curare. Or le curare agit sur la jonction neuromusculaire, c'est-à-dire la réalisation par le muscle de la commande envoyée par les motoneurones, et pas directement sur le muscle. Cela suggère déjà que c'est au niveau des neurones et non des muscles que naît la sensation d'effort subjectif.

En 2015, Alexandre Zénon va demander à 12 participants d'exercer différents niveaux de force alors qu'ils subissent soit 1) une condition contrôle, soit 2) une inhibition de leur cortex moteur primaire gauche (les sujets exercent une force avec leur main droite) via une stimulation magnétique transcrânienne, soit 3) une inhibition de leur **aire supplémentaire motrice (SMA)** (Zénon *et al.*, 2015). Cette zone a déjà été identifiée, avec le dmPFC, comme corrélant avec l'intensité d'un effort physique (Cramer *et al.*, 2002; Kurniawan *et al.*, 2013; Pessiglione *et al.*, 2007; Spraker *et al.*, 2007). Zénon *et al.* vont construire une variable dépendant de plusieurs mesures (évaluations de la pénibilité des efforts, taille de la pupille, etc.) pour déterminer la perception subjective de l'effort effectué par les sujets, indépendamment de l'effort effectué, qui reste le même, qu'ils subissent ou non une inhibition. Ils vont montrer que seule une perturbation de la SMA, mais pas du cortex moteur, va impacter cette variable qui représente

## Introduction

la perception subjective de l'effort (Zénon et al., 2015). Leur hypothèse est que la SMA envoie une copie de la commande motrice aux aires somatosensorielles et que c'est là où la sensation d'effort va naître. Il est intéressant au passage de constater que cette zone est légèrement postérieure et semble même se superposer légèrement au dmPFC qu'on a évoqué précédemment. Il n'est donc pas exclu que le dmPFC soit impliqué aussi dans cette sensation.

### II.3.d Liens bidirectionnels entre réseau des valeurs et effort physique

Pour ce qui est de la valeur subjective décomptée par l'effort physique, peu d'articles s'y sont pour l'instant intéressés directement, mais il paraîtrait que là encore, **le réseau des valeurs, en particulier le vmPFC, encode bien la valeur décomptée par l'effort physique** et pas seulement la valeur des gains (Arulpragasam et al., 2018; Seaman et al., 2018).

Les ganglions de la base, qui existent depuis plusieurs millions d'années dans les cerveaux des Vertébrés (Grillner and Robertson, 2016), jouent un rôle important dans la génération des mouvements (Vibert et al., 2005). En effet, leur dégénérescence entraîne différents troubles moteurs, avec par exemple des mouvements involontaires irrépressibles comme dans la chorée de Huntington. Le noyau accumbens, qui fait partie des ganglions de la base, joue un rôle dans les mécanismes de volition vis-à-vis d'un but (*wanting*) (voir *Histoire de la découverte des centres cérébraux du plaisir* et (Berridge et al., 2009)). La dopamine qui est envoyée à ce noyau serait nécessaire, mais pas suffisante, pour surmonter les coûts liés à une action qui permettrait d'obtenir une récompense importante (Kurniawan, 2011; Phillips et al., 2007). Enfin, le striatum ventral médie l'effet des enjeux monétaires sur la force exercée dans des tâches où la récompense monétaire de l'expérience dépend de la force exercée à chaque essai (Pessiglione et al., 2007; Schmidt et al., 2012). Le noyau accumbens reçoit en effet d'un côté des informations sur les récompenses en jeu via la dopamine envoyée par la substance noire et l'aire tegmentale ventrale, et aussi via ses connections avec le vmPFC (voir *Bases neurales de la valeur*). De l'autre côté, le noyau accumbens est aussi fortement connecté aux autres ganglions de la base qui sont plus directement impliqués dans la génération de commandes motrices. Il a été suggéré que, selon si l'effort à effectuer est de type mental ou physique, il entraîne soit l'activité du noyau caudé soit celle du putamen, respectivement (Schmidt et al., 2012). Enfin, s'ils sont toujours capables de générer des mouvements de différentes intensités, quand on leur demande de le faire, les patients atteints de déficit d'auto-activation psychique lié à des lésions striato-pallidales sont incapables de faire un effort proportionnel à la récompense en jeu (voir **Figure 15** et (Schmidt et al., 2008)). En conclusion, non seulement **le réseau des valeurs**



**encode la valeur d'une récompense décomptée par le coût de l'effort** à faire, mais il est aussi impliqué dans **la génération de l'effort** lui-même.



**Figure 15 : Impact des lésions aux ganglions de la base sur la capacité à doser son effort en fonction des récompenses en jeu.** Figures du haut : tâche où on demande explicitement aux participants d'effectuer différents niveaux d'effort en serrant une poignée de force. Figures du bas : tâche où la rémunération monétaire des participants dépend de l'intensité avec laquelle ils auront serré la poignée de force. Les sujets contrôles à gauche dosent leur effort en fonction des récompenses en jeu, de même que les patients contrôles atteints de la maladie de Parkinson à droite. Les patients atteints de déficit d'auto-activation psychique dû à des lésions striato-pallidales, dans les figures du milieu, sont incapables de doser leur effort à moins que cela soit explicitement demandé dans les instructions. Figure issue de (Schmidt et al., 2008).

### Résumé intermédiaire

Une partie conséquente de notre système nerveux est dévolue à la génération de mouvements et d'activité physique à la suite de plusieurs millions d'années de sélection par l'évolution. L'activité physique va de pair avec des coûts énergétiques (dépendance d'ATP, de réserves de glucides et de lipides) et de production de déchets (lactate) dans l'organisme. L'évolution semble avoir sélectionné un mécanisme qui donne un proxy aux individus de ces coûts via la sensation subjective d'effort et qui leur permet de limiter leurs dépenses en énergie. Cette sensation n'est cependant pas une lecture directe et parfaite des coûts engagés dans

## Introduction

l'activité physique, mais bien un proxy qui les reflète sans en donner une image exacte. Ce processus est soumis à des failles puisqu'après une certaine durée et intensité d'effort, le corps libère des endorphines dans le cerveau qui peuvent rendre certaines personnes dépendantes de la sensation ressentie en exerçant un effort qui devient alors appétitif au lieu d'être aversif, du moins dans les conditions où il amène à la libération de ces endorphines. Dans le cerveau, l'aire supplémentaire motrice pourrait être une zone dont l'un des rôles est de mesurer ce proxy des coûts engagés par l'effort et qui le rendrait aversif.

Une récompense paraît d'autant moins intéressante qu'elle est liée à une action demandant une activité physique importante en durée ou en intensité ce qui témoigne de cette aversion pour l'effort. Le lien entre valeurs subjectives et effort n'est cependant pas évident puisque d'un côté le réseau des valeurs est impliqué dans l'évaluation des récompenses décomptées par les efforts, et de l'autre il incite les systèmes moteurs à effectuer l'effort quand la récompense en vaut la peine, en réduisant ainsi la pénibilité subjective. Différentes composantes du réseau des valeurs agissent peut-être séparément sur ces deux aspects. En effet, seul le vmPFC, et pas le striatum ventral, a été identifié comme corrélant avec la valeur subjective décomptée par les efforts (Arulpragasam et al., 2018; Seaman et al., 2018), alors qu'à l'inverse c'est surtout le striatum ventral, et les ganglions de la base plus généralement, qui sont identifiés comme un moteur de l'action physique (Pessiglione et al., 2007; Schmidt et al., 2008, 2012). La pénibilité estimée de l'activité physique, c'est-à-dire l'effort prospectif, est probablement basé sur les expériences subjectives passées. Il reste cependant à voir où cette estimation de l'effort prospectif est encodée, si elle est réellement encodée indépendamment de la valeur subjective.

## II.4 L'effort mental

### II.4.a L'activité mentale est subjectivement coûteuse

« Περιστώμεθα δὴ καθάπερ ἀθληταὶ πρὸς τοῦτον αὖ τὸν λόγον. »<sup>21</sup>

41b, *Philèbe*, Platon, (Platon, 2018)

Comme dans l'extrait ci-dessus du *Philèbe* de Platon, où Socrate propose d'aborder un problème comme des athlètes, énormément de langues, depuis plusieurs millénaires comme nous pouvons le voir avec cet extrait, utilisent le vocabulaire dévoué à l'effort physique pour parler d'un effort de nature mentale. Guillaume Ferrero, dans son article *L'inertie mentale et la*

---

<sup>21</sup> « Dressons-nous donc, comme des athlètes, pour enserrer ce nouveau problème. »

## Introduction

*loi du moindre effort*, essaye ainsi de montrer que dans le domaine mental, tout comme dans le domaine physique, les individus cherchent à minimiser une certaine forme de travail mental. En résumé, la **loi du moindre effort** qui a été identifiée dans le champ de l'activité physique s'appliquerait aussi au champ de l'activité mentale. Ferrero dit « *Tout effort mental répugne instinctivement à l'homme. [...] lorsque les nécessités de l'existence le contraignent à travailler avec le cerveau, l'homme cherche toujours à accomplir l'effort le plus petit, à employer les processus psychologiques qui lui coûtent la moindre fatigue et qui pour cela ne sont pas douloureux.* » (Ferrero, 1894). En linguistique, l'idée que les individus minimisent une certaine forme d'effort dans le langage va prendre une place importante. Dès 1897, dans son ouvrage *Essai de Sémantique*, Michel Bréal, considéré comme le fondateur de la sémantique moderne, écrit : « *Il y a là une tendance générale du langage, qui est de se faire comprendre aux moindres frais, je veux dire avec le moins de peine possible.* » (Bréal, 1897). Cependant, il ne développe pas plus en avant cette réflexion. Une cinquantaine d'années plus tard, George Zipf développera une réflexion plus importante autour de l'idée d'un coût que les individus chercheraient à minimiser dans le langage. Il proposera ainsi l'idée d'un principe du moindre effort en linguistique (Zipf, 1949). Cette idée sera reprise plus tard par le linguiste André Martinet dans son principe d'économie en linguistique (Martinet, 1955). André Martinet écrit : « *L'évolution linguistique en général peut être conçue comme régie par l'antinomie permanente des besoins communicatifs et expressifs de l'homme et de sa tendance à réduire au minimum son activité mentale et physique. Sur le plan des mots et des signes, chaque communauté linguistique trouve à chaque instant un équilibre entre les besoins d'expression qui demandent des unités plus nombreuses, plus spécifiques et proportionnellement moins fréquentes, et l'inertie naturelle qui pousse vers un nombre plus restreint d'unités plus générales et d'emploi plus fréquent.* » (Martinet, 1955). Néanmoins, le langage passe par des actions musculaires en plus d'une activité mentale, ce qui rend la distinction entre la part dévolue à l'effort physique et à l'effort mental difficiles à discerner.

Plus tard, l'idée que les individus cherchent à minimiser leur activité mentale va trouver de nombreuses preuves dans différents domaines de l'activité mentale. Des tâches de choix ont permis de montrer que les individus préfèrent des tâches qui requièrent moins de mémoire de travail (Benoit et al., 2019; Kool et al., 2010; McGuire and Botvinick, 2010; Westbrook et al., 2013), des calculs mentaux moins complexes (Choe et al., 2019; Nagase et al., 2018), des projections spatiales plus simples (Nagase et al., 2018), moins d'alternance entre différentes tâches à effectuer (*task switching*) (Apps et al., 2015; Kool et al., 2010; Monsell, 2003) et moins

## Introduction

de focus attentionnel (Apps et al., 2015). **Certains types d'activités mentales sont donc perçus comme coûteux.** L'ensemble de ces activités mentales qui sont subjectivement perçues comme coûteuses a été regroupé dans ce qu'on appelle les **fonctions exécutives**. Les fonctions exécutives se divisent en trois principaux groupes (Diamond, 2013):

- **l'inhibition** que ce soit pour inhiber un comportement spontané ou pour diriger son attention et ne pas se laisser distraire ;
- **la mémoire de travail ;**
- **la flexibilité mentale.**

Subjectivement, ces fonctions sont associées à un coût : la sensation aversive d'effort mental. Cependant, il a été montré que posséder des capacités plus élevées dans l'ensemble de ces fonctions est associé à une meilleure qualité de vie sur le plan tant physique (Miller et al., 2011; Riggs et al., 2010, 2012; Will Crescioni et al., 2011) que mental (Brown and Landgraf, 2010). Ces capacités sont aussi associées à un meilleur taux de réussite au niveau social et à un meilleur développement psychologique, cognitif et social (Diamond, 2013). Il faut noter dès à présent que la notion de **contrôle cognitif** qui a amplement été développée par d'autres chercheurs est relativement proche de la notion de fonction exécutive (Botvinick et al., 2001; MacDonald et al., 2000). Elle est liée à l'idée d'un mécanisme qui permet de réguler les réactions automatiques ou de mettre en place une réaction comportementale adaptée lorsque aucun automatisme préexistant ne permet de réagir face à une situation donnée. Dans le même ordre d'idée, Daniel Kahneman, reprenant des termes introduits par Keith Stanovich et Richard West en psychologie (Stanovich and West, 2000), développera l'idée qu'il existe deux systèmes dans le cerveau dénommés simplement le Système 1 et le Système 2 (Kahneman, 2012). Pour lui, le Système 1 est un mode de pensée lié justement à l'implémentation d'automatismes, souvent soumis à des biais psychologiques. A l'inverse, le Système 2 est un système qui permettrait de corriger ces biais et de réguler les automatismes du Système 1. Par la suite, si nous utilisons les termes d'effort mental, de fonctions exécutives, de contrôle cognitif ou de Système 2, il faudra les comprendre comme synonymes de l'idée d'un mécanisme lié à la sensation aversive d'effort mental en vue d'implémenter une régulation du comportement, en particulier des réactions automatiques. Il faut cependant bien sûr noter que, rigoureusement parlant, ces différentes notions, bien qu'elles se superposent amplement sur certains aspects, ne renvoient pas exactement aux mêmes concepts.

## Introduction

Il est intéressant de constater que l'impact de l'effort mental, ou plutôt de la difficulté en l'occurrence, sur la valeur subjective d'une récompense ne répond pas à la même fonction que celle liée à l'effort physique. En effet, Chong *et al.* ont montré que l'impact de l'effort mental sur la valeur subjective était mieux décrit par une fonction de type hyperbolique comme c'était le cas pour le délai (voir *L'aversion pour le délai*) :  $SV = \frac{R}{1+k \cdot E}$  où  $SV$  est la valeur subjective,  $R$  le montant de récompense,  $E$  la quantité d'effort mental qui dans ce cas correspond à la difficulté de la tâche et  $k$  le paramètre subjectif (Chong et al., 2017). Cependant, comme dans le cas du délai (voir *L'aversion pour le délai*), cette fonction hyperbolique présente une limite essentielle : elle suppose que l'implémentation d'un effort mental, même s'il est élevé, vaut mieux que de ne rien faire. En effet, elle n'autorise pas la  $SV$  à être inférieure ou égale à zéro, c'est-à-dire qu'elle implique que, à partir du moment où il existe une récompense à obtenir cela vaut nécessairement le coup d'agir.

Il faut noter cependant que, comme dans le cas de l'activité physique, l'activité mentale semble parfois être perçue comme quelque chose de positif. Certains individus ressentiraient ainsi un besoin interne d'employer leurs fonctions exécutives. Ils pourraient même se sentir frustrés et/ou s'ennuyer s'ils n'ont pas l'occasion d'exercer une activité mentale nécessitant du contrôle cognitif (Inzlicht et al., 2018). Ce « **besoin de cognition** » (*need for cognition*) a été identifié en 1955 par Arthur Cohen qui a identifié des différences interindividuelles dans ce besoin (Cacioppo and Petty, 1982; Cohen, 1955). Ce besoin pourrait être dû à plusieurs causes. Par exemple les individus qui vont jouer à des jeux nécessitant des fonctions exécutives (échecs, jeux de société, sudoku, etc.) cherchent aussi une gratification liée à leur performance. Ils tireraient ainsi du plaisir d'un côté à observer leur niveau de progression individuelle au cours du temps et, de l'autre, à observer leur classement relatif aux autres joueurs. Le plaisir tiré de l'activité mentale serait ainsi non pas lié à l'activité mentale, en tant que telle, mais à la récompense induite par l'évolution de la performance individuelle et/ou au sein d'un groupe. De plus, le contrôle sur soi est, depuis des siècles, grâce au travail des philosophes à travers le monde notamment, perçu comme quelque chose qui apporte des bienfaits. Aujourd'hui, les liens entre santé physique et mentale et utilisation des fonctions exécutives sont connus (Diamond, 2013). Tout un business économique s'est d'ailleurs développé autour de l'idée d'applications ludiques censées améliorer la résistance aux pathologies neurodégénératives en poussant l'utilisateur à employer du contrôle cognitif. De même que l'activité physique est valorisée au niveau sociétal pour l'amélioration de la qualité de vie qu'elle apporterait, l'activité mentale est valorisée dans nos sociétés pour les mêmes raisons.

## Introduction

Une autre hypothèse par rapport à la valeur positive qui est attribuée à l'activité mentale propose qu'au cours de la vie les individus aient appris à associer, selon leur vécu, un certain niveau d'effort à un certain niveau de récompense obtenu. Ce procédé est nommé par l'expression « *learned industriousness* » en anglais (Eisenberger, 1992; Inzlicht et al., 2018; Kool and Botvinick, 2018). Comme dans le conditionnement opérant (voir *Psychologie et physiologie*), les individus associeraient alors l'exécution d'un certain niveau d'activité mentale à l'obtention d'un certain niveau de récompense, même quand cette association n'est pas pertinente. En résumé, l'activité mentale, tout comme l'activité physique, peut être associée à une valeur positive en fonction des individus et du contexte, malgré la sensation aversive d'effort qui lui est associée.

### II.4.b De quoi l'aversion pour l'activité mentale nous protège-t-elle ?

*« les plaisirs tirés d'activités qui ne sont pas celles qui nous occupent font obstacle aux activités qui nous occupent. En effet, les amateurs de flûte sont incapables de prêter attention aux arguments qu'on leur présente s'ils entendent, serait-ce faiblement, un joueur de flûte, parce qu'ils ont plus de joie à cet art qu'à la présente activité. Le plaisir correspondant au jeu de flûte détruit donc l'activité que suppose l'argumentation. Et une chose pareille arrive encore, quelles que soient les circonstances, chaque fois qu'on exerce en même temps deux activités. La plus agréable, en effet, tend à chasser l'autre et si son plaisir l'emporte de beaucoup, c'est d'autant plus vrai, au point qu'il n'est plus possible de maintenir l'autre activité. C'est pourquoi si une chose, quelle qu'elle soit, nous plaît fort, nous ne pouvons pas vraiment en faire une autre et si nous en faisons d'autres, c'est que la première nous séduit médiocrement, exactement comme sont dans les théâtres ceux qui mangent des friandises : c'est quand les compétiteurs sont mauvais qu'ils se livrent surtout à cette occupation. »*

3.3.1.2, Livre X, *Éthique à Nicomaque*, Aristote (Aristote, 2004)

Dans le domaine physique, il est clair que le corps a des contraintes matérielles liées à la nature des muscles : l'activité musculaire consomme des ressources énergétiques (ATP, glycogène, lipides) et elle génère des toxines (lactate). Un effort trop intense pendant une durée prolongée peut provoquer des lésions tissulaires. Il est donc important que le corps s'en prémunisse. Une solution que l'évolution a trouvée pour cela est de se représenter un proxy du coût métabolique engendré par l'activité physique via la sensation aversive d'effort. On peut de plus mesurer la contraction musculaire directement et voir comment celle-ci est corrélée à une évaluation subjective d'un effort physique par le sujet. Pour ce qui est de l'activité mentale par

## Introduction

contre, il est bien plus dur de savoir quelle serait la ressource consommée ou les déchets produits lorsque les fonctions exécutives sont engagées. Il semble pourtant qu'il y ait bien une ressource en jeu puisque des tâches requérant l'utilisation de fonctions exécutives, comme la mémoire de travail, pendant un temps long induisent une baisse objective de la performance (Blain et al., 2016; Gergelyfi et al., 2015; Lorist et al., 2000) et **une sensation subjective de fatigue mentale** (Gergelyfi et al., 2015; van der Linden et al., 2003).

Par le passé, cette idée avait mené à l'apparition de l'idée d'épuisement du moi (*ego depletion*) selon laquelle un des aspects des fonctions exécutives, l'inhibition, en particulier l'inhibition des comportements spontanés, aussi appelée la maîtrise de soi (*self control*), repose sur un réservoir très limité de ressources mentales et qu'une fois que ce réservoir est vidé, il est impossible d'effectuer à nouveau du contrôle sur soi tant que la ressource en question n'a pas été restaurée (Baumeister et al., 1998). Les tenants de cette théorie avaient imaginé que le glucose sanguin aurait pu être cette ressource nécessaire à l'exécution de l'inhibition des comportements automatiques (Gailliot and Baumeister, 2007; Gailliot et al., 2007), mais il a depuis été montré que cela était surtout lié à un effet psychologique sur les participants qui ingéraient du glucose plus qu'un effet physiologique du glucose en tant que tel (Job et al., 2013; Kurzban, 2010; Lange and Eggert, 2014; Molden et al., 2012). La théorie de l'épuisement de soi est globalement en désuétude car une méta-analyse a prouvé que l'effet en question n'était probablement pas significatif une fois corrigé pour le biais de publication de résultats positifs (Carter and McCullough, 2014). Sur des temps plus longs néanmoins, il semble qu'il existe réellement un phénomène de fatigue mentale reflété subjectivement et objectivement par une baisse en performance comme on l'a vu précédemment.

Plusieurs possibilités ont été évoquées pour déterminer quelle pourrait être la ressource requise par les fonctions exécutives (Kool and Botvinick, 2018; Shenhav et al., 2017; Zénon et al., 2019), nous allons en lister les principales ci-dessous :

- 1) **ressource énergétique** : c'est l'idée que les tenants de la théorie de l'épuisement de soi avaient en tête quand ils pensaient avoir identifié le glucose comme la ressource déterminant la capacité à exercer du contrôle sur soi. Si cette théorie a depuis été invalidée comme on l'a vu, il n'est pas impossible que d'autres ressources métaboliques interviennent. En effet si l'apport énergétique global au cerveau est constant, la répartition du sang et des ressources entre les différentes zones du cerveau varie au cours du temps (c'est le principe même de l'effet BOLD qu'on observe en IRMf) et peut-être que le cerveau a besoin de réguler l'apport reçu par une zone donnée au-delà d'une

certaines durée. Le glycogène stocké dans les neurones a aussi été proposé comme une ressource éventuelle qui pourrait être limitée (Christie and Schrater, 2015) ;

- 2) **production de déchet(s) à nettoyer** : on sait désormais le rôle primordial de l'accumulation de certaines protéines dans certaines dégénérescences cérébrales, comme par exemple l'accumulation de peptide amyloïde beta dans l'hippocampe chez les patients atteints de la maladie d'Alzheimer. Il a été proposé que, quand les fonctions exécutives sont utilisées, du peptide amyloïde beta s'accumule dans l'espace extracellulaire et qu'à partir d'un certain stade, il faille le « nettoyer » afin d'éviter de provoquer des lésions des neurones (Holroyd, 2015) ;
- 3) **coût d'opportunité** : la sensation aversive liée à l'utilisation des fonctions exécutives serait liée à une forme de pression du cerveau pour pousser l'individu à agir en vue de récompenses à court terme plutôt qu'à long terme (Inzlicht et al., 2014; Kurzban, 2016; Kurzban et al., 2013). L'idée étant que les actions qui engagent les fonctions exécutives sont souvent rentables sur le long terme mais pas sur le court terme. Plus le temps à employer du contrôle cognitif sans obtenir de récompense passe, plus la sensation aversive de coût augmenterait afin de pousser l'individu à s'investir dans une autre activité qui serait plus rentable à court terme. Dans ce cadre, une activité mentale liée à l'utilisation des fonctions exécutives paraît d'autant plus subjectivement coûteuse que le taux d'obtention de récompenses dans l'environnement sans avoir besoin d'utiliser les fonctions exécutives est élevé ;
- 4) **coût de l'information** : le coût subjectif de l'activité mentale dépendrait de la différence entre les *a priori* du sujet sur le monde et les *a posteriori* du sujet après avoir mis à jour ses *a priori* (Kolling et al., 2016a; O'Reilly et al., 2013; Zénon, 2019; Zénon et al., 2019). On appelle cette différence la divergence de Kullback-Leibler et, en l'occurrence, cela reviendrait à dire que le coût lié à l'effort mental est lié la quantité d'information déviant des *a priori* que les neurones doivent encoder ;
- 5) **sous-produit de la contrainte à avoir un comportement automatique optimal** : une étude en intelligence artificielle a montré que pour avoir un système qui optimise ses comportements par défaut dans différentes situations afin d'avoir un comportement autant que possible automatisé, il était nécessaire d'attacher un coût aux processus « *top-down* » qui cherchent à réadapter le comportement sans cesse plutôt que de se baser sur des comportements stables optimaux à long terme (Kool and Botvinick, 2018; Teh et al., 2017). Ceci, bien que très théorique, peut faire sens en ce qu'il a été montré que, tôt



## Introduction

dans une nouvelle tâche, recruter les fonctions exécutives est positivement corrélé à la performance. Cependant, c'est l'inverse quand on regarde plus tard dans la tâche, comme si ceux qui n'avaient pas réussi à automatiser la tâche avec un comportement par défaut optimal n'avaient pas une très bonne performance (Diamond, 2013).

La nature de la ressource qui est utilisée dans l'effort mental est donc encore loin d'être élucidée, mais toutes ces hypothèses donnent de la matière pour de futures études qui permettront peut-être enfin de savoir pourquoi l'activité mentale paraît subjectivement coûteuse. Il faut, au passage, noter que la plupart de ces hypothèses ne sont pas incompatibles les unes avec les autres. Si l'on pense à l'effort physique sur lequel l'idée de ressource pour l'effort mental a été établie par analogie, la contraction des muscles passe autant par une production de déchets métaboliques (le lactate) que par une consommation des ressources énergétiques (réserves de glycogène, ATP, etc.) et, de plus, la mobilisation d'un muscle dans une action donnée (écrire sur son ordinateur par exemple) l'empêche d'effectuer d'autres actions qui pourraient être plus gratifiantes sur le court terme (saisir un biscuit par exemple).

L'une des difficultés principales dans l'étude de l'effort mental, en comparaison à l'effort physique notamment, est l'impossibilité de mesurer directement la variable dépensée dont il est le proxy puisque, comme nous venons de le voir, on ne connaît pas, pour l'instant, la ressource sur laquelle il est basé. Wouter Kool et Matthew Botvinick appellent cela le **problème économétrique** dans la recherche sur l'effort mental (Kool and Botvinick, 2018). En effet, dans le domaine physique, on peut mesurer la contraction musculaire ou la sécrétion de lactate par exemple, pour voir dans quelle mesure l'effort subjectif varie par rapport à ces variables observables. Dans le domaine de l'activité mentale par contre, à part en mesurant la **performance** ou en faisant varier la **difficulté**, il est dur de voir quelle variable objective mesurer pour voir dans quelle mesure la sensation subjective d'effort varie ou non avec elle. Plusieurs méthodes ont cependant été établies au cours de ces dernières années pour tenter d'en mesurer des proxy : la **dilatation de la pupille** (Alnaes et al., 2014; Bijleveld, 2018; McCloy et al., 2017; Peysakhovich et al., 2015, 2017; Takeuchi et al., 2011; van der Wel and van Steenbergen, 2018; Zénon, 2019), les **temps de réaction** (Alnaes et al., 2014; Ferguson et al., 2013; Haji et al., 2015; Kamiński et al., 2017; Pals et al., 2015; Robinson et al., 1997; Takeuchi et al., 2011; Vassena et al., 2014), des **reports subjectifs de la pénibilité de l'effort** effectué par les sujets (Haji et al., 2015; Robinson et al., 1997; Vassena et al., 2014) ou encore des **choix** impliquant différents degrés d'efforts mentaux (Apps et al., 2015; Benoit et al., 2019; Kool et al., 2010; McGuire and Botvinick, 2010; Westbrook et al., 2013). C'est sur la base de ces

## Introduction

mesures et de variations dans la difficulté qu'est généralement mesuré l'effort mental et que des zones cérébrales liées à l'effort mental ont pu être identifiées bien que toutes ces mesures présentent des limites en ce qu'il s'agit de proxy qui varient aussi avec d'autres variables et non pas de l'effort mental en tant que tel (Shenhav et al., 2017).

### II.4.a Bases neurales de l'effort mental

Les recherches en neuroimagerie ont permis d'identifier un réseau d'aires principalement constitué du **cortex dorsolatéral préfrontal (dlPFC)** bilatéralement, d'un cluster bilatéral dans le **cortex pariétal** et d'un cluster dans le **dmPFC** qui serviraient de supports neuraux aux fonctions exécutives (voir **Figure 16** et (Niendam et al., 2012)). Le rôle de chacune de ces aires est encore fortement débattu mais une idée de plus en plus répandue est que **le dlPFC** servirait à la **sélection des réponses/stratégies appropriées** en fonction de la tâche en cours, **l'inhibition des stratégies non-appropriées** et au **maintien des règles de la tâche en cours**, alors que les régions pariétales traiteraient l'information relative à la saillance des items et les associations entre stimulus et réponse et que **le dmPFC** servirait à signaler les **conditions dans lesquelles il est nécessaire d'exercer plus d'effort mental** (Botvinick et al., 1999, 2004; Collette and Salmon, 2014; Duverne and Koechlin, 2017; Koechlin et al., 2003; MacDonald et al., 2000; Niendam et al., 2012).

L'activité du **dlPFC** en particulier a en effet été corrélée avec **l'inhibition** (van Belle et al., 2014; Wager et al., 2005), **la mémoire de travail** (Barch et al., 1997; Braver et al., 1997; Cohen et al., 1997; Courtney et al., 1997; Rottschy et al., 2012) et **la flexibilité mentale** (Ravizza and Carter, 2008), c'est-à-dire les trois principales fonctions exécutives énoncées précédemment (voir *L'activité mentale est subjectivement coûteuse*). Le dlPFC semble bien être au cœur de ces trois grandes catégories des fonctions exécutives (voir **Figure 16**). Il a d'ailleurs été observé que l'activité du cortex dorsolatéral préfrontal diminue avec la fatigue (Blain et al., 2016). C'est donc l'activité de cette zone, ou du moins de ce réseau, qui semble déterminante dans la sensation subjective d'effort mental. Une possibilité serait donc que les neurones dans le dlPFC épuisent une ressource ou au contraire accumulent des déchets quand ils sont utilisés pour une durée prolongée. La sensation subjective d'effort mental existerait ainsi en vue d'éviter que le dlPFC soit actif trop intensément trop longtemps. Il faut préciser cependant que, comme pour l'effort physique, la sensation subjective d'effort pourrait ne pas être générée directement par l'activité de cette zone mais par la lecture de cette activité par une autre zone, comme c'était le cas pour l'effort physique, par exemple le dmPFC ou la SMA. Notez que, là aussi, la nomenclature peut varier selon les auteurs et que certains appellent aussi cette zone le

## Introduction

gyrus frontal moyen sur la base de l'anatomie (Blain et al., 2016). Il s'agit pourtant bien de ce que la plupart des auteurs désigne comme le cortex dorsolatéral préfrontal.



**Figure 16 : Méta-analyse montrant les zones liées aux fonctions exécutives sur la base de 193 études chez l'adulte sain.** En rouge, les zones activées par des fonctions exécutives à travers les études et en vert les clusters correspondants à 9 études où une conjonction a été effectuée entre les trois sous-catégories des fonctions exécutives (flexibilité, inhibition et mémoire de travail). On distingue un cluster bilatéral au niveau du cortex dorsolatéral préfrontal, un autre cluster bilatéral au niveau du cortex pariétal et enfin un cluster localisé au niveau du dmPFC. Figure issue de (Niendam et al., 2012).

Dès 1997, l'équipe de Jonathan Cohen dissocie le rôle du dmPFC de celui du dlPFC. Ils montrent que l'activité du dmPFC corrèle avec la difficulté d'une tâche de mémoire de travail, la tâche AX-CPT. La difficulté était manipulée en altérant la lisibilité des stimuli dans la tâche : 85% des pixels étaient retirés aléatoirement dans les essais difficiles. Le dlPFC, lui, était plus actif dans les essais où l'information devait être maintenue en mémoire plus longtemps, du fait de délais plus longs dans la tâche (Barch et al., 1997). Par la suite, de la fin des années 1990 au début des années 2000, l'équipe de **Jonathan Cohen** principalement mais aussi d'autres équipes vont progressivement établir que l'activité du dmPFC augmente en réponse à des situations où il existe un conflit interne sur quelle est la meilleure décision/action à effectuer (Botvinick et al., 1999, 2001, 2004; Carter et al., 2000; Casey et al., 2000; Durston, 2003; Kerns

## Introduction

et al., 2004; MacDonald et al., 2000). Ce travail va amener Jonathan Cohen à proposer, avec ses successeurs Matthew Botvinick et Amitai Shenhav, une théorie du dmPFC où ils proposent que le dmPFC encode **la valeur attendue d'utiliser du contrôle cognitif** (*expected value of control*) selon la formule :

$$EVC = P * R - C$$

où *EVC* est la valeur attendue d'utiliser les fonctions exécutives (le dIPFC), *R* la récompense espérée, *P* la probabilité estimée d'obtenir la récompense en question et *C* le coût d'utiliser les fonctions exécutives (Shenhav et al., 2013). L'idée simple de cette théorie est que le dmPFC calculerait la valeur nette d'employer les fonctions exécutives en estimant d'un côté les bénéfices d'utiliser une fonction exécutive en termes de récompense pour l'organisme et de l'autre le coût neural d'utiliser cette ressource. En fonction du résultat de ce calcul, il enverrait un signal pour recruter le dIPFC et utiliser ses capacités de contrôle cognitif proportionnellement au besoin de la situation en cours (Shenhav et al., 2013, 2016a). Cette théorie reste encore très controversée (voir *Hypothèses alternatives*) mais elle a l'avantage de permettre de comprendre plusieurs résultats dans un schéma théorique global et de faire d'éventuelles prédictions sur le rôle du dmPFC. Il est d'ailleurs généralement admis que le dmPFC puisse envoyer un signal au dIPFC pour le recruter (Kerns et al., 2004), du fait notamment de leur connectivité fonctionnelle (Jin et al., 2018; Margulies et al., 2007; Zarr and Brown, 2016), mais la nature du (ou des) calcul(s) effectué par le dmPFC est controversée (Kolling et al., 2016b). Le dmPFC corrèle, en outre, significativement avec les différents indicateurs d'effort mental évoqués plus haut, comme les temps de réaction (Kamiński et al., 2017; McGuire and Botvinick, 2010) ou le diamètre de la pupille au repos (Murphy et al., 2014; Schneider et al., 2016) et les changements de la pupille pendant une tâche (Schneider et al., 2018). Ces résultats suggèrent qu'il pourrait être le siège de la perception subjective de l'effort mental. De plus, en 2005, Naccache *et al.* ont montré, chez une patiente atteinte d'une lésion unilatérale du côté gauche incluant le dmPFC, une dissociation entre la capacité à utiliser une fonction exécutive et la sensation subjective d'effort mental. Elle était en effet capable de réaliser une tâche de Stroop qui requiert une fonction exécutive, l'inhibition d'une réponse spontanée dans le cas des essais incongruents, aussi bien que des sujets contrôles. Il semblerait d'ailleurs que son dmPFC droit se soit bien activé en réponse aux essais incongruents. Par contre, elle était incapable de discriminer les essais incongruents des essais congruents en termes d'effort mental ressenti. Contrairement aux sujets sains, la patiente n'avait pas l'impression d'avoir exercé plus d'effort pour les essais incongruents (Naccache et al., 2005).

## Introduction

Cette expérience nous éclaire au moins sur un point : il est possible de dissocier l'exécution d'une activité nécessitant du contrôle cognitif de la sensation subjective d'effort mental qui l'accompagne habituellement. En effet, le comportement était équivalent à celui de sujets contrôles en termes de performance, mais n'allait plus de pair avec une sensation subjective d'effort. De plus, le dlPFC, classiquement identifié comme le siège de l'exécution des fonctions exécutives, semble avoir été préservé par la lésion. Il est ainsi possible que ce soit une autre zone, comme le dmPFC, qui se charge de surveiller et de réguler le niveau d'activité du dlPFC en fonction de la quantité de ressources physiologiques et/ou de déchets qui pourraient s'y être accumulés.

### II.4.b Réseau des valeurs et effort mental

« *Jen K'iou dit : « Maître, ce n'est pas que votre doctrine me déplaît ; mais je n'ai pas la force de la mettre en pratique. » Le Maître répondit : « Celui qui vraiment n'a pas assez de forces, tombe épuisé à moitié route. Pour vous, il n'en est pas ainsi, vous vous prescrivez des limites que vous ne voulez pas dépasser ; ce n'est pas la force, mais la volonté qui vous manque. » »*

*Ioung Ie, Louen Yu, Confucius (Confucius and Lao-Tseu, 1973)*

A ma connaissance, bien que plusieurs études aient démontré l'impact d'un effort mental prospectif sur la valeur subjective au niveau comportemental comme on l'a vu précédemment (voir *L'activité mentale est subjectivement coûteuse*), peu ont étudié comment l'activité du réseau des valeurs était impactée par un effort mental prospectif. Une étude révèle que le striatum ventral voit son activité diminuer au moment de l'obtention d'une récompense si un effort mental plus élevé a dû être effectué pour obtenir un niveau de récompense équivalent (Botvinick et al., 2009), mais il s'agit là d'une évaluation rétrospective et non pas d'un compromis coûts/bénéfices prospectif. De futures études nous diront si l'activité du réseau des valeurs est bien décomptée par l'anticipation d'un effort mental.

De plus, comme dans le cas de l'effort physique, **le réseau des valeurs semble capable de déclencher l'utilisation de ce réseau exécutif quand son utilisation permet d'augmenter la récompense en jeu** (Locke and Braver, 2008; Schmidt et al., 2012). Cela a notamment été montré dans le cas de la mémoire de travail qui peut être améliorée via des projections des ganglions de la base vers le cortex préfrontal (Chatham and Badre, 2015). On en trouve aussi une preuve dans le cas de l'inhibition dans une tâche de Stroop numérique où la performance semble dépendre de l'activité du striatum ventral (Schmidt et al., 2012). Si l'on accepte que le

## Introduction

dmPFC calcule l'équation de l'EVC proposée par Jonathan Cohen, Matthew Botvinick et Amitai Shenhav, cela reviendrait juste à dire que le réseau des valeurs fournit l'information de la valeur  $P * R$  de l'équation au dmPFC qui verrait donc son activité, et ainsi, la probabilité de solliciter le dlPFC, augmentées. Ceci est d'ailleurs cohérent avec une vision motivationnelle de la fatigue (Hockey, 2011; Inzlicht et al., 2014; Kurzban, 2016; Kurzban et al., 2013; Müller and Apps, 2019), comme celle proposée par la théorie du coût de l'opportunité. En effet, l'idée principale de ces différentes théories motivationnelles de la fatigue est que plus le temps passé à exercer un effort mental augmente, plus la perception du coût de cet effort augmente. Cela serait dû non pas au fait que l'effort mental serait limité en termes de ressources physiologiques, mais au fait que la ressource serait le temps. Un temps long passé à utiliser ces ressources cognitives sans obtenir de récompense conséquente serait perçu comme un coût subjectif croissant. Tout cela prédit que proposer des récompenses plus élevées permettrait d'atténuer, voire de faire disparaître, l'effet de la fatigue sur la performance. A l'époque où les théories sur l'« épuisement de soi » foisonnaient, une étude avait montré qu'un paiement plus élevé pouvait contrecarrer les effets de la « déplétion du moi » (Muraven and Slessareva, 2003) ce qui irait dans ce sens. Cependant, étant donné la durée des manipulations censées provoquer cette déplétion (5 minutes en essayant de retenir de montrer ses émotions pour l'une des expériences de l'article par exemple) et les critiques qui ont émergées depuis et que l'on a déjà présentées, il est difficile de parler d'un effet de fatigue mentale qui serait compensé par une motivation augmentée. De plus, une étude a essayé d'observer cet effet après avoir induit de la fatigue cognitive en demandant aux sujets de résoudre des sudokus pendant 2h. Aucun impact des récompenses plus élevées sur la performance des sujets n'a été observé une fois qu'ils étaient fatigués (Gergelyfi et al., 2015). Ce résultat infirme, au moins en partie, l'idée que la ressource des fonctions exécutives soit purement liée à une représentation motivationnelle de la tâche. De plus, là encore, le biais de publication pourrait masquer d'autres résultats invalidant ces théories motivationnelles de la fatigue. Cependant, d'autres études sont nécessaires pour y voir plus clair.

## II.5 Les punitions

### II.5.a Qu'est-ce qu'une punition ?

*« lorsque nous disons que le plaisir est la fin, nous ne voulons pas parler des « plaisirs des fêtards » ni des « plaisirs qui se trouvent dans la jouissance », comme le croient certains qui, par ignorance, sont en désaccord avec nous ou font à nos propos un mauvais accueil, mais de l'absence de douleur en son corps, et de trouble en son âme. »*

Une définition simple d'une punition est de considérer comme une punition tout processus qui réduit la probabilité de réexercer une action donnée (Rolls, 2018), c'est-à-dire tout évènement extérieur qui va être spontanément évité. On peut en distinguer au moins deux types : les punitions d'ordre **physique** qui impliquent une **douleur**, et celles d'ordre **psychologique**. La douleur correspond à une sensation générée par les fibres nociceptives qui remonte par la moelle épinière jusqu'au cerveau. Les punitions d'ordre psychologique correspondent principalement à une **frustration** liée à la survenue d'un évènement non-souhaité, comme une perte à un jeu d'argent ou le fait de ne pas avoir obtenu une récompense désirée par exemple. De plus, il faut probablement distinguer les punitions de par leur intensité. Le sentiment de **peur** en anticipation d'une punition ne naît que face à une punition d'intensité élevée comme un électrochoc chez le rongeur par exemple. Des punitions d'intensité moindre, comme un souffle d'air sur l'œil, bien qu'avérsives et éventuellement aussi stressantes, ne provoqueront par contre pas de sensation de peur (Izquierdo et al., 2016). Une littérature extensive a étudié les bases comportementales et neurales de la peur chez le rongeur, via notamment l'association pavlovienne entre certains lieux (stimulus conditionné) et le fait de subir un électrochoc (stimulus non-conditionné), c'est-à-dire une douleur physique. Les rongeurs vont, s'ils en ont la possibilité, éviter les endroits liés à ce genre d'expérience douloureuse. Cet évitement se fera soit par **évitement passif** (*passive avoidance*), soit via un **évitement actif** (*active avoidance*). L'évitement passif consiste à éviter de se diriger dans la direction où les rongeurs savent qu'ils risquent de subir un électrochoc. L'évitement actif consiste à se déplacer activement hors de la zone où ils savent qu'ils vont subir un électrochoc s'ils ne bougent pas. Quand les rongeurs ne peuvent pas éviter l'électrochoc, ils restent sur place et montrent une attitude typique en restant immobiles (*freezing*) à l'idée de la punition qui va leur arriver (Izquierdo et al., 2016). Ces comportements d'anticipation persistent généralement même une fois que le lieu en question n'est plus associé au fait de subir un électrochoc.

Un phénomène intéressant révélé par ces études chez le rongeur est qu'un seul électrochoc peut suffire à générer ce conditionnement à la peur (Izquierdo et al., 2016), là où dans le domaine des récompenses plusieurs répétitions sont généralement nécessaires jusqu'à ce que l'association entre le stimulus conditionné et le stimulus non-conditionné soit stable. Or, nous avons déjà vu comment, d'après la théorie des perspectives, l'utilité était considérée par rapport à un point de référence et comment le même montant n'est pas perçu de la même manière selon qu'on peut le gagner ou le perdre. Les individus ont un décompte plus élevé de

## Introduction

la valeur dans le domaine des pertes que dans le domaine des gains en moyenne (voir **Figure 9**). Il semblerait donc que les individus apprennent à créer des associations plus rapidement dans le domaine des punitions que dans le domaine des gains. Ceci est d'ailleurs cohérent avec une pression évolutive pour la survie puisqu'un animal a tout intérêt à apprendre très vite à éviter des stimuli pouvant mettre sa vie en danger. En effet, une seule rencontre avec le mauvais prédateur ou un aliment toxique pourrait mettre fin à ses jours, alors que, dans le domaine des gains, à part dans des conditions extrêmes, il est généralement tolérable de rater une occasion d'obtenir une récompense.

### II.5.b Dévaluation par les punitions

*« Les pugilistes en effet trouvent agréable la fin pour laquelle ils combattent (la couronne et les honneurs). Or les coups reçus sont douloureux pour des gens faits de chair et ils leur sont pénibles, ainsi que tout le travail ardu qu'ils se donnent. »*

2.6.2, Livre IV, *Éthique à Nicomaque*, Aristote (Aristote, 2004)

Si on peut imaginer quelques exemples dans la nature où les animaux vont accepter de subir des punitions pour obtenir une récompense en affrontant physiquement le mâle dominant du groupe pour récupérer son statut par exemple ou en acceptant la frustration liée à la non-obtention d'un désir pour en obtenir un autre plus grand, ce type de situations sont globalement moins étudiées en laboratoire par rapport aux autres dévaluateurs évoqués jusqu'ici, notamment pour des raisons éthiques. Néanmoins, il a été montré que des rongeurs sont prêts à subir un électrochoc pour pouvoir auto-stimuler leurs structures du plaisir tant que le nombre d'électrochocs subis pour ce faire n'est pas trop élevé (Valenstein and Beer, 1962). Chez l'homme, deux études ont regardé comment les punitions dévaluaient les récompenses avec une tâche d'apprentissage où les individus devaient associer différents visages à l'obtention d'un choc électrique ou d'une caresse de leur main associé ou non à l'obtention d'une récompense monétaire (Park et al., 2011; Talmi et al., 2009). Les auteurs ont utilisé un modèle de la forme :  $SV = \beta_r * R^{\alpha R} + \beta_S * S^{\alpha S} + \beta_{S*r} * R^{\alpha R} * S^{\alpha S}$  où  $SV$  est la valeur subjective,  $\beta_r$  la sensibilité aux récompenses en jeu,  $R$  le montant des récompenses en jeu,  $\alpha R$  un paramètre au cas où la relation entre  $SV$  et  $R$  ne soit pas linéaire,  $\beta_S$  la sensibilité aux chocs électriques en jeu,  $S$  le niveau du choc électrique,  $\alpha S$  un paramètre au cas où la relation entre  $SV$  et  $S$  ne soit pas linéaire et  $\beta_{S*r}$  un terme d'interaction entre  $S$  et  $R$  (Park et al., 2011). Les auteurs ont pu montrer qu'un modèle non-linéaire, c'est-à-dire où  $\alpha R$  et  $\alpha S$  sont différents de 1, et surtout avec une interaction entre les punitions et les récompenses, c'est-à-dire où  $\beta_{S*r}$  est différent de zéro, expliquait mieux



## Introduction

les données issues du comportement et de la neuroimagerie (Park et al., 2011; Talmi et al., 2009). Ce résultat suggère que les punitions ne sont pas simplement équivalentes à des récompenses de valence négative qui seraient sur la même échelle. Malgré ces résultats, une autre étude plus récente où les individus devaient faire un compromis entre l'obtention de différentes sommes d'argent dévaluées par le fait d'être soumis à des chocs électriques de différentes intensités sur la main ont simplement modélisé la valeur subjective comme une combinaison linéaire des montants en jeu et de l'intensité des chocs :  $SV = \beta_0 + \beta_r * R + \beta_s * S$  où  $\beta_0$  est une constante,  $\beta_r$  la sensibilité aux récompenses en jeu,  $R$  le montant des récompenses en jeu,  $\beta_s$  la sensibilité aux chocs électriques en jeu et  $S$  le niveau du choc électrique (Shapiro and Grafton, 2020). Dans le cas des choix binaires très vastement étudiés en laboratoire, on peut aussi d'une certaine manière observer ces choix comme des cas où l'individu accepte une punition (la non-obtention de l'item non-choisi) afin d'obtenir une récompense (l'obtention de l'item choisi). Dans ce type de tâches, le réseau des valeurs est régulièrement identifié comme corrélant à la valeur de l'option choisie dévaluée par la valeur de l'option non-choisie (voir *Bases neurales du choix*).

Il faut noter que, comme pour les efforts physiques et mentaux, les punitions de l'ordre physique peuvent avoir, dans le cas du **masochisme**, une valence positive difficilement explicable d'un point de vue évolutif. Dans l'extrait du marquis de Sade cité précédemment (voir *Philosophie*), Sade propose déjà que ce qui est recherché dans le masochisme, c'est « *d'ébranler la masse de nos nerfs par le choc le plus violent possible* » (Sade, 2012). Une hypothèse pourrait être que les individus qui ont un niveau faible d'éveil au niveau basal chercheraient à atteindre un niveau optimal d'éveil, comme si l'éveil était en soi une récompense, et que, de ce fait, ils seraient prêts à prendre des risques ou à ressentir de la douleur si tant est que leur niveau d'éveil est augmenté (Hebb, 1955; Schmidt et al., 2013). Ce ne serait donc que par association avec une augmentation de l'éveil que les punitions physiques seraient alors vues comme positives. Cela n'explique par contre pas pourquoi l'éveil en lui-même serait perçu comme agréable.

Certains ont depuis aussi identifié une forme psychologique de masochisme, mais qui semble plutôt être liée à des profils pathologiques en ce que les individus qui y sont soumis sont plus sujets à la dépression et semblent moins heureux dans leur vie (Békés et al., 2018).

### II.5.c Bases neurales des punitions

Par le passé, un traitement un peu radical de la douleur chronique correspondait à effectuer une chirurgie pour retirer des bouts du cortex préfrontal, c'est l'américain Van Wagenen qui va démarrer ce type d'opérations dans les années 1940. Ces opérations cependant semblent plutôt rendre les sujets indifférents à la douleur ressentie qu'à supprimer vraiment la douleur, sans compter les altérations de personnalité qu'elles entraînent. Dans les années 1950-1960, grâce au docteur Le Beau, ces opérations vont petit à petit se restreindre au cortex cingulaire, c'est ce que l'on appelle une cingulotomie (Lévêque, 2016). Depuis les années 1970-80, ce type de méthodes pour réduire la douleur semble abandonné ou en tout cas très peu utilisé, mais cela nous éclaire sur le rôle de cette zone dans la douleur. Les aires cérébrales de la sensation de douleur physique ont depuis été regroupées dans ce qui est communément appelé la « **matrice de la douleur** » (*pain matrix*). Cette matrice est principalement constituée du **cortex cingulaire**, de l'**insula** et du **thalamus**, mais aussi d'autres aires que nous ne listerons pas ici (voir <https://neurosynth.org/analyses/terms/pain/>, **Figure 17** et (Craig et al., 1996; Hayes and Northoff, 2012; Krishnan et al., 2016; Wager et al., 2013)).

L'activité du cortex cingulaire antérieur et de l'insula, avant de subir un léger choc physique sur la peau par un laser, prédit l'intensité de la douleur ressentie par le sujet (Boly et al., 2007), suggérant que, comme pour le réseau des valeurs, **l'état initial de ce réseau** peut biaiser la perception d'un stimulus.

Des enregistrements de l'activité des neurones dans le **cortex cingulaire humain** ont montré que certains neurones réagissaient à des stimuli provoquant des sensations de douleur (Hutchison et al., 1999). Ces résultats sont cependant à nuancer par le fait qu'il a été montré qu'il était impossible de dissocier ce réseau de la douleur d'un réseau qui encodait seulement la saillance (Mouraux et al., 2011) et que stimuler des patients implantés avec des électrodes intracrâniennes au niveau des neurones du cortex cingulaire antérieur ne provoque pas de sensation de douleur (Caruana et al., 2018; Hutchison et al., 1999). D'un autre côté, depuis 2014, des électrodes implantées dans le cortex et qui stimulent le cortex cingulaire antérieur permettent d'alléger les sensations de douleur chronique chez les patients qui en sont atteints (Boccard et al., 2014; Levi et al., 2019; Russo and Sheth, 2015; Spooner et al., 2007). Le rôle précis du cortex cingulaire dans la perception de la douleur reste donc à élucider. Dans tous les cas, il est clair que son activité peut être élicitée par des stimuli jugés douloureux et qu'il joue un rôle dans le ressenti subjectif de la pénibilité d'un stimulus douloureux. Le dmPFC, que nous avons déjà évoqué, est légèrement plus dorsal par rapport à l'épicentre des clusters liés à la

## Introduction

douleur (voir <https://neurosynth.org/analyses/terms/pain/>), mais il est régulièrement impliqué aussi dans sa perception (voir **Figure 17**).

**L'insula** est liée à l'intéroception, c'est-à-dire la perception subjective des états internes du corps (Craig, 2002; Namkung et al., 2017). Elle s'active dans des situations liées à une sensation psychologique de dégoût ou des mots évoquant ce genre d'émotion (Ziegler et al., 2018). Son activité corrèle avec l'intensité de la douleur perçue par un individu (Baliki et al., 2009). De plus, un nombre conséquent d'études où l'insula était stimulée électriquement chez l'homme ont montré qu'activer les neurones de l'insula suffisait à provoquer une sensation subjective de douleur (Afif et al., 2010; Frot and Mauguière, 2003; Isnard et al., 2004; Mazzola et al., 2012, 2017, 2019; Stephani et al., 2011). De plus, au niveau cortical au moins, il semblerait que seules des stimulations électriques de l'insula et de la région somatosensorielle secondaire qui lui est adjacente aient élicitée une sensation de douleur (Mazzola et al., 2012) ce qui suggère que l'insula est non seulement suffisante, mais aussi nécessaire pour la perception subjective de la douleur. Ces résultats sont cependant à nuancer par des études chez des patients ayant subi des lésions unilatérales de l'insula qui non seulement ressentent toujours de la douleur, mais y sont même plus sensibles que des sujets contrôles (Feinstein, 2013; Starr et al., 2009). Certains patients avec des lésions dans l'insula peuvent même développer un syndrome de douleur qui a été nommé syndrome de douleur operculo-insulaire (Garcia-Larrea et al., 2010). Le rôle précis de l'insula n'est donc pas encore complètement clarifié en lien avec la douleur, mais il est certain qu'elle joue un rôle primordial dans le ressenti de la douleur et que, quand elle est lésée, la perception de la douleur s'en retrouve altérée.



**Figure 17 : Aires cérébrales liées à la sensation de douleur.** Ces zones sont d'autant plus actives que la douleur était jugée intensément dans une tâche où différents niveaux de chaleur (46, 47 et 48°C) étaient exercés sur leur bras gauche et sur leur pied gauche. Les zones principales qu'on peut observer sont le cortex cingulaire avec un cluster qui inclue aussi le dmPFC et l'aire supplémentaire motrice et l'insula. Figure issue de (Krishnan et al., 2016).

## Introduction

Les **punitions d'ordre psychologique**, comme la perte d'argent ou le rejet social, induisent une activité dans l'**insula antérieure**, le **striatum dorsal**, le **cortex cingulaire** et le **dmPFC** (Bartra et al., 2013; Eisenberger and Lieberman, 2004; Palminteri et al., 2012). En outre, une équipe a montré que la sensibilité de l'insula antérieure en anticipation à des pertes dans une tâche monétaire permettait de prédire la capacité des sujets à apprendre spécifiquement dans le domaine des pertes dans une autre expérience monétaire qui a eu lieu 8 à 10 mois plus tard (Samanez-Larkin et al., 2008). De plus, des patients avec des lésions dans l'insula antérieure ou dans le striatum dorsal présentent des déficits dans l'apprentissage par renforcement qui sont spécifiques aux pertes, mais pas aux récompenses (Palminteri et al., 2012).

Ces résultats suggèrent qu'il existe bien un réseau commun des punitions, qu'elles soient d'ordre physique ou psychologique, et que ce réseau inclue au moins l'insula, le cortex cingulaire et le dmPFC (voir **Figure 18**). Cependant, le rôle respectif de chacune des zones de ce réseau n'est pas encore totalement compris. L'idée d'un réseau qui s'active en réponse à n'importe quel type de punition a amené certains chercheurs à supposer l'existence d'un système neural d'alarme commun qui réagirait à tout type de menace (Eisenberger and Lieberman, 2004). Les auteurs proposent que le dmPFC identifie des situations où il existe des déviations entre ce qui est désiré et ce qui se passe dans l'environnement afin de signaler au reste du cerveau de prendre des mesures afin de minimiser cette différence.



**Figure 18 : Réseaux communs entre punitions physiques et punitions psychologiques.** Figure du haut : en jaune, les clusters corrélant avec l'aversion non-douloureuse, en bleu, les clusters corrélant avec la sensation de douleur ; figure du bas : en vert, zones actives autant pour la douleur que pour les punitions d'ordre psychologique. On peut notamment distinguer le cortex cingulaire, le dmPFC et l'insula comme autant de clusters présents dans les deux types de punition. Figure issue d'une méta-analyse de (Hayes and Northoff, 2012).

La sensation de **peur** c'est-à-dire l'appréhension d'une punition à venir, a été très fortement corrélée sur les modèles animaux comme chez l'homme avec une activité importante dans l'**amygdale** (Izquierdo et al., 2016; LeDoux, 2000, 2014). Des lésions à l'amygdale chez l'homme diminuent la capacité de conditionnement à la peur, mesurée par la conductance dermale (Bechara et al., 1995). De plus, le volume de l'amygdale est corrélé avec la magnitude de l'effet de la peur sur la conductance dermale chez des sujets sains (Cacciaglia et al., 2015). D'autres régions connectées à l'amygdale, telles que l'hippocampe, semblent cependant aussi jouer un rôle dans la réaction de peur bien qu'elles soient moins centrales (Izquierdo et al., 2016). Il est intéressant, au passage, de constater que, d'après une méta-analyse, le dmPFC semble impliqué dans l'inhibition d'une réaction de peur (Fullana et al., 2018). La peur est conçue comme l'appréhension d'une menace perçue dans le présent. L'anxiété, par contre, est vue comme une réaction d'anticipation de menaces dans l'environnement. Certains chercheurs ont émis l'hypothèse que l'insula, pas spécialement reliée à la réaction de peur, soit particulièrement recrutée dans l'anxiété suggérant qu'une intéroception altérée dans l'insula pourrait expliquer en partie le syndrome anxieux (Paulus and Stein, 2006).

Enfin, **l'évitement d'une punition peut activer les réseaux de la récompense de manière similaire à l'obtention d'une récompense** (Kim et al., 2006; Tom et al., 2007). Plusieurs études ayant inclus des gains monétaires décomptés par l'éventualité d'une punition physique (un choc électrique sur la main gauche) ont montré que là encore **le réseau des valeurs, notamment le vmPFC, représentait la valeur subjective des récompenses dévaluée par l'éventualité du choc électrique** (Park et al., 2011; Shapiro and Grafton, 2020; Talmi et al., 2009). Il semblerait donc bien que dans ce cas aussi le réseau des valeurs prenne en compte ce dévaluateur, les punitions, dans l'évaluation des actions.

## II.6 Le dmPFC : un système de réaction ?

### II.6.a Liens entre les différents dévaluateurs

Beaucoup de chercheurs ont cherché à voir s'il existait des points communs, autant en termes computationnels qu'en termes d'aires cérébrales concernées, entre ces différents dévaluateurs de la valeur subjective. Différentes équipes ont souvent cherché à comparer ces différents dévaluateurs au sein d'une même étude. Il y a par exemple des études comparant le décompte intertemporel et le décompte par l'effort physique (Klein-Flügge et al., 2015; Prevost et al., 2010; Sugiwaka and Okouchi, 2004), le décompte intertemporel et le décompte par l'effort mental (Massar et al., 2015; Vassena et al., 2014), le décompte par l'effort mental et

## Introduction

l'effort physique (Chong et al., 2017; Schmidt et al., 2012; Smit et al., 2005), le décompte par les probabilités et par le délai (Andrade and Petry, 2012; Ohmura et al., 2006; Rachlin et al., 1991; Takahashi et al., 2007) ou encore le décompte intertemporel, le décompte par les probabilités et le décompte par l'effort physique (Seaman et al., 2018).

Au niveau computationnel, il semblerait que la plupart de ces dévaluateurs agissent par une formule similaire. Comme on l'a déjà évoqué, Chong *et al.* ont montré que le décompte par l'effort physique était mieux décrit par une fonction de décompte de la forme  $SV = R - k * E^2$ , alors que le décompte par l'effort mental était mieux décrit par une fonction hyperbolique  $SV = \frac{R}{1+k*E}$  (Chong et al., 2017). Le décompte intertemporel des récompenses et le décompte par les probabilités dans le cas du risque sont eux aussi généralement décrit par une fonction hyperbolique de la forme  $SV = \frac{R}{1+k*D}$  pour le délai où  $D$  est le délai (voir *L'aversion pour le délai*) et  $SV = \frac{R}{1+k*(\frac{1}{p}-1)}$  où  $p$  est la probabilité de la récompense  $R$  d'être obtenue (voir *L'incertitude est subjective*). Les punitions, par contre, semblent interagir avec les récompenses dans la perception de la valeur subjective (voir *Dévaluation par les punitions*). À l'exception de l'effort physique et des punitions, la plupart de ces modèles sont tous de type **hyperbolique**. Il faut noter cependant qu'il n'y a que dans le cas des choix intertemporels où une telle fonction est vraiment consensuelle. Cela paraît d'ailleurs logique puisque c'est le seul cas, avec les probabilités, où avoir une fonction nécessairement positive a du sens. Il vaut mieux obtenir une récompense dans un délai très long que de ne rien obtenir. A l'inverse, dans le cas des efforts ou des punitions, il paraît absurde que la fonction ne puisse jamais prendre de valeurs négatives, et qu'elle implique qu'il vaille mieux subir une punition ou effectuer un effort pénible plutôt que de ne rien faire, même face à une récompense minime.

Plusieurs chercheurs se sont donc posé la question de savoir si les paramètres subjectifs utilisés dans chacun de ces modèles étaient corrélés entre les différents types de dévaluateurs, en particulier dans les cas où la fonction de décompte est similaire. Il a notamment été proposé que le décompte des délais lointains soit seulement dû à un décompte probabiliste lié au fait que plus un événement est lointain, plus la probabilité qu'il ait lieu est faible (Mischel, 1966; Rachlin et al., 1991; Rotter, 1954). Dans la même série d'idée, la sensation subjective et aversive liée à un effort a été liée à une dévaluation du temps passé à effectuer une action qui ne fournit pas un taux suffisamment élevé de récompenses à court terme (Kurzban, 2016; Kurzban et al., 2013). Il a cependant été montré que les paramètres subjectifs issus de ces différents modèles n'étaient pas ou très peu corrélés que ce soit entre le délai et l'effort physique

## Introduction

(Klein-Flügge et al., 2015; Sugiwaka and Okouchi, 2004), les probabilités et le délai (Myerson et al., 2003; Olson et al., 2007; Reynolds et al., 2004; Takahashi et al., 2007) ou l'effort physique et l'effort mental (Chong et al., 2017), bien que d'autres études aient pu trouver une corrélation significative entre les paramètres du décompte par l'effort physique et ceux du décompte par l'effort mental (Białaszek et al., 2017), notamment quand les efforts étaient fictifs (Nishiyama, 2016). Seaman *et al.* ont effectué un protocole avec 75 participants devant effectuer une tâche de décompte par l'effort physique, une tâche de décompte par le délai et une tâche de décompte par la probabilité, toutes modélisées par une fonction hyperbolique, et ils ont montré que les paramètres de décompte de chaque tâche n'étaient pas corrélés à travers les participants (Seaman et al., 2018). Ceci ne permet bien sûr pas de dire que ces concepts sont encodés dans différentes parties du cerveau ou qu'ils n'interagissent pas, mais cela montre qu'ils sont partiellement indépendants les uns des autres. En résumé, au niveau computationnel, la plupart de ces études montrent une **absence de corrélation entre les différents paramètres de décompte**.

D'un autre côté, un phénomène curieux est qu'une fatigue au niveau cognitif semble provoquer une augmentation de la perception subjective du coût d'un effort physique et une réduction de la performance et de l'endurance liées à une activité physique (Marcora et al., 2009; Van Cutsem et al., 2017). Réciproquement, un surentraînement au niveau physique impacte les capacités exécutives des sujets dans une étude de choix intertemporels (Blain et al., 2019). Ces résultats suggèrent donc qu'il existe une ressource commune, au moins entre l'effort physique et l'effort mental. Pourtant, nous avons vu que la sensibilité à l'effort physique n'est pas corrélée à la sensibilité pour l'effort mental dans la plupart des expériences où les chercheurs ont tenté de corréler ces paramètres chez les mêmes participants. Ce résultat paradoxal pourrait s'expliquer par le fait que certains efforts physiques requièrent du contrôle cognitif pour être correctement exécutés. Des activités physiques nécessitant d'être précis ou d'inhiber des réactions automatiques pourraient ainsi nécessiter l'allocation accrue d'attention et de ressources liées à l'exécution d'un effort mental. Certains auteurs ont même proposé que le coût subjectif d'une activité physique soit principalement lié à un coût mental en termes d'inhibition d'une réaction automatique, plutôt qu'en termes d'activité musculaire (Shenhav et al., 2017). Le lien entre les ressources requises par des activités physiques et par des activités mentales reste encore à explorer pour mieux comprendre dans quelle mesure elles sont liées ou indépendantes. Quoiqu'il en soit, il paraît que, au niveau comportemental en tout cas, les sensibilités pour ces dévaluateurs des récompenses ne soient pas vraiment corrélées. Cela

## Introduction

suggère que les capacités ainsi que l'apprentissage lié à chacun de ces dévalueurs au cours de la vie se font de manière indépendante. Cela n'empêche d'ailleurs pas que ce soit les mêmes structures cérébrales qui aient en charge de considérer la sensibilité à chacun des dévalueurs. Cela veut juste dire que, si c'est le cas, la sensibilité à chacun d'entre eux est encodée de manière relativement indépendante.

### II.6.b Deux systèmes neuraux avec des rôles différents

Au niveau comportemental, les liens entre ces différents dévalueurs existent peut-être, mais ils ne semblent pas être très robustes. Au niveau neural cependant, il est intéressant de constater qu'on retrouve régulièrement des structures similaires pour encoder autant les dévalueurs directement que la valeur subjective décomptée par les dévalueurs.

En effet, dans à peu près toutes ces situations, **la valeur subjective dévaluée est systématiquement représentée dans le réseau des valeurs, en particulier au sein du vmPFC** (Arulpragasam et al., 2018; Hsu et al., 2009; Kable, 2014; Kable and Glimcher, 2007; Levy et al., 2010; Park et al., 2011; Peters and Büchel, 2011; Seaman et al., 2018; Shapiro and Grafton, 2020; Talmi et al., 2009). Il encode visiblement bien la valeur des récompenses dévaluée par les différents types de dévalueurs et non pas seulement la valeur des récompenses en jeu.

Le réseau des valeurs semble aussi impliqué dans **l'incitation à exercer un effort** quand le calcul bénéfices/coûts est positif. De plus, c'est lui qui indique quelle intensité d'effort effectuer en fonction de la récompense en jeu (Pessiglione et al., 2007; Schmidt et al., 2008), indépendamment de la nature physique ou mentale de l'effort à exercer (Schmidt et al., 2012).

En résumé, et pour faire le lien avec la théorie de Kent Berridge (Berridge et al., 2009), on pourrait dire que ce réseau des valeurs permet 1) d'évaluer un rapport bénéfices/coûts lié à un objectif donné et 2) d'inciter à agir en vue d'atteindre ces objectifs (la composante « *wanting* » de Kent Berridge).

Enfin, on retrouve un groupe d'aires cérébrales régulièrement associées à l'intensité des coûts cérébraux dans la plupart de ces études qui comprend principalement **l'insula antérieure et le dmPFC**. On retrouve en effet le dmPFC, régulièrement accompagné de l'insula antérieure, en lien avec l'incertitude (Behrens et al., 2007; Poudel et al., 2020; Stern et al., 2010; Sun et al., 2017; Volz et al., 2004; White et al., 2014), l'estimation du temps (Coull et al., 2011; Kolling and O'Reilly, 2018; Protopapa et al., 2019; Radua et al., 2014), la difficulté et l'intensité d'un effort mental à effectuer (Barch et al., 1997; Engström et al., 2015; Ridderinkhof



## Introduction

et al., 2004; Shenhav et al., 2014, 2016b; Wager et al., 2005), l'intensité d'un effort physique à exécuter (Burke et al., 2013), et même en corrélation avec le fait de subir des punitions (Krishnan et al., 2016; Shackman et al., 2011; Wager et al., 2013). En faisant une conjonction entre les clusters liés aux mots « *uncertainty* » (incertitude), « *effort* », « *punishment* » (punition) et « *delay* » (délai) dans [Neurosynth](#), on retrouve d'ailleurs en effet le dmPFC et l'insula antérieure bilatéralement (voir **Figure 19** haut). De plus, même en testant à nouveau cette conjonction après avoir remplacé l'effort par sa composante physique avec le mot « *force* » et sa composante mentale avec le mot « *cognitive control* » (contrôle cognitif) et les punitions par leur composante physique avec le mot « *pain* » (douleur) et leur composante mentale avec le mot « *loss* » (pertes), on retrouve toujours le dmPFC et l'insula antérieure bilatéralement (voir **Figure 19** bas).



**Figure 19 : Le délai, l'incertitude, l'effort et les punitions sont liés à l'activité du dmPFC et de l'insula antérieure.** Figures du haut : Conjonction entre les voxels qui corrént avec les mots « *effort* », « *punishment* », « *uncertainty* » et « *delay* » dans [Neurosynth](#) via le test d'uniformité pour chacun de ces termes (correction pour le taux de faux positifs FDR pour les comparaisons multiples  $p.value < 0.01$ ). Figures du bas : Conjonction entre les voxels qui corrént avec les mots « *uncertainty* », « *delay* », « *force* », « *cognitive control* », « *pain* » et « *loss* » dans [Neurosynth](#) via le test d'uniformité pour chacun de ces termes (correction pour le taux de faux positifs FDR pour les comparaisons multiples  $p.value < 0.01$ ). En rouge, on voit uniquement les voxels présents systématiquement pour tous les termes mentionnés pour chaque image. Figures de gauche : zoom sur le cluster dans le dmPFC aux coordonnées ( $x = -2$  ;  $y = 10$  ;  $z = 46$ ). Figures de droite : zoom sur le cluster dans l'insula antérieure droite aux coordonnées ( $x = 34$  ;  $y = 20$  ;  $z = 4$ ). Comme on peut le voir sur les deux images de droite, l'insula antérieure est présente de manière bilatérale. Dans les deux tests, on observe aussi un cluster bilatéral dans le thalamus ventro-latéral qui est impliqué dans la génération du

## Introduction

mouvement via ses connexions au cortex primaire moteur et au cortex prémoteur<sup>22</sup> ainsi qu'un cluster dans le dlPFC gauche et dans le cortex pariétal postérieur gauche. Dans la figure du haut, on constate en outre un cluster dans le striatum ventral et un cluster dans le vmPFC, ainsi qu'un cluster dans le dlPFC droit et dans le cortex pariétal postérieur droit. Dans la figure du bas, on observe aussi un cluster dans l'amygdale gauche. Les clusters issus de la conjonction sont affichés sur un template de SPM qui est basé sur la moyenne des scans T1 de 152 cerveaux humains. Les coordonnées affichées sont présentées dans l'espace MNI (Montreal Neurological Institute).

Le rôle du dmPFC en lien avec ces différents dévaluateurs est cependant loin d'être complètement élucidé. Une possibilité évoquée par une théorie qui a maintenant plus de 20 ans est qu'il détecte les situations conflictuelles où il est nécessaire d'employer de l'effort mental et qu'il recrute alors le dlPFC afin de résoudre la situation (Botvinick et al., 1999, 2001, 2004; Carter and van Veen, 2007; Kerns et al., 2004; Ridderinkhof et al., 2004). McKlveen propose, sur la base des études chez le rongeur, que le cortex frontal médian serve à mettre en place un comportement contexte-dépendant et à recruter des ressources énergétiques en mobilisant le système nerveux autonome pour s'adapter face au stress (McKlveen et al., 2015). De la même manière, on pourrait interpréter l'activation du dmPFC face à ces différents concepts aversifs comme une réaction afin d'allouer des ressources pour s'adapter face à eux. Il pourrait réagir à ces événements aversifs en recrutant plus de contrôle cognitif pour résoudre la situation comme cela a été proposé depuis 20 ans par l'équipe de Jonathan Cohen et ses successeurs principaux que sont Matthew Botvinick et Amitai Shenhav (Shenhav et al., 2013). Le contrôle cognitif pourrait alors servir, entre autres, à évaluer les bénéfices d'abandonner la stratégie d'exploiter l'espace environnant et de se remettre à explorer l'environnement. Il a d'ailleurs été montré que le dmPFC s'active d'autant plus que la valeur subjective d'explorer d'autres environnements par rapport à l'environnement présent est attrayante (Kolling et al., 2012, 2016b). Cette théorie expliquerait pourquoi le dmPFC s'active proportionnellement à l'intensité d'une douleur physique sans pour autant que son activité, quand il est stimulé, provoque de sensation subjective de douleur (Caruana et al., 2018; Hutchison et al., 1999). Cela est aussi cohérent avec l'hypothèse d'**un système neural d'alarme**, qui réagirait en réponse à toute menace, émise par Eisenberger et Lieberman (Eisenberger and Lieberman, 2004). Il a d'ailleurs déjà été proposé que la corrélation entre le dmPFC et l'incertitude soit principalement liée au fait que le dmPFC cherche ainsi à identifier les situations qui requièrent l'implémentation de contrôle cognitif (Mushtaq et al., 2011). Le dmPFC ne réagirait cependant pas qu'aux menaces réelles, comme c'est le cas dans les punitions, mais aussi dans les cas où des difficultés sont identifiées

---

<sup>22</sup> [https://en.wikipedia.org/wiki/Ventral\\_lateral\\_nucleus](https://en.wikipedia.org/wiki/Ventral_lateral_nucleus)

## Introduction

dans l'environnement et qu'il faut les surmonter. Parvizi *et al.* ont stimulé deux patients implantés avec des électrodes intracrâniennes dans le cortex cingulaire près du dmPFC. Ils ont pu montrer que stimuler le cortex à cet endroit pouvait générer chez leurs patients une sensation d'être prêt à affronter tout défi qui se présenterait à eux (Parvizi et al., 2013). Ce résultat est cependant à nuancer du fait qu'il n'y a que deux patients inclus dans cette étude, que les électrodes sont plus ventrales que le cluster qu'on a caractérisé jusque-là sous le nom de dmPFC et, enfin, qu'une étude incluant 329 patients qui a caractérisé l'impact de stimuler le cortex cingulaire en différents endroits n'a pas pu répliquer ces résultats quelle que soit la zone stimulée (Caruana et al., 2018). L'étude de Caruana *et al.*, même si elle n'est pas centrée sur le dmPFC mais plutôt sur la partie du cortex cingulaire qui lui est ventrale, a néanmoins permis de montrer que toutes les stimulations dans la zone qui correspond plus ou moins au dmPFC tel qu'on l'a défini jusqu'à présent, générait des mouvements assez cohérents avec la théorie développée ici. En effet, des stimulations autour de cette zone provoquait des enchaînements d'action qui donnaient l'impression que les patients cherchaient à gérer une menace dans leur environnement ou à fuir de la situation présente. Les patients stimulés cherchaient par exemple à se lever comme s'ils cherchaient à fuir une situation. Un des patients a dit qu'il a senti qu'il avait envie de s'en aller sans être capable de fournir une justification à cette sensation. D'autres réactions typiques des patients étaient qu'ils se mettaient à explorer l'environnement avec leurs yeux ou à avoir des mouvements vers leur visage comme s'ils cherchaient à le protéger d'une menace (Caruana et al., 2018).

En résumé, le dmPFC, en lien avec l'insula antérieure, pourrait constituer un système de réaction à des situations nécessitant l'allocation de contrôle cognitif. Il se chargerait ainsi non seulement d'identifier les situations requérant l'allocation de plus de ressources, mais aussi de calculer le ratio optimal de ressources à allouer à une situation donnée. La ressource en question pourrait être unique et consister dans l'exécution de contrôle cognitif, mais elle pourrait aussi varier en fonction de la situation. On peut par exemple facilement imaginer des situations où la ressource consiste dans le temps alloué à une tâche donnée.

### II.6.c Hypothèses alternatives pour le dmPFC

- **Un réseau appétitif et un réseau aversif**

Dans le cas de l'humeur, il a été montré qu'il existe **deux systèmes opposés**, avec d'un côté un réseau qui encode l'humeur positive, et de l'autre un réseau qui encode l'humeur négative. Le réseau de l'humeur positive correspond au réseau des valeurs. Il inclut en effet,

## Introduction

selon les études, soit le vmPFC et le cortex postérieur cingulaire (Vinckier et al., 2018), soit le striatum ventral (Knutson et al., 2014; Rutledge et al., 2014). Le réseau de l'humeur négative inclut l'insula antérieure (Knutson et al., 2014; Vinckier et al., 2018) et le dmPFC (Vinckier et al., 2018). Dans le champ des valeurs, il se pourrait de même que ces deux réseaux soient opposés avec, d'un côté, **le réseau des valeurs spécialisés dans l'encodage des valeurs positives** et, de l'autre, **un réseau constitué du dmPFC et de l'insula antérieure qui encoderait les valeurs négatives** (Bartra et al., 2013; Fouragnan et al., 2018). Ce lien entre le dmPFC et l'insula antérieure avec les valeurs négatives est nettement attesté au moment du traitement des conséquences d'une action, le *feedback* en anglais. Néanmoins, ce lien est moins évident au moment de la décision ou de la prise de décision, où c'est le striatum dorsal qui semble mener la danse. Il a été proposé que cet encodage de la valeur négative par le dmPFC et l'insula antérieure au moment du feedback serve à la fois à initier une réaction précoce pour réagir à l'évènement négatif rapidement et en même temps à moduler l'apprentissage sur un temps légèrement plus long (Fouragnan et al., 2015). Ceci permettrait d'améliorer l'identification de situations favorables à approcher par rapport à des situations qu'il faudrait éviter. On voit donc que ce résultat n'est pas totalement incompatible avec notre proposition d'un système réagissant à une menace en vue de mettre en place une réponse adaptée.

Une proposition légèrement différente propose aussi que le dmPFC encode, non pas la valeur subjective dans le domaine aversif, mais la valeur nette après avoir intégré les coûts et les bénéfices associés à une action. L'insula antérieure encoderait les coûts associés à une action alors que le vmPFC encoderait la valeur des gains. Le dmPFC, lui, se chargerait de calculer la valeur nette après avoir dévalué les gains par les coûts (Pessiglione and Delgado, 2015; Pessiglione et al., 2018). Le dmPFC enverrait ensuite un signal aux aires motrices pour mettre le corps en mouvement et agir en fonction de la valeur nette issue de ce calcul (Pessiglione et al., 2018). Dans cette même logique, il a été proposé que, lors d'un choix binaire, le dmPFC intègre les informations envoyées par le réseau des valeurs au fur et à mesure quant à la valeur associée à chacune des options afin de pouvoir effectuer un choix (Pisauro et al., 2017). Cependant, cette hypothèse semble incompatible premièrement avec le fait que le vmPFC encode directement les gains décomptés par les coûts, comme on l'a déjà vu pour les différents dévalueurs (voir *Le réseau des valeurs est impacté par l'incertitude*, *Bases neurales du délai*, *Liens bidirectionnels entre réseau des valeurs et effort physique* et *Réseau des valeurs et effort mental*). Deuxièmement, le fait que le dmPFC corrèle mieux avec la saillance des enjeux plutôt qu'avec la valeur nette (Litt et al., 2011) plaide aussi contre l'idée qu'il s'agirait d'une aire en

## Introduction

bout de chaîne qui intégrerait les coûts et les bénéfices liés à l'action. Cependant, plus d'études sont nécessaires en vue d'y voir plus clair.

- **L'encodage et la mise à jour des modèles internes du monde**

Une autre hypothèse propose que le dmPFC se contente de mettre à jour des **modèles internes du monde** (Kolling et al., 2016a; O'Reilly et al., 2013). Ceci serait d'ailleurs cohérent avec la théorie proposant que ce qui est coûteux dans l'effort mental, c'est le **coût de l'information** encodée par les neurones (Zénon et al., 2019). En effet, on pourrait envisager que lorsque les modèles internes du monde doivent être mis à jour, le coût de l'information, c'est-à-dire la divergence de Kullback-Leibler entre les *a posteriori* et les *a priori* selon cette théorie, est plus élevée. Cela expliquerait aussi pourquoi le dmPFC corrèle avec la volatilité dans l'environnement (Behrens et al., 2007), puisqu'il s'agit typiquement de situations où les modèles internes du monde doivent constamment être mis à jour. Alexandre Zénon, dans sa revue de 2019, montre comment cette théorie du dmPFC, basée sur la théorie de l'information, est compatible avec la plupart des tâches où le dmPFC est actif (Zénon et al., 2019).

- **Un rôle unique ou des fonctions multiples ?**

Un débat ancien et bien plus général que l'aire du dmPFC tourne autour de la question du rôle assigné à chaque zone cérébrale. Plusieurs théories s'opposent très fortement et depuis longtemps. D'un côté, il y a les tenants d'une organisation du cerveau en **modules anatomo-fonctionnels**. Chaque zone cérébrale aurait ainsi un rôle fonctionnel unique qu'il suffirait de découvrir (Anderson, 2010; Bergeron, 2007). De l'autre côté, les tenants de la **réutilisation neurale** proposent que le cerveau fonctionne en différents réseaux et qu'une même zone cérébrale puisse avoir des fonctions très différentes en parallèle selon les réseaux dans lesquelles elle est impliquée (Anderson, 2010). Cette dernière hypothèse semble de plus en plus confirmée par les faits de manière générale en ce que l'hypothèse de la réutilisation neurale a pu trouver des preuves dans le langage, l'écriture ou dans l'arithmétique qui toutes deux recrutent des aires cérébrales aussi impliquées dans d'autres processus (Anderson, 2010; Dehaene and Cohen, 2007), mais il reste possible que certaines aires y soient plus sujettes que d'autres. Ce débat se reflète au niveau du dmPFC par deux visions très différentes entre les chercheurs américains de Princeton avec principalement Jonathan Cohen, Matthew Botvinick et Amitai Shenhav et, de l'autre, les chercheurs d'Oxford avec principalement l'équipe de Matthew Rushworth et Nils Kolling. Les premiers cherchent depuis plus de 20 ans à expliquer l'activité du dmPFC par une **théorie unique**. Ils ont d'abord postulé que le dmPFC identifiait

## Introduction

les situations conflictuelles, c'est-à-dire celles où plusieurs réponses cognitives et/ou motrices sont possibles et semblent équivalentes, afin d'indiquer au reste du cerveau quand ces situations conflictuelles ont lieu et d'allouer des ressources en fonction pour résoudre le conflit (Botvinick et al., 2001, 2001, 2004; Carter et al., 1998; Kerns et al., 2004). Désormais, leur théorie tourne autour de la valeur attendue d'exercer du contrôle cognitif que nous avons déjà évoqué à plusieurs reprises (Shenhav et al., 2013, 2016a). Les seconds en revanche proposent que le dmPFC soit indépendamment inclus dans plusieurs réseaux fonctionnels (Kolling et al., 2016b) et qu'il s'active en lien avec la valeur d'explorer l'environnement (Kolling et al., 2012, 2016a) ou encore avec la mise à jour de modèles internes de l'environnement (Kolling et al., 2016a) de manière indépendante. Ils revendiquent de ne pas chercher à regrouper toutes les fonctions identifiées du dmPFC sous une seule théorie (Kolling et al., 2016b). Le débat n'est donc pas encore tranché et comprendre le rôle précis du dmPFC reste encore une sorte de Graal pour plusieurs chercheurs. Cependant, après plus de 20 ans d'études autour de cette zone, force est de constater que de nets progrès ont été faits pour mieux comprendre son rôle en lien avec le reste du cerveau et avec le comportement. Ces différentes théories, qu'elles soient ou non vérifiées, poussent à tester différentes hypothèses et donc à stimuler des recherches fructueuses sur ce sujet, malgré les tensions qu'elles génèrent parfois entre les différentes équipes qui cherchent à défendre chacune sa théorie. Peut-être un jour aurons-nous enfin la réponse de savoir quel(s) rôle(s) le dmPFC remplit dans le cortex cérébral.

### II.6.d Système de réaction et éveil

**Le dmPFC est régulièrement associé à des composantes du système nerveux autonome, en lien avec l'éveil et l'excitabilité.** Le dmPFC corrèle par exemple avec le diamètre de la pupille au repos (Murphy et al., 2014; Schneider et al., 2016), les changements de la pupille pendant une tâche (Schneider et al., 2018), ou l'activité du locus coeruleus (Alnaes et al., 2014). Plus généralement, l'activité du dmPFC et de l'insula antérieure ont été corrélées au fonctionnement de deux systèmes clés du système nerveux autonome qui régulent l'état d'éveil : les systèmes orthosympathique et parasympathique (Beissner et al., 2013; Critchley et al., 2003; Valenza et al., 2019). Il a de plus été montré que le dmPFC corrélait avec des événements saillants, comme la surprise dans le cas des erreurs de prédiction (Garrison et al., 2013; Vassena et al., 2017) ou encore avec des enjeux élevés, indépendamment de s'il s'agit de récompenses ou de punitions (Litt et al., 2011). Il a régulièrement aussi été associé à l'identification d'erreurs ayant été commises dans le comportement (Iannaccone et al., 2015; McGuire and Botvinick, 2010). Plus généralement, le dmPFC, ainsi que l'insula antérieure qui

## Introduction

y est souvent fonctionnellement associée, ont été proposés comme constituants un réseau de la saillance qui détecterait les stimuli qui sont importants pour l'individu afin d'initier une réponse comportementale adaptée (Seeley et al., 2007; Uddin, 2015).

De plus, **la plupart des dévaluateurs que nous avons évoqués jusqu'ici entraînent une augmentation de l'activité du système nerveux autonome.** Il a par exemple été montré que non seulement la sensation subjective d'effort physique (Borg, 1970; Zénon et al., 2014), mais aussi celle d'effort mental corrèle avec une augmentation de l'activité cardiovasculaire (Hidalgo-Muñoz et al., 2018; Richter et al., 2008, 2016; Silvestrini, 2017) et du diamètre pupillaire (Alnaes et al., 2014; Kahneman and Beatty, 1966; Klingner et al., 2008; Lempert et al., 2016; Peysakhovich et al., 2015). L'incertitude aussi provoque une augmentation du diamètre pupillaire (Colizoli et al., 2018; Friedman et al., 1973; Geng et al., 2015; Urai et al., 2017; Zénon, 2019). De plus, les différents types d'incertitudes ont été liés aux deux neuromodulateurs principaux du système nerveux autonome avec d'un côté l'acétylcholine, liée au système orthosympathique, qui encoderait le risque et de l'autre la noradrénaline, liée au système parasympathique, qui encoderait l'ambiguïté (Yu and Dayan, 2003). Dans le cas des choix intertemporels, le lien est un peu moins clair, mais il semblerait que la taille de la pupille permette de prédire quand la récompense différée sera choisie (Lempert et al., 2016). Il semblerait en outre que les individus plus impulsifs, c'est-à-dire ceux qui préfèrent les récompenses à court terme dans ce type de tâches, auraient aussi des niveaux d'éveil autonome plus faibles par rapport aux autres (Herman et al., 2018). En 2011, Mouraux *et al.* expliquaient que l'activité du dmPFC en lien avec la douleur s'expliquait mieux par un lien avec la saillance liée au ressenti d'un événement douloureux (Mouraux et al., 2011). De plus, dans l'expérience de Gallagher et ses collègues déjà citée précédemment, lorsque la jonction neuromusculaire est inhibée par du curare, mais que plus d'effort est nécessaire subjectivement pour atteindre le même niveau d'effort, cela va de pair avec une augmentation du pouls cardiaque, de la pression artérielle et de la noradrénaline sanguine en comparaison aux sujets contrôles (Gallagher et al., 2001). Ces résultats suggèrent donc un lien direct entre l'aspect aversif des différents dévaluateurs et le fait qu'ils soient saillants. Leur saillance stimule le système nerveux autonome dont l'activité peut, ou non, être liée à l'activité du dmPFC.

Le lien entre l'activité du dmPFC et l'éveil n'est pas du tout incompatible avec ce que nous avons développé précédemment, puisqu'un système neural en réaction à des événements aversifs permettrait justement de recruter des ressources afin de répondre aux menaces présentes dans l'environnement. L'activité du système nerveux autonome a d'ailleurs justement



## Introduction

été proposée comme un indicateur des ressources neurales disponibles (Matthews and Davies, 2001; Smit et al., 2004, 2005). Augmenter l'activité de ce système reviendrait donc à augmenter les ressources potentiellement utilisables par l'individu pour réagir face à une récompense ou une menace.

Si le lien entre l'activité du dmPFC et l'éveil semble relativement établi, le lien de causalité n'est cependant pas tout à fait clair. Il se pourrait en effet que des événements qui excitent le système d'éveil activent le dmPFC. Le locus coeruleus est la seule structure dans le cerveau qui produit de la noradrénaline et il est de ce fait l'un des composants principaux du système nerveux parasympathique. Or, le locus coeruleus possède des projections vers l'ensemble du cortex (Aston-Jones and Waterhouse, 2016; Benarroch, 2018) et pourrait ainsi augmenter l'activité du dmPFC. Il se pourrait aussi, comme on l'a suggéré, que ce soit l'activité du dmPFC qui recrute des ressources pour réagir aux menaces en cours. Il a notamment été montré que le cortex préfrontal peut influencer l'activité du locus coeruleus (Aston-Jones and Waterhouse, 2016) auquel il envoie des projections, au moins sur les modèles animaux (Arnsten and Goldman-Rakic, 1984; Aston-Jones and Waterhouse, 2016). Ces deux hypothèses n'étant pas nécessairement incompatibles, il se pourrait aussi qu'elles soient toutes deux valables, c'est-à-dire que des événements qui déclencheraient l'activation des circuits de l'éveil provoquent l'activité du dmPFC et réciproquement. Plusieurs preuves, surtout sur les modèles animaux, montrent en effet que l'insula tout comme le dmPFC pourraient aussi bien lire l'activité du système autonome, et donc réagir en fonction de celle-ci, que la modifier directement (Cechetto, 2014; Kano et al., 2020; McKlveen et al., 2015; Verberne and Owens, 1998).

Cependant, le lien entre l'éveil et les capacités, en termes d'effort mental notamment, varie en U inversé de sorte que, au-delà d'un certain stade, augmenter l'éveil est délétère pour la performance (Berridge and Spencer, 2017). Il faudrait voir si l'activité du dmPFC varie linéairement avec l'activité du système nerveux autonome, en l'activant ou en étant activé par lui, ou si, au contraire, il s'agit bien de deux processus distincts. Le dmPFC pourrait réguler l'activité du système nerveux autonome de manière instrumentale, c'est-à-dire l'augmenter uniquement dans les cas où c'est pertinent pour améliorer la performance et au contraire essayer de la réduire lorsque c'est pertinent. Certains perçoivent le dmPFC et le cortex cingulaire antérieur comme des modulateurs du système nerveux autonome qui permettraient justement de l'inhiber quand cela est adapté, et dont l'inhibition pourrait être dysfonctionnelle dans certaines pathologies (Kano et al., 2020). Le sens de l'interaction entre le dmPFC et les zones impliquées dans l'éveil, en particulier le locus coeruleus, reste encore à mieux identifier pour

## Introduction

savoir s'il est unidirectionnel ou bidirectionnel et, dans le cas où il serait unidirectionnel, dans quel sens il est dirigé, mais il est probable qu'ils soient liés directement ou indirectement.

En résumé, le dmPFC, avec l'insula antérieure, constitue un système dont l'activité augmente dans des situations nécessitant une réaction comportementale qui n'est pas automatisée dans le cerveau. Que ce soit face à une menace, comme dans le cas des punitions, face à l'exécution d'une activité physique ou mentale, face à une situation incertaine, ou dans le cas où un délai est nécessaire dans l'obtention d'une récompense et donc où une incertitude liée à l'obtention de cette récompense, on remarque systématiquement l'activité de ce réseau. Ce réseau va, de plus, de pair avec l'activité du système autonome. Si le lien de causalité n'est pas complètement établi. Il est possible soit que 1) le système autonome signale à ce réseau des situations saillantes dans lesquels une réponse comportementale adaptée doit être mise en place à l'aide de contrôle cognitif, soit que 2) le réseau constitué par le dmPFC et l'insula antérieure recrute le système autonome pour faciliter la mise en place d'une réaction comportementale plus efficace en termes de vitesse de réaction et d'intensité. Ces deux propositions ne sont d'ailleurs bien évidemment pas incompatibles. On peut en effet tout à fait imaginer que les deux propositions soient vraies selon les situations.

### **Résumé intermédiaire**

Le réseau des valeurs encode le désir d'un item donné, mais son activité semble aussi dévaluée directement par les différents coûts nécessaires à l'obtention de cet item. Ces coûts peuvent correspondre à l'exécution d'une activité physique ou mentale, associée à la sensation aversive d'effort, mais aussi au fait de subir une punition de forme physique, comme la douleur, ou mentale, comme une frustration par exemple. Enfin, ce réseau des valeurs voit aussi son activité, en réponse à l'obtention d'une récompense donnée, dévaluée par l'incertitude associée à l'obtention de cette récompense et aussi par le délai plus ou moins long dans lequel la récompense sera obtenu.

De plus, la prise en compte de chacun de ces dévaluateurs paraît associée à une activité accrue d'un réseau constitué du dmPFC et de l'insula antérieure. On peut facilement imaginer deux rôles pour cette relation. Une première possibilité serait que ce réseau calcule la valeur négative par laquelle chacun de ces dévaluateurs va impacter la valeur subjective. Il enverrait alors le résultat de ce calcul au réseau des valeurs pour que la valeur subjective qui va diriger le comportement intègre aussi cette composante négative. Une autre possibilité serait que ce

réseau réagisse à des situations associées avec des valeurs aversives afin de mettre en place une réponse comportementale adaptée.

### III La dynamique de la prise de décision

*“¿Debe ser bonito tomar tus propias decisiones verdad?”*

- *A mí cada vez me queda más claro, brujita, que, pues, que lo único que tenemos son las decisiones que tomamos.*<sup>23</sup>”

*Épisode 4, Saison 2, Narcos Mexico*

Nous avons amplement étudié les processus liés à l'évaluation autant pour ce qui est des récompenses que des dévaluateurs de celle-ci comme l'incertitude, le délai, l'effort physique, l'effort mental et les punitions. Nous avons aussi vu comment les différents dévaluateurs pouvaient impacter l'évaluation des récompenses au niveau comportemental, comme cérébral. En effet, nous avons pu identifier un réseau de zones cérébrales liées à l'évaluation qui comprend pour membres principaux le cortex préfrontal ventromédian et le striatum ventral, et dont l'activité augmente avec la valeur anticipée des récompenses et décroît avec les dévaluateurs qui peuvent y être associés. De plus, nous avons pu étudier un autre réseau de zones cérébrales liées à l'identification de stimuli aversifs qui y réagissait comme pour amener le reste du cerveau à mettre en place un plan d'action pour réagir à ces stimuli aversifs. À présent, nous allons étudier plus précisément les processus dynamiques qui mènent à la prise de décision. Pour ce faire, nous allons d'abord étudier les différents modèles de prise de décision qui ont pu être proposés ces dernières décennies. Ensuite, nous allons nous pencher sur les modèles dynamiques de la prise de décision dans le cas des choix binaires.

#### III.1 Décomposer la motivation

##### III.1.a Étapes de la motivation

*« une fois qu'on a posé la fin, on regarde la question de savoir comment et par quels moyens on peut l'atteindre et si plusieurs moyens paraissent en mesure de l'atteindre, on examine quel est le plus facile et le plus beau. »*

*2.3.3, Livre III, Éthique à Nicomaque, Aristote (Aristote, 2004)*

<sup>23</sup>

*« - Ça doit être bon de prendre ses propres décisions.*

*- Il devient de plus en plus clair pour moi que nous n'avons rien d'autre que nos décisions. »*

## Introduction

Dans la vision de Kent Berridge et de Morten Kringelbach que nous avons déjà évoqué, on peut décomposer le comportement motivationnel en 3 étapes (Berridge et al., 2009; Georgiadis and Kringelbach, 2012):

- 1) **le désir (*wanting*)** où le sujet désire quelque chose et est incité à l'obtenir par différentes zones cérébrales, dont le striatum ventral, qui le poussent à agir pour satisfaire ce désir ;
- 2) **la consommation** et l'appréciation de l'objet du désir (*liking*) où le sujet ressent du plaisir pour avoir obtenu ce qu'il désirait ;
- 3) **la satiété et l'apprentissage (*learning*)** où le sujet ne désire plus le but qu'il recherchait mais par contre il apprend puisqu'il peut réévaluer la valeur du but qu'il désirait en fonction de l'intensité du plaisir ressenti pendant la consommation et des coûts liés aux actions nécessaires pour obtenir ce plaisir.

Antonio Rangel et ses collègues vont encore plus décomposer ce processus en identifiant 5 étapes clés (voir **Figure 20** et (Rangel et al., 2008)):

- 1) **se représenter l'ensemble des actions possibles ;**
- 2) être capable **d'évaluer la valeur nette** (= bénéfice espéré – coût estimé de l'action) associée à chaque action de l'ensemble identifié en (1) en fonction des états internes (ressources disponibles, faim, soif, etc.) ;
- 3) **sélectionner**, parmi l'ensemble d'actions identifiées en (1), l'action avec la valeur nette estimée en (2) la plus élevée ;
- 4) **évaluer les conséquences de l'action effectuée** (bénéfice obtenu et coût effectué) ;
- 5) **apprendre**, c'est-à-dire mettre à jour les coûts et bénéfices associés à chaque action en les comparant à ce qui était initialement attendu, via ce que l'on appelle des erreurs de prédiction, afin de raffiner sa sélection pour l'avenir.

Rangel développe donc juste la première étape liée au désir du modèle de Kent Berridge et Morten Kringelbach en la décomposant en 3 sous-étapes qui permettent de prendre en compte le fait que plusieurs buts sont souvent désirés en parallèle et qu'il faut choisir parmi ces buts lesquels satisfaire en priorité. Ceci implique donc d'attribuer une valeur à chacun de ces buts pour pouvoir les évaluer sur une échelle commune et sélectionner en fonction le but à satisfaire en premier.



**Figure 20: Différentes étapes dans les mécanismes de prise de décision d'après (Rangel et al., 2008).**

Peut-être que l'une des choses qui manque à ces deux modèles est d'incorporer une étape clé entre la notion de désir et la satisfaction du désir. Cette étape correspond à la mise en place d'une série d'actions mentales et/ou motrices en vue d'atteindre le but en question. La plupart des modèles développés par la suite vont considérer cette étape mais toujours de manière séparée de l'évaluation. Ils partent tous du principe que le cerveau va d'abord effectuer une évaluation des différents plans d'action puis choisir lequel est le meilleur et s'y tenir avant de produire le mouvement (Cisek, 2012).

Un débat a cependant eu lieu entre différentes théories à propos de la nature de ce qui est évalué et du moment où un plan d'action est mis en place pour atteindre l'objet du désir. **La théorie « fondée sur les biens » (*good-based model*)** postule que les agents évaluent d'abord chaque offre indépendamment en prenant en compte les gains potentiels et les coûts associés à ces biens, dont les coûts de l'action. Ce serait seulement une fois qu'une décision sur quel bien désirer en priorité a été prise qu'un plan d'action serait mis en place en vue d'atteindre l'objet du désir (Cisek, 2012; Padoa-Schioppa, 2011). L'un des points principaux de cette théorie est de séparer d'un côté la représentation des biens qui serait utilisée pour l'évaluation et de l'autre

## Introduction

la représentation sensorimotrice de la série d'actions à effectuer en vue d'atteindre ces biens qui ne viendrait qu'après. Un point curieux de cette théorie est qu'elle propose que le coût des actions soit envisagé dans l'espace des biens avant d'envisager un plan d'action. En d'autres termes, le cerveau estimerait un coût approximatif des actions à effectuer avant même de savoir quelles actions il va entreprendre pour atteindre l'objet de son désir (Cisek, 2012; Padoa-Schioppa, 2011). Une autre théorie propose que ce qui est comparé est directement dans **le plan des actions (*action-based model*)**. Cette théorie propose que les plans d'actions liés à chaque option envisagée soient pris en compte au moment de l'évaluation (étape 2) et au moment de la sélection de la meilleure option (étape 3). Une fois la meilleure série d'action choisie, les individus la mettraient alors en route en vue d'atteindre leur objectif (Cisek, 2012; Rangel and Hare, 2010).

**Paul Cisek** a proposé une autre théorie qu'il a nommé **l'hypothèse de la compétition des moyens** (« *affordance competition hypothesis* »). Les animaux évoluent depuis toujours dans des environnements constamment changeants où il faut prendre des décisions vite pour survivre. L'idée principale de cette théorie est donc que les systèmes impliqués dans les différentes étapes évoquées précédemment œuvrent en parallèle plutôt que de manière radicalement séquentielle afin d'optimiser le temps de réaction des individus. Plus précisément, ce modèle propose que le système correspondant à l'étape (1) du modèle de Rangel, ainsi que le système permettant de se représenter la valeur associée à chacune des actions correspondant à l'étape (2) du modèle de Rangel agiraient en parallèle et en continu. On aurait ainsi des systèmes continuellement mis à jour en fonction de l'évolution des états internes de l'organisme et de l'environnement externe pour représenter, au mieux, les différentes actions possibles et les valeurs associées à chacune de ces actions, de manière dynamique. Ce modèle propose donc que d'un côté le système sensorimoteur permette de se représenter les différents plans d'actions possibles au vu de l'environnement de manière dynamique. De l'autre, le réseau des valeurs biaiserait le comportement vers un de ces plans d'actions en fonction de la valeur associée à chacun de ces plans d'actions à chaque instant. Ce fonctionnement de manière parallèle plutôt que de manière radicalement séquentielle permettrait ainsi à l'organisme de réagir de manière dynamique au cas où les contingences évoluent dans l'environnement (Cisek, 2007, 2012). L'alternative à ce fonctionnement en parallèle serait que les systèmes associés à chaque étape se mettent au repos une fois que l'étape suivante a commencé. Imaginons à l'instant  $t1$ , les parents du petit Bruce Wayne confrontés à deux options pour rentrer chez eux lors de l'étape (1) du modèle : une petite ruelle à droite (action A) ou une grande route à gauche (action B). À

## Introduction

l'étape (2), ils vont alors attribuer une valeur à chacune des actions possibles et découvrir qu'emprunter la ruelle leur permet d'arriver chez eux plus rapidement ( $V_A > V_B$ ). Une fois engagés sur cette ruelle, à l'instant  $t_2$  ( $t_2 > t_1$ ), ils vont alors apercevoir une personne en face d'eux un pistolet à la main. Selon un schéma purement séquentiel, ils vont poursuivre leur route en ignorant cette nouvelle information puisque l'action A initiée à l'étape (3) du modèle n'est pas encore achevée. Dans la vision d'un fonctionnement parallèle de Cisek, les parents de Wayne possèdent alors un nouveau schéma d'actions (étape 1) et une valeur associée à chacune de ces actions (étape 2) pour essayer de survivre à cette situation avec par exemple : A) continuer vers l'individu au pistolet d'un air calme, B) s'éloigner en marchant, C) s'éloigner en courant, D) se préparer à combattre la personne, etc. Cisek propose ainsi que l'action qui va être sélectionnée à chaque instant ne découle pas d'une décision préalable d'un système séparé qui serait dévolu à la sélection des actions, mais plutôt d'une intégration continue de l'information issue de différentes zones cérébrales. L'action sélectionnée dépendrait ainsi de biais issus du vmPFC qui enverrait des informations sur la valeur subjective associée à chaque stimulus, du cortex cingulaire antérieure qui enverrait des informations sur la valeur associée aux différentes actions possibles, du dIPFC pour connaître le contexte et les règles qui y sont associées et des ganglions de la base pour ce qui est de l'attente plus ou moins forte d'obtenir une récompense. Paul Cisek nomme ce processus un consensus distribué, dans le sens où c'est l'intégration de ces différentes entrées de manière dynamique qui, à ses yeux, permettrait la sélection d'une action, et pas une évaluation centrée dans l'espace des biens ou dans l'espace des actions (Cisek, 2012). Si une vision très simpliste en plusieurs étapes de type 1) *évaluation*, 2) *sélection*, 3) *action* semble certes séduisante et pratique pour pouvoir en faire l'étude en laboratoire, elle ne semble pas correspondre aux besoins écologiques pour lesquels nous avons été sélectionnés par les processus évolutifs. En ce sens, l'idée d'un système permettant de s'adapter continuellement et parallèlement semble plus convaincante par rapport à un système qui ferait d'abord un choix puis s'y tiendrait jusqu'à ce que l'objet de son choix ait été atteint. En effet, l'idée d'un système séquentiel, où la prise de décision précéderait et serait complètement dissociée de la mise en action fait penser à des cas qu'on aurait plutôt tendance à qualifier de pathologiques pour leur manque de flexibilité mentale, plus qu'à des comportements adaptés et sélectionnés par l'évolution. Dans tous les cas, l'idée que le cerveau fonctionne en plusieurs étapes lorsqu'il prend une décision, telles les étapes identifiées par Rangel, paraît plausible. L'idée que chacune de ces étapes soit actualisée en permanence comme le suggère Cisek paraît cependant être un mécanisme coûteux pour le cerveau.

## Introduction

Néanmoins, on en comprend vite l'intérêt au niveau évolutif. Des tâches permettant de mieux capturer la prise de décision dans des environnements évoluant de manière dynamique permettront peut-être de mieux comprendre ce point. De plus, la question de l'évaluation dans l'espace des biens ou dans l'espace des actions ou d'un mélange comme celui proposé par Cisek et des supports cérébraux de ces différentes étapes reste encore à compléter.

### III.1.b Différents systèmes d'évaluation

« *De plus, l'incontinent exécute ce que veut son appétit sans en prendre la décision, alors que le continent, à l'inverse, décide ce qu'il exécute sans en avoir l'appétit.* »

2.2.1, Livre III, *Éthique à Nicomaque*, Aristote (Aristote, 2004)

Au moins trois grands systèmes d'évaluation ont été caractérisés avec des propriétés très différentes et des bases neurales qui semblent aussi différer (Balleine et al., 2009; Dayan, 2008; Lopez-Persem, 2016; Rangel et al., 2008). Ces trois systèmes sont :

- le système **pavlovien** que nous avons déjà évoqué (voir *Psychologie et physiologie*) qui est basé sur des systèmes innés sélectionnés depuis longtemps par l'évolution et qui comprend un ensemble de réactions automatiques à des récompenses primaires (stimulus non-conditionné) ou des stimuli permettant de prédire leur survenue (stimuli conditionnés). Son rôle est de prédisposer le corps à la consommation de récompenses primaires, grâce à des réactions typiques comme la salivation en réponse à des stimuli prédisant l'obtention de nourriture, comme une odeur par exemple ;
- le système des **habitudes** : ce système serait principalement basé sur un renforcement basé sur la répétition d'une association entre un stimulus et un set d'actions comportementales qui vont mener à l'obtention d'une récompense, c'est-à-dire une réponse comportementale. Ce renforcement est basé sur un apprentissage par essai-erreur et de répétitions de l'association entre la série d'actions ayant mené à la récompense et la consommation de la récompense. Un des points clés est que ce système, contrairement à celui que nous allons évoquer après, n'implique pas de représentation d'un but. Il agirait un peu comme une machine à laver qu'on démarrerait : à la vue d'un stimulus (l'appui sur la touche démarrer), une série d'actions est entamée (mettre de l'eau à une certaine température, tourner à certains moments, etc.), mais la machine ne se représente pas le but final (laver les vêtements), bien que ce soit ce qui est obtenu à la fin.



## Introduction

Il a récemment même été proposé que ce système pourrait être complètement dissocié d'une représentation de la valeur et ne serait renforcé que par la répétition d'une série d'actions, indépendamment de leur lien ou non avec la valeur (Miller et al., 2019). Le point clé de ce système est qu'il associe à un stimulus (ou un set de stimuli) donné, une réponse automatique. Ce système serait basé sur le **striatum dorsolatéral** (Miller et al., 2019; Rangel et al., 2008; Yin and Knowlton, 2006) ;

- le système **orienté vers un but** : ce système considère la valeur des actions sur la base de la valeur subjective des récompenses qu'elles permettent d'obtenir. Face à un changement interne (satiété par exemple) ou externe (cet arbre n'a plus de fruits, ce restaurant a changé de cuisinier), il s'adapterait en réévaluant les valeurs des actions associées pour maximiser le taux de récompense. Ce système se différencie d'un côté du système des habitudes puisqu'il est orienté vers un but contrairement au système des habitudes. Il se différencie aussi du système pavlovien en ce qu'il implique un comportement instrumental actif, là où les réactions du système pavlovien sont des réactions passives à des contingences de l'environnement.

Une preuve, parmi d'autres, suggérant que ces systèmes peuvent en effet être dissociés est que, par exemple, dans le cas des patients souffrant du syndrome de perte d'auto-activation psychique que nous avons déjà évoqué (voir *Impact des lésions du réseau des valeurs*), les patients se trouvent incapables d'initier une action par eux-mêmes en vue d'un but, alors qu'ils sont capables de le faire si on leur demande de le faire (Schmidt et al., 2008), et qu'ils sont capables de s'engager dans des attitudes répétitives et stéréotypiques sans vraiment savoir pourquoi (Laplane and Dubois, 2001). Une hypothèse possible est que le système des habitudes de ces patients est intact alors que leur système orienté vers un but serait lésé. Les patients en question continueraient d'associer le fait que, dans le contexte où ils se trouvent, il faut faire ce que dit l'expérimentateur quand on leur donne des instructions, alors même qu'ils n'arrivent pas à se représenter un but justifiant de le faire.

De plus, plusieurs expériences sur des modèles animaux, comme chez l'homme, ont permis de dissocier ces différents systèmes au niveau cérébral sur la base d'expériences où différentes zones du cerveau sont lésées et où la récompense obtenue à la suite d'une action est dévaluée. Quand la récompense est dévaluée, le système orienté vers un but devrait s'adapter en dévaluant aussi la valeur de l'action qui y est associée, alors que le système des habitudes,

## Introduction

ayant été fortement renforcé à effectuer l'action en question va continuer de pousser l'individu à effectuer l'action même quand elle est dévaluée (Rangel et al., 2008). Le cortex orbitofrontal et le dlPFC seraient impliqués dans la motivation orientées vers des buts et capables de s'adapter aux changements de contingences dans l'environnement, alors que le système des habitudes serait plutôt lié au striatum dorsal. Le système pavlovien, lui, serait lié au striatum ventral, l'amygdale basolatérale et aussi le cortex orbitofrontal (Daw et al., 2005; Rangel et al., 2008). Quoiqu'il en soit, au vu des techniques actuelles de mesure de l'activité cérébrale, il paraît difficile, chez l'homme, de différencier le système des habitudes et le système pavlovien qui semblent tous deux liés au réseau des valeurs, notamment au striatum ventral. L'évolution des outils de mesure permettra peut-être de mieux distinguer les différentes parties du striatum ventral selon leur appartenance à chacun de ces deux systèmes des valeurs.

En résumé, il existerait au moins trois systèmes impliqués dans les processus motivationnels. Le premier, le système pavlovien, permettrait des réactions automatiques et typiques face à différents types de stimuli sélectionnés par l'évolution. Par exemple, la peur des araignées et des serpents serait innée vu qu'on peut observer des réactions déjà chez des enfants de 6 mois (Hoehl et al., 2017). De même, le sucre en cas de baisse de la glycémie ou l'eau en cas de soif serait des stimuli déclenchant spontanément une réaction d'approche. Par apprentissage par renforcement, ces stimuli non-conditionnés peuvent être associés à d'autres stimuli, les stimuli conditionnés, qui vont une fois l'association opérée, eux aussi provoquer ces réactions typiques. Le deuxième système, le système par habitudes, est principalement basé sur un renforcement par répétitions qui permet d'associer un plan d'actions à un stimulus. Le troisième enfin, le système orienté vers un but, qu'on appelle aussi basé sur un modèle du monde, est basé sur l'idée qu'il adapte de manière flexible un modèle interne du monde en vue d'optimiser continuellement les coûts et les bénéfices liés au fait d'atteindre un but.

### **III.2 Choix binaires et modèles**

Probablement du fait que la neuroéconomie et l'étude de la prise de décision sont héritières de l'économie, mais aussi parce que c'est les cas de choix les plus simples à étudier, la neuroéconomie s'est principalement penchée sur l'étude des choix binaires entre une option A et une option B. En particulier, si l'on se réfère au modèle d'Antonio Rangel développé précédemment (voir *Étapes de la motivation*), les recherches se sont concentrées sur les étapes (2) et (3) des processus motivationnels, c'est-à-dire sur l'évaluation des options présentée lors d'un choix binaire, les valeurs  $V_A$  et  $V_B$  associées aux options A et B, et la sélection subséquente

## Introduction

entre les deux options selon la différence entre  $V_A$  et  $V_B$  (choix de A si  $V_A > V_B$  et réciproquement).

### III.2.a Modèle de diffusion

Historiquement, la modélisation de choix binaires commence avec la **théorie de la détection du signal** en 1954 (Peterson et al., 1954; Tanner and Swets, 1954). Le but principal de cette théorie est de modéliser les choix des individus dans un cas où ils doivent dire si oui ou non ils ont détecté un signal sur la base d'informations perceptuelles (la vision en l'occurrence) avec de l'incertitude liée à la difficulté de percevoir le signal. Par la suite, un modèle qui va devenir très populaire pour expliquer la prise de décision dans le domaine perceptuel s'appelle le **modèle de diffusion** (*drift-diffusion model*). Les bases de ce modèle apparaissent dans un article de Roger Ratcliff de 1978 (Ratcliff, 1978). L'un des principes fondamentaux de ce modèle est l'idée que de l'information est accumulée au cours du temps en faveur de chacune des options à disposition jusqu'à ce qu'un seuil soit atteint pour l'une des options ce qui va déterminer laquelle des options sera choisie. Le modèle va ensuite se développer progressivement pour intégrer plusieurs modalités (Ratcliff et al., 2016). Ce modèle va particulièrement être utilisé pour modéliser les choix et les temps de réaction dans des tâches basées sur la perception. Les deux composantes principales de ce modèle sont 1) un **seuil**  $a$  qui détermine à partir de quel niveau d'information, c'est-à-dire de confiance, le choix peut être effectué, 2) le **taux d'accumulation d'information**  $v$  quant à l'orientation des points et 3) le **point de départ**  $z$  qui peut être biaisé ou non vers l'une des options avant même que l'accumulation commence (Ratcliff, 1978). Au cours des dernières décennies, différentes variantes de ce modèle ont été établies (voir **Figure 21** et (Ratcliff et al., 2016)).



**Figure 21 : Schéma général du modèle de diffusion (*drift-diffusion model*).** Au cours du temps, de l'évidence est accumulée en faveur de chacune des options A et B présentées dans le choix. Dans cet exemple, l'évidence pour choisir B plutôt que A atteint le seuil de décision plus rapidement et l'individu va donc choisir l'option B. Les composantes principales du modèle sont 1) **le seuil de décision**, c'est-à-dire la distance  $a$  entre le trait en pointillés et les traits associés aux options A et B ; 2) **le taux de diffusion (*drift rate*)**  $v$  qui correspond à la vitesse où l'information est accumulée et qui va déterminer la vitesse à laquelle l'un des deux seuils va être atteint. Par la suite, différentes variantes ont été introduites, avec, de ce fait, des paramètres additionnels : 3) **le point de départ (*starting point*)**  $z$  qui peut être biaisé vers l'une des options. Dans certaines variantes, ce point peut varier entre plusieurs essais auquel cas il faut aussi définir une distribution avec une moyenne  $\mu_z$  et une variance  $\sigma_z$  dont le point de départ serait issu à chaque essai ; 4) **le temps de non-décision (*stimulus encoding time*)** qui correspondrait à un temps où les stimuli sont encodés au niveau sensoriel avant que l'accumulation d'information qui va servir à la prise de décision soit initiée ; 5) **le temps de l'exécution motrice** qui serait lié à la réponse motrice du fait que dans ce genre de tâche une action motrice (appuyer sur un bouton généralement) est nécessaire pour répondre et que ce temps est considéré comme étant postérieur à la prise de décision ; 6) **un paramètre  $r$  indiquant comment les seuils se réduisent au cours du temps** (Milosavljevic et al., 2010); et probablement encore plusieurs autres paramètres que j'ai oublié d'évoquer ici tellement il existe de variantes qui ont été conceptualisées. Certaines variantes notamment ne modélisent qu'un seul accumulateur au lieu de deux (Ratcliff et al., 2016). Figure issue de (Dutilh and Rieskamp, 2016).

## Introduction

A partir des années 2010, le modèle de diffusion que l'on vient de présenter va être de plus en plus utilisé dans le cas de la prise de décision basée sur les valeurs (Basten et al., 2010; Krajbich and Rangel, 2011; Krajbich et al., 2010; Milosavljevic et al., 2010). Le cerveau accumulerait de l'information à propos de chaque option au cours du temps, sur la base des états internes (faim, soif, etc.), d'indices dans l'environnement (estimation des différentes composantes des aliments en termes de différents nutriments sur la base de leur aspect visuel (Suzuki et al., 2017) ou en prenant en compte l'impact des aliments en terme de santé (Hare et al., 2011a) par exemple) et de la mémoire (souvenir plus ou moins agréable la dernière fois que l'on a mangé quelque chose de similaire). Comme pour la décision basée sur la perception, une fois qu'un seuil serait atteint pour une des options, la décision est prise. Ce processus permet ainsi de décrire le processus dynamique qui donne lieu à l'évaluation (étape (2) du modèle de Rangel, voir *Étapes de la motivation*) et à la sélection (étape (3) du modèle de Rangel, voir *Étapes de la motivation*).

Au **niveau neuronal**, Wang a proposé un modèle dans lequel différents neurones encoderaient l'évidence pour chaque option. **Par compétition et inhibition réciproques**, une des deux populations de neurones finirait par l'emporter permettant ainsi à l'individu de savoir laquelle des options choisir (Wang, 2002). Ce modèle a reçu des preuves dans la prise de décision basée sur la perception, notamment via les travaux de l'équipe de Michael Shadlen chez le primate non-humain qui a montré que l'aire latérale intrapariétale faisait typiquement ce type de calcul (Huk and Shadlen, 2005; Wong, 2007). Ce résultat a cependant été critiqué par l'équipe de Jonathan Pillow qui a montré que ce résultat était lié au fait de moyenniser l'activité des neurones à travers plusieurs essais plutôt que de regarder l'activité essai par essai. Au niveau de chaque essai, le cortex pariétal ne ferait pas ce type de calcul puisque, pour un essai donné, les neurones réagiraient plutôt de manière binaire, c'est-à-dire qu'ils passent, d'un coup, à un niveau d'activité prédisant si l'individu va choisir de répondre à gauche ou à droite, mais il n'y a pas d'accumulation vers l'une ou l'autre option (Latimer et al., 2015). Ce résultat invalide l'idée d'accumulateur dans le cortex pariétal, mais il est toujours possible que de l'accumulation ait lieu dans d'autres zones du cortex.

A l'heure actuelle, la question se pose encore de savoir si dans le cas de la prise de décision basée sur la valeur, une (ou plusieurs) aire(s) cérébrale(s) effectuent aussi une accumulation dynamique pour chaque option afin de pouvoir prendre une décision, comme proposé dans ce type de modèle. Dans le cas de la prise de décision sur la base d'informations perceptuelles, la source principale d'information est basée sur l'environnement même si des *a*

## Introduction

*priori* basés sur les essais précédents peuvent jouer dans l'estimation des informations perceptuelles. Les populations neurales en compétition sont ainsi liées aux variables observées dans l'environnement. Dans le cas de la prise de décision basée sur les valeurs, par contre, deux grandes hypothèses s'opposent. La première hypothèse est que les valeurs associées à chaque action seraient apprises par l'expérience et **mémorisées**. Il suffirait donc d'extraire de la mémoire la valeur associée à chaque option à disposition de la mémoire pour pouvoir prendre une décision. La compétition au niveau neural aurait alors lieu sur la base de l'information stockée en mémoire pour la valeur associée à chaque option. La seconde hypothèse part du principe que **les valeurs sont systématiquement recalculées** de manière dynamique quand une prise de décision doit être effectuée. C'est cette seconde hypothèse qui semble désormais admise par la plupart des chercheurs autant sur la base d'évidence expérimentale qui montre que la valeur subjective dépend des états internes (faim/soif/satiété/etc.) que sur la base de l'intuition (Padoa-Schioppa, 2011). Dans tous les cas, les neurones du réseau des valeurs sont régulièrement supposés être le lieu où s'effectueraient ce type de calcul dynamique. Différentes études chez le primate non-humain suggèrent en effet que les neurones du vmPFC (Strait et al., 2014), du cortex orbitofrontal latéral (Ballesta and Padoa-Schioppa, 2019) et même du striatum ventral (Strait et al., 2015) pourraient en effet fonctionner comme prédit par la théorie de Wang. Différentes populations neuronales encoderaient la valeur de chacune des options considérées. Par inhibition mutuelle entre les différentes populations, la population où l'encodage de la valeur est le plus élevé, c'est-à-dire la population pour laquelle la valeur est la plus élevée, finirait par gagner la compétition signalant ainsi au reste du cerveau quelle option il faut choisir.

### III.2.b Modèles de choix alternatifs

D'autres modèles relativement proches existent pour expliquer comment se construit dynamiquement la prise de décision basée sur la valeur. Nous ne les évoquerons pas tous pour des raisons de concision car la liste est très longue (voir (Ratcliff et al., 2016) pour une liste relativement exhaustive des grandes familles de modèles de prise de décision dynamiques), mais nous allons mentionner les spécificités de certains de ces modèles. Cela nous permettra d'avoir une idée générale des différentes propositions existantes quant à la modélisation du processus dynamique qui donne lieu à un choix.

- **Modèles de course (*race models*)** : le principe est le même que le modèle de diffusion, deux accumulateurs vont accumuler de l'information sur la valeur de chaque option et le premier accumulateur qui atteint le seuil va déclencher la décision. La seule différence avec le modèle de diffusion est que ces modèles partent du principe que la vitesse

d'accumulation d'information pour chaque accumulateur est indépendante (Gold and Shadlen, 2007).

- **Modèle d'urgence (*urgency-gating model*)** : nous l'avons déjà évoqué, ce modèle, développé initialement par Paul Cisek (Cisek et al., 2009; Thura et al., 2012), est conceptuellement similaire au modèle de diffusion qui inclue des seuils évoluant dynamiquement avec le temps (Hawkins et al., 2015). L'idée de ce modèle est que le temps réduit les seuils de décision indépendamment de la position des accumulateurs d'évidence en faveur des différentes options.
- **Modèle de diffusion attentionnel (*attentional drift diffusion model*)** : le principe de cette variante est que l'accumulateur d'évidence accumule de l'évidence pour l'option qui est en train d'être observée. L'idée sous-jacente de ce modèle, principalement porté par les équipes d'Ian **Krajbich** et d'Antonio **Rangel**, est que la valeur subjective dépend de l'attention. Un poids plus grand serait donné à l'option qui est en train d'être observée. En forçant les individus à regarder chaque option pour un temps déterminé et à des moments déterminés au cours de la prise de décision, ils ont montré que leur modèle pouvait prédire les choix effectués par les individus (Krajbich and Rangel, 2011; Krajbich et al., 2010). Ils ont aussi montré que l'activité du cortex orbitofrontal, une des composantes clés du réseau des valeurs comme nous l'avons déjà vu (voir *L'anatomie du réseau des valeurs*), chez l'homme (Lim et al., 2011) comme chez le primate non-humain (McGinty et al., 2016) dépendait de l'attention prêtée à chaque stimulus. Ces résultats suggèrent donc que, non seulement l'attention dépend de la valeur subjective ce qui a été montré par ailleurs (Anderson, 2016, 2019; Anderson et al., 2011), mais aussi que **la valeur subjective peut être biaisée par l'attention**. À l'heure actuelle, il est encore difficile de savoir par quels mécanismes cette association a lieu. Il se pourrait que ce soit lié à une forme d'apprentissage au cours de la vie : de la même manière que l'on a tendance à regarder ce qui a plus de valeur étant donné que l'attention est biaisée par la valeur, on finirait par accorder plus de valeur à ce qu'on regarde, même quand cela n'est pas associé à l'obtention d'une récompense. Il pourrait aussi s'agir d'un biais qui est inné mais il faudrait essayer de comprendre quelles pressions évolutives ont pu donner lieu à ce biais.

Une des limites de tous ces modèles est que, bien qu'ils permettent certes de modéliser les temps de réaction et les choix sur la base des valeurs assignées à chaque option, ils ne permettent pas de faire de prédiction sur comment l'information accumulée pour chaque option

## Introduction

est transformée en valeur subjective. En résumé, ils décrivent comment la valeur est accumulée au cours du temps jusqu'à prendre une décision, mais pas sur quelles bases elle est construite.

### III.2.c Une représentation « efficace » des valeurs ?

En 1961, Horace **Barlow**, dans son article *Possible Principles underlying the Transformations of Sensory Messages*, proposait que les neurones du système sensoriel fonctionnent pour **représenter « efficacement » l'information**, c'est-à-dire en minimisant l'encodage des informations redondantes et en détectant et en amplifiant les informations qui sont pertinentes pour la survie de l'animal (Barlow, 2012). L'idée principale de la représentation efficace (*efficient coding*) de l'information est de maximiser la quantité d'information véhiculée en minimisant le nombre d'unités nécessaires, c'est-à-dire le nombre de potentiels d'action des neurones dans le cas du cerveau, pour la véhiculer. Cette théorie doit aussi beaucoup au développement de la théorie de l'information qu'on fait généralement remonter à l'article de 1948 de Claude **Shannon** sur la théorie de la communication (Shannon, 1948). Cette proposition théorique a depuis été validée dans le domaine sensoriel par plusieurs études qui ont pu montrer que les aires cérébrales liées à l'encodage des informations sensorielles fonctionnaient bien selon une représentation « efficace » de l'information compatible avec la théorie de Shannon (Gervain and Geffen, 2019; Ma and Wu, 2011; Ming and Holt, 2009). Là encore, comme pour le modèle de diffusion, les chercheurs en neuroéconomie vont s'inspirer des travaux et des théories issues du domaine de la perception.

En 2012, Kenway **Louie** et Paul **Glimcher** vont proposer que la représentation de la valeur répond elle aussi à un encodage « efficace » (Louie and Glimcher, 2012). L'une des propositions qu'ils vont faire dans ce sens est que les neurones encodant la valeur lors d'un choix encodent la valeur d'un objet donné via une « **normalisation divisive** ». La « normalisation divisive » consiste à diviser la valeur de l'option considérée par la somme de toutes les valeurs disponibles :  $SV_A = \frac{V_A}{\sum_i V_i}$  où  $SV_A$  est la valeur subjective de l'option A après la normalisation,  $V_A$  est la valeur de A avant la normalisation et  $\sum_i V_i$  la somme des valeurs de toutes les options à disposition. L'idée est que les neurones ayant des capacités finies en termes de quantités de potentiels d'action qu'ils peuvent émettre par unité de temps, ils ont besoin de réadapter l'encodage de l'information en fonction du rang des valeurs à évaluer. Pour estimer le rang des valeurs, Louie et Glimcher proposent aussi que, non seulement les options présentes, mais aussi le **contexte récent** influe sur cet encodage. Par exemple, imaginons une situation où j'ai envie de boire du vin et que je viens de boire un très bon vin. Il n'y a plus de très bon vin à



## Introduction

disposition, je considèrerai alors l'idée de boire un verre de mauvais vin différemment par rapport à une situation où je viens de boire d'un autre mauvais vin par exemple. Depuis leur proposition, Louie et Glimcher ont montré que les choix de sujets humains sains pouvaient être expliqués par cette théorie (Khaw et al., 2017; Louie et al., 2013). Plusieurs études chez le primate non-humain ont montré en outre que les neurones du cortex orbitofrontal médian et/ou latéral adaptaient leur activité en fonction des récompenses potentielles dans l'environnement (Conen and Padoa-Schioppa, 2018; Kobayashi et al., 2010; Padoa-Schioppa, 2009; Yamada et al., 2018).

L'une des limites de cette théorie, reconnue dès le début par les auteurs, est qu'il est difficile de savoir *a priori* quels sont les paramètres qui déterminent la distribution des valeurs à partir de laquelle le système des valeurs va effectuer sa normalisation (Louie and Glimcher, 2012). En particulier, il paraît difficile de savoir sur quelle **échelle de temps** le système des valeurs décide d'établir cette distribution. En effet, en laboratoire dans des conditions contrôlées, l'expérimentateur définit *a priori* l'échelle de temps qu'il va considérer pour son analyse. Il peut ainsi définir un certain nombre d'essais précédents sur lesquels la normalisation va avoir lieu par exemple. Cela ne nous dit pas cependant si le cerveau utilise en effet la même échelle de temps. De plus, l'intégration pourrait se faire sur une échelle de temps fixe ou sur une échelle de temps dynamique. Si l'échelle de temps est dynamique, il faudra aussi chercher à comprendre sur la base de quels critères la durée d'intégration est sélectionnée par le cerveau. On peut d'ailleurs facilement imaginer des interactions entre la valeur et l'échelle de temps considérée. Par exemple, un item avec une valeur très élevée, qu'il soit appétitif ou aversif, pourrait impacter la distribution des valeurs pour un temps plus long qu'un item de valeur faible car il reste présent dans l'espace mental. Cela reste cependant à vérifier. De plus, il a récemment été montré que, dans un cas de choix ternaire, une modulation de l'attention expliquait mieux les choix effectués par les participants en comparaison à une normalisation divisive de chaque option par la somme des valeurs des options (Gluth et al., 2020). Le débat reste donc ouvert pour mieux comprendre les mécanismes des choix et comment la valeur de chaque option est considérée en fonction du contexte.

### III.2.d Bases neurales du choix

Dans le cas des choix binaires, le réseau des valeurs, en particulier le **vmPFC**, a été identifié en neuroimagerie comme corrélant à différents paramètres liés au choix. Le vmPFC a pu être observé en corrélation avec la **valeur de l'option choisie** (Arulpragasam et al., 2018; Larsen and O'Doherty, 2014; Seaman et al., 2018; Wunderlich et al., 2010), **la différence entre**

**la valeur de l'option choisie et la valeur de l'option non-choisie** (Boorman et al., 2009; Jocham et al., 2012), la différence en valeur absolue entre les deux options (De Martino et al., 2013) ou encore avec **la valeur du choix**, c'est-à-dire la somme (ou la moyenne) des valeurs des options présentées (Hare et al., 2011b; Wunderlich et al., 2010). Certaines études n'ont, à l'inverse, pas observé de lien entre le vmPFC et la valeur non-choisie ou la différence entre la valeur choisie et la valeur non-choisir (Wunderlich et al., 2010). Il existe d'ailleurs probablement un biais de publication qui masque tous les résultats où le vmPFC n'est pas observé avec ces paramètres. Ces résultats divergents ne sont d'ailleurs pas si choquants si l'on considère le rôle du vmPFC et du réseau des valeurs. En effet, si l'on accepte, comme dans la théorie de Berridge (voir *Histoire de la découverte des centres cérébraux du plaisir*), que le réseau des valeurs, et en particulier le vmPFC, se spécialise dans l'incitation vers un plaisir, c'est-à-dire le désir, il est dur de savoir ce qu'il fait dans les choix binaires tels que ceux auxquels il est soumis en laboratoire qui ne sont peut-être pas les plus écologiques. En effet, il doit alors à la fois représenter la désirabilité de chaque item présenté et, à la fois, classer les items sur la base de cette désirabilité pour pouvoir prendre une décision. De plus, dans le contexte des choix de laboratoire, choisir l'une des options revient à perdre l'autre option. Autrement dit, il ne suffit pas de prendre en compte la désirabilité escomptée de l'option choisie, mais aussi de prendre en compte le regret éventuel associé au fait de ne pas avoir choisi l'autre option. Les mécanismes pourraient en effet alors être légèrement différents que dans un cas où il s'agirait par exemple de définir l'ordre dans lequel deux récompenses vont être consommées sans que le choix de l'une implique de ne plus du tout avoir accès à l'autre.

En 2012, sur la base de résultats en magnétoencéphalographie, Laurence Hunt a proposé que l'activité du vmPFC évolue de manière dynamique au cours du choix. Il a fondé cette proposition sur les prédictions théoriques du modèle de Wang (voir *Modèle de diffusion*) où deux populations neuronales, chacune associée à l'encodage de la valeur subjective d'un des deux stimuli présentés, sont en compétition. **Le vmPFC encoderait ainsi d'abord la valeur du choix, basée sur la somme des options présentées, puis, par inhibition mutuelle, il se mettrait à représenter la différence de valeurs entre les options présentées** (Hunt et al., 2012). Hunt et ses collaborateurs ont depuis partiellement répliqué ce résultat en IRMf (Jocham et al., 2014). Un pattern similaire a aussi été observé dans le striatum ventral de primates non-humains (Cai et al., 2011). La résolution temporelle de l'IRMf (de 1 à 3 secondes selon les paramètres en général) permet difficilement d'observer cette dynamique et cela expliquerait peut-être pourquoi les résultats divergent autant entre les études.

## Introduction

D'un autre côté, comme nous l'avons déjà évoqué (voir *Impact des lésions du réseau des valeurs*), des lésions au vmPFC n'empêchent pas les sujets d'effectuer des choix basés sur la valeur des options présentées, mais elles impactent seulement la cohérence des sujets dans leurs choix. Le rôle causal du vmPFC dans les choix et son rôle précis restent donc encore à déterminer. De plus, le vmPFC est impliqué dans l'humeur et la sensation subjective d'être heureux ou satisfait et l'humeur elle-même peut impacter le comportement (Vinckier et al., 2018). Il se pourrait ainsi que le vmPFC ne joue pas de rôle causal dans le choix, essai par essai, mais plutôt un rôle passif. Le sujet serait alors juste d'autant plus heureux que les options présentées sont intéressantes ( $V_A + V_B$  élevé), que le choix était facile ( $|V_A - V_B|$  élevé) et que l'option choisie est intéressante ( $V_{option\ choisie}$  élevée), sans pour autant avoir un rôle direct dans le choix qui va être effectué autrement que par un effet lié à l'humeur. On a déjà vu en effet que l'activité basale du vmPFC pouvait avoir un impact sur la valeur perçue d'une option et sur les choix effectués (voir *Le réseau des valeurs et ses propriétés*). Son rôle reviendrait alors à refléter un état global de satisfaction lié à l'environnement présent. Cela permettrait de savoir si cela vaut le coup de continuer à profiter de l'environnement présent ou à chercher à changer d'environnement. Le vmPFC est d'ailleurs régulièrement identifié comme une zone s'activant avec la valeur de rester dans un environnement donné par rapport à changer d'environnement (Gläscher et al., 2009).

## **IV Les variables méta-décisionnelles : le coût décisionnel et la confiance**

Comme nous l'avons vu au chapitre précédent, la prise de décision implique un processus dynamique en plusieurs étapes et accompagné de bruit dans l'estimation de la valeur de chaque item qui est évalué. De ce fait, en particulier dans le cas où un choix est nécessaire entre deux items de valeurs proches, le cerveau a besoin d'être capable d'effectuer un arbitrage pour savoir à quel moment il doit prendre une décision. En effet, la prise de décision en elle-même pourrait s'accompagner d'un coût que ce soit en termes de savoir combien de temps allouer à une décision donnée (coût d'opportunité) ou de savoir quelle quantité de ressources neurales allouer à la prise de décision (coût des fonctions exécutives). Dans ce cas, le cerveau a besoin d'être capable d'effectuer un arbitrage pour savoir quel coût il est prêt à dépenser pour une décision donnée et en vue de quels bénéfices ce coût va être dépensé. Le bénéfice retiré de la décision pouvant consister aussi bien dans la récompense que l'on espère obtenir à la suite de la décision, que dans le gain de confiance dans la décision obtenue suite à l'allocation de plus de ressources.

## Introduction

Dans cette partie, nous allons étudier plus en détails deux variables décisionnelles : le coût associé à la prise de décision et la confiance associée à la décision. Nous allons, de plus, chercher à mieux comprendre les bases neurales de chacune de ces variables méta-décisionnelles.

### IV.1 La confiance dans la décision

Pour apprendre dans quelle mesure l'association entre une action ou un événement passif et une récompense ou une punition est pertinente (dernières étapes du modèle de Rangel vu précédemment, voir *Étapes de la motivation*), les organismes disposent de plusieurs stratégies. Dans le cas du système des habitudes et du système pavlovien (voir *Différents systèmes d'évaluation*), tout événement spatio-temporellement concomitant à une récompense peut y être associé. L'association, c'est-à-dire l'apprentissage, sera principalement basée sur la **répétition**, surtout dans le cas des récompenses, et sur la **saillance**, notamment dans le cas des punitions où une seule association peut suffire à créer une association (Izquierdo et al., 2016), comme dans la peur par conditionnement ou dans les expériences de stress post-traumatique. Pour le système orienté vers des buts, l'apprentissage va être modulé grâce à une composante additionnelle. Cette composante permet aux individus d'estimer dans quelle mesure les actions qu'ils ont effectués sont bien les plus efficaces pour obtenir des événements positifs et éviter des situations aversives : la **confiance**. La confiance est peut-être aussi engagée dans les deux autres systèmes motivationnels, mais cela reste à prouver. De plus, lorsque les individus prennent une décision donnée, il semblerait qu'ils agissent non pas seulement en vue de maximiser leurs gains, mais aussi de maximiser leur confiance dans leur décision. Ce phénomène se reflète dans les modèles dynamiques de prise de décision que nous avons vus précédemment (voir *Choix binaires et modèles*) où la décision n'est prise qu'une fois qu'un seuil de confiance dans la décision est atteint. Dans cette partie, nous allons ainsi nous intéresser à cette composante majeure du système motivationnel orienté vers des buts qui permet potentiellement de définir un seuil de décision.

#### IV.1.a La confiance

Nous avons déjà étudié comment les individus et leurs cerveaux estiment l'incertitude, autant celle liée au risque que celle liée à l'ambiguïté (voir *L'incertitude*). La confiance cependant se distingue de l'estimation de l'incertitude. En effet, l'incertitude est liée uniquement à l'estimation des probabilités dans l'environnement externe, alors que la confiance est un processus dit « métacognitif », c'est-à-dire un jugement sur sa propre réflexion. La

## Introduction

définition que nous allons adopter ici pour la confiance est de la considérer comme **l'estimation subjective de la probabilité d'avoir donné une réponse correcte** (Pouget et al., 2016). Plusieurs définitions et sous-catégories existent pour la confiance, mais nous n'allons pas nous aventurer à les développer ici pour des raisons de concision et afin de nous concentrer sur la confiance dans la décision qui vient pendant ou après la décision. Le lecteur est invité à consulter (Bioud, 2019; Meyniel et al., 2015) pour une vision plus détaillée.

Comme suggéré dans l'introduction de cette partie, il semblerait que la confiance soit utilisée comme une métrique permettant de savoir dans quelle mesure le comportement doit ou non être renforcé ou au contraire modifié pour l'avenir. Il a été démontré par exemple que dans le cas d'un choix où la confiance finale dans la décision était faible, les individus avaient plus de chance de changer d'avis la seconde fois où le choix était présenté (Folke et al., 2016). Dans la même logique, il a été montré récemment que les individus adaptaient leur comportement en fonction de leur niveau de confiance dans une tâche impliquant de l'exploration : quand leur confiance dans le fait que la stratégie d'exploitation était la meilleure était faible, ils se remettaient à explorer l'environnement (Boldt et al., 2019). Il a aussi été montré que, quand la confiance dans un choix donné est faible, les individus ont tendance à adapter leur ratio vitesse/précision lors des essais suivants afin d'améliorer leur précision, c'est-à-dire la confiance espérée, quitte à réduire leur vitesse (Desender et al., 2019). Tous ces résultats confirment bien que la confiance joue un rôle dans l'apprentissage et dans le comportement en aidant l'individu à adapter son comportement. Le comportement futur peut être adapté, quand la confiance est faible, soit en changeant de stratégie soit en la modifiant lorsque la confiance est faible.

Les individus, en plus de chercher à optimiser la valeur obtenue lors de leurs actions et de leurs choix, semblent aussi chercher à maximiser leur confiance dans leurs décisions. Ce phénomène se reflète dans les modèles dynamiques de prise de décision que nous avons vus précédemment (voir *Choix binaires et modèles*) où la décision n'est prise qu'une fois qu'un seuil de confiance dans la décision est atteint. Il est généralement admis que les individus cherchent à maximiser un **compromis vitesse/précision** dans leurs décisions. Moins ils sont sûrs de la bonne décision à prendre, plus ils vont passer de temps à se décider et réciproquement. Cela explique pourquoi des évaluations de la confiance sont régulièrement négativement corrélées avec les temps de réactions dans la décision (de Gardelle et al., 2016; Lebreton et al., 2015; Rahnev et al., 2020; Weidemann and Kahana, 2016). Il a récemment été proposé par des membres de notre équipe que les individus chercheraient à optimiser le **ratio confiance/effort**

## Introduction

en essayant de minimiser l'effort mental effectué lors d'une prise de décision par rapport au gain de confiance dans la décision qu'ils espèrent obtenir du fait de cet effort (Lee and Daunizeau, 2019). Les liens entre confiance et temps de réaction ne sont cependant pas unilatéraux puisqu'il semblerait, d'après une étude chez 6 participants, que la confiance subjective consiste en une estimation du temps mis à répondre, indépendamment de la performance (Kiani et al., 2014). Cette proposition reste néanmoins à confirmer dans des groupes de participants plus importants pour évaluer dans quelle mesure biaiser le temps de réponse sans changer les autres variables permet en effet de faire varier l'estimation subjective de confiance.

En résumé, la confiance dans la décision pourrait jouer aussi bien sur les décisions en cours en vue de favoriser les ajustements permettant de la maximiser

### IV.1.b Y a-t-il une zone cérébrale de la confiance ?

Nous avons déjà étudié les aires cérébrales liées à l'estimation subjective de l'incertitude dans l'environnement (voir *Les bases neurales de l'incertitude*). Le dmPFC est une zone clé dans ce processus. Pour l'estimation de la confiance, par contre, on retrouve le réseau des valeurs avec, notamment, le **vmPFC**, en corrélation avec la confiance dans des tâches de choix binaires basées sur la valeur (De Martino et al., 2013; Shapiro and Grafton, 2020), la mémoire (Chua et al., 2006; Kuchinke et al., 2013; Moritz et al., 2006) ou sur la perception (Fleming et al., 2018; Gherman and Philiastides, 2018; Rolls et al., 2010) et aussi dans des tâches d'évaluation où il n'y a pas de choix à effectuer (Lebreton et al., 2015). Une première conclusion serait donc de dire que le vmPFC encode non seulement la valeur subjective dans une dimension commune, peu importe ce qui est considéré (voir *Bases neurales de la valeur*), mais qu'il en fait aussi de même pour la confiance qu'il encoderait indépendamment du type de tâche effectuée. L'idée d'une **échelle commune** pour la confiance, indépendamment de la tâche considérée, semble de plus confirmée au niveau comportemental (de Gardelle and Mamassian, 2014; de Gardelle et al., 2016). Une méta-analyse a d'ailleurs identifié le vmPFC comme corrélant avec la confiance à travers plusieurs études regroupant différents types de tâches (Vaccaro and Fleming, 2018).

Cependant, les choses n'ont pas l'air si simples. Tout d'abord, on pourrait argumenter au moins deux choses :

- 1) le vmPFC encode la **valeur subjective**, il pourrait donc n'être activé quand la confiance est plus élevée que parce que l'individu est aussi plus heureux parce qu'il

## Introduction

a plus confiance dans sa décision. Cela dit, conceptuellement, la confiance pourrait aussi ne pas être autre chose que justement le plaisir ressenti à avoir une probabilité subjective plus élevée d'avoir pris la bonne décision. Cela expliquerait pourquoi les individus cherchent à maximiser leur confiance ;

- 2) Dans beaucoup de tâches où le vmPFC est identifié comme corrélant avec la confiance, il le fait quand on demande au sujet de rapporter sa confiance dans la décision. L'activation du vmPFC pourrait de ce fait n'être due qu'au fait de devoir rapporter son ressenti subjectif de confiance et non au ressenti en lui-même. Cependant, plusieurs études montrent que son activité corrèle avec la facilité d'une décision même quand aucune question sur la confiance n'est posée (Lebreton et al., 2015; Rolls et al., 2010), ce qui confirme qu'il encode la confiance de manière automatique, indépendamment de si elle est, ou non, demandée.

Enfin, plusieurs études ont identifié des **clusters très différents** en corrélation avec la confiance et, ironiquement, il s'agit de travaux venant parfois des mêmes équipes que celles ayant identifié le vmPFC en corrélation avec la confiance. Par exemple, dans le cas de choix basés sur la perception, certains auteurs ont identifié un cluster légèrement plus dorsal et postérieur que le vmPFC en corrélation avec la confiance rapportée par les sujets. Ce cluster serait situé dans le **cortex cingulaire antérieur périgénual**, c'est-à-dire juste devant le genou du corps calleux (Bang and Fleming, 2018; Gherman and Philiastides, 2018). D'autres ont identifié d'autres zones **sans que le vmPFC soit impliqué**. Par exemple, plusieurs études en lien avec la mémoire identifient un cluster dans le **cortex préfrontal ventrolatéral** (Chua et al., 2009; Qin et al., 2011), ou encore une modulation de l'activité de l'**hippocampe** ou du **cortex pariétal** avec la confiance (Hayes et al., 2011), sans aucune activation du vmPFC détectée. Quelques études ont aussi identifié un autre composant du réseau des valeurs, le **striatum ventral**, en lien avec la confiance, mais pas le vmPFC (Hebart et al., 2016; Molenberghs et al., 2016). Enfin, une méta-analyse pointe vers le fait que le cluster associé à la confiance serait plus dorsal que le vmPFC sans pour autant être collé au genou du corps calleux comme suggéré précédemment (White et al., 2014). Ces résultats ont aussi été confirmés par une étude qui montre que, dans une tâche d'évaluation, on pouvait identifier **un gradient dans le cortex préfrontal médian avec un encodage plus marqué pour la valeur subjective au niveau ventral et l'inverse pour la confiance** (voir **Figure 22** et (De Martino et al., 2017)). Ceci est d'ailleurs cohérent avec le résultat de plusieurs études qui montrent que deux clusters, un dans le vmPFC et un autre plus dorsal, parfois regroupés dans un seul grand cluster, sont régulièrement associés avec la

## Introduction

confiance (Kuchinke et al., 2013; Lebreton et al., 2015; Moritz et al., 2006; Vaccaro and Fleming, 2018).



**Figure 22 : Un gradient dans le cortex préfrontal médian : de l'encodage de la valeur subjective vers l'encodage de la confiance.** La figure de gauche montre d'où ont été extraites les régions d'intérêt utilisées pour l'analyse de droite. La figure de droite montre la différence moyenne à travers les sujets entre la réponse paramétrique à l'évaluation de la confiance (C2) par rapport à la réponse paramétrique à l'évaluation de la désirabilité d'un item (R2) en fonction de la localisation de chacune des régions d'intérêts classées de la plus ventrale à la plus dorsale. Plus la région d'intérêt sélectionnée est dorsale, plus l'encodage de la confiance y est important par rapport à l'encodage de la valeur subjective et réciproquement. L'étude est basée sur 22 participants. Figure issue de (De Martino et al., 2017)

En résumé, s'il paraît relativement démontré que le cortex préfrontal médian corrèle avec l'estimation subjective de la confiance à travers plusieurs types de tâches et de conditions, il reste à mieux comprendre quelle zone spécifique est impliquée dans cet encodage. En effet, anatomiquement, les différentes parties du cortex préfrontal médian qui ont été identifiées font partie de populations neurales ayant des propriétés différentes. On peut identifier au moins trois zones anatomiques qui sont en lien avec les clusters fonctionnels identifiés par les différentes études mentionnées précédemment (voir **Figure 23**). Il reste donc à mieux comprendre s'il existe une seule zone de la confiance à travers toutes les situations ou si cette zone est tâche-dépendante ce qui expliquerait les différences de localisation observées pour le cluster en lien avec la confiance.





**Figure 23 : Comparaison entre clusters anatomiques dans le cortex préfrontal médian et clusters paramétriquement corrélés à la confiance.** Figures du haut : Distinctions anatomiques de différentes parties du cortex préfrontal médian d'après (Córcoles-Parada et al., 2017). Figures du bas : à gauche clusters corrélant à la confiance dans une tâche d'évaluation d'après (De Martino et al., 2017), au milieu clusters corrélant à des évaluations comportementales de la confiance dans une tâche de perception (Gherman and Philiastides, 2018) et à droite figure issue de la même étude montrant le lien entre le vmPFC et une mesure de confiance dérivée d'un signal électrophysiologique plutôt que directement des ratings (Gherman and Philiastides, 2018). Ces différentes figures illustrent à quel point la localisation précise des clusters liés à la confiance varie. Bien qu'elle reste toujours centrée dans le cortex préfrontal médian, d'une étude et d'une méthode d'analyse à l'autre, les clusters se superposent à différentes aires anatomiques. Sens des sigles : mPFC : cortex préfrontal médian, pgACC : cortex antérieur cingulaire pré-génual, vmPFC : cortex préfrontal ventromédian. Les coordonnées sont indiquées dans l'espace MNI.

Les différences spatiales dans la localisation du cluster lié à la confiance pourraient, de plus, être liées à la dynamique des différentes zones du cortex préfrontal médian. Il a récemment été montré que le vmPFC encodait d'abord la valeur subjective puis la confiance en deux temps séparés et non en parallèle (voir **Figure 24** et (Shapiro and Grafton, 2020)). Si l'ensemble du cortex préfrontal médian encode la confiance mais avec des dynamiques temporelles différentes, il est possible que les différences anatomiques liées aux résultats de ces différentes études soient principalement liées à des différences de résolution temporelle entre ces études, mais cela reste à explorer pour l'avenir.



**Figure 24 : Dynamique du vmPFC en lien avec la valeur subjective (SV) et la confiance (confidence) au cours du temps.** Cette figure est issue d'une tâche de prise de décision basée sur la valeur d'après (Shapiro and Grafton, 2020).

Les individus cherchent à maximiser leur confiance dans leurs décisions. Ils adaptent le temps qu'ils mettent à répondre dans une décision donnée en vue de maximiser cette mesure de confiance. Au niveau cérébral, les choses ne sont pas encore tout à fait claires. Le cortex préfrontal médian, en particulier sa partie ventrale que nous avons déjà identifié comme étant membre du réseau des valeurs, est en tout cas assez régulièrement lié à l'estimation de la confiance dans une décision. La même zone responsable de l'estimation de la valeur subjective serait ainsi responsable de l'estimation de la confiance dans cette première estimation.

## **IV.2 Les coûts de la prise de décision**

### **IV.2.a La prise de décision est coûteuse**

**Le système pavlovien et le système des habitudes ne « coûtent rien ».** Ils sont perçus comme opérant de manière automatique et ne demandant, de ce fait, pas d'effort. Ils n'impliquent d'ailleurs pas vraiment de rapport bénéfices/coûts mais plutôt une réponse

## Introduction

automatisée en réponse à un set de stimuli donné. Ceci est d'ailleurs cohérent avec le fait que ces systèmes ne semblent pas s'accompagner de mesure de confiance dans leur fonctionnement. Ceci a mené au développement de plusieurs théories qui parlent de l'utilisation d'une stratégie « sans modèle du monde » (*model-free*) et avec un coût neural très faible qui s'opposerait à une stratégie « basée sur un modèle du monde » (*model-based*) qui implique des computations neurales avec un coût (Daw et al., 2005). La stratégie sans modèle du monde est ainsi partiellement équivalente à l'utilisation du système des habitudes.

A l'inverse, lorsque les individus emploient le système orienté vers un but, ils cherchent à maximiser leur niveau de confiance dans leurs décisions comme nous venons de le voir. Ceci va de pair avec un temps de délibération accru pour que le niveau de confiance soit plus élevé, ainsi qu'avec l'allocation de ressources neurales accrue. Cela implique qu'augmenter le niveau de la confiance dans une décision donnée va aussi de pair avec un coût. **L'emploi du système orienté vers un but s'accompagne ainsi d'un coût subjectif.** Ce coût se reflète d'ailleurs partiellement dans des modèles comme le modèle d'urgence (*urgency gating model*) de Paul Cisek qui partent du principe que le cerveau lancerait une sorte de minuteur une fois qu'il démarre une tâche (ou l'essai d'une tâche en laboratoire). Ce minuteur indiquerait une « urgence à répondre » qui augmenterait linéairement en fonction du temps « dépensé » dans une décision donnée. Indépendamment de la tâche en cours, le cerveau forcerait l'individu à ne pas passer trop de temps sur une tâche particulière en le poussant à répondre même s'il n'a pas confiance dans son choix (Cisek et al., 2009; Ditterich, 2006; Drugowitsch et al., 2012; Thura et al., 2012). D'autres modèles, comme le modèle de *drift-diffusion* (voir *Modèle de diffusion*), ont aussi incorporé ce type de signal d'urgence (Milosavljevic et al., 2010). Ces modèles qui incorporent un signal d'urgence ont été partiellement contestés par Hawkins *et al.*, en particulier sur l'idée que les individus ressentiraient systématiquement ce type de signal dans n'importe quelle tâche. Les auteurs proposent en effet que ce type de signal soit spécifique à des tâches où les individus ont été beaucoup entraînés et que généralement la présence ou non de ce type de signal d'urgence dépende de la tâche en cours (Hawkins et al., 2015). De plus, la plupart de ces études sont basées sur des études basées sur une prise de décision perceptuelle et il est dur, pour l'instant, de savoir dans quelle mesure il existe ou non un signal d'urgence dans le cas de la prise de décision basée sur la valeur lorsqu'on utilise le système orienté vers un but.

### IV.2.b La nature des coûts liés à une décision

Le système orienté vers un but s'adapte continuellement aux contingences changeantes dans le corps (faim/satiété, soif ou non, libido ou non, etc.) et/ou dans l'environnement

## Études expérimentales

(présence plus ou moins élevée de récompenses et/ou de punitions). Il requiert donc au moins une des fonctions exécutives que nous avons déjà évoquées : la **flexibilité**. De plus, il va de pair avec **l'inhibition** et l'emploi de contrôle cognitif afin de résister aux actions automatisées proposées par les deux autres systèmes. Or, l'utilisation des fonctions exécutives, comme nous l'avons déjà vu (voir *L'activité mentale est subjectivement coûteuse*), s'accompagne d'une **sensation subjective d'effort mental**. Comme nous l'avons vu précédemment (voir *De quoi l'aversion pour l'activité mentale nous protège-t-elle ?*), il existe au moins 5 grandes hypothèses sur la nature des coûts liés à la sensation d'effort mental. Si l'on reprend ces points en lien avec la prise de décision plus spécifiquement, la nature des coûts associés à la prise de décision peut s'expliquer selon les théories suivantes :

- Le temps passé à prendre la décision (coût d'opportunité) ;
- La dépense de ressources énergétiques par les aires cérébrales nécessaires à la prise de décision ;
- L'accumulation de déchets dans les aires cérébrales nécessaires à la prise de décision ;
- Le coût de l'information en termes de différence entre l'estimation initiale de la valeur subjective associée à chaque item au début du processus de prise de décision et l'estimation à un instant donné de la prise de décision ;
- Une contrainte qui serait imposée pour favoriser l'emploi des procédures automatisées, comme celles associées au fonctionnement du système des habitudes.

Quoiqu'il en soit, le système orienté vers un but s'accompagne d'un coût en termes de sensation subjective d'effort mental. Le coût de l'utilisation de ce système se superpose aux coûts de l'action en tant que telle. Dans cette thèse, nous allons, entre autres, essayer d'identifier les bases neurales de ce coût lié à la décision elle-même.

## Chapitre 2: Études expérimentales

En introduction, nous avons vu d'un côté qu'il existe un réseau constitué principalement du cortex préfrontal ventromédian, le vmPFC, et du striatum ventral dont l'activité corrèle positivement avec la valeur subjective issue du compromis coûts/bénéfices. Ce réseau est, de ce fait, régulièrement dénommé le réseau des valeurs (voir *L'estimation des valeurs positives : le réseau des valeurs*). D'un autre côté, il existe un réseau principalement constitué du cortex préfrontal dorsomédian, le dmPFC, et de l'insula antérieure dont l'activité augmente avec tout ce qui dévalue la valeur subjective d'une action ou d'un item (voir *Les dévalueurs de la valeur*

## Études expérimentales

*subjective*). Cela a permis de penser que ces deux réseaux pouvaient fonctionner en tant que deux réseaux opposants, l'un étant lié aux affects positifs et l'autre aux affects négatifs (Knutson et al., 2014). D'un autre côté, le réseau constitué du dmPFC et de l'insula antérieure a été impliqué dans l'évaluation de la valeur d'utiliser du contrôle cognitif et dans la régulation de l'activité du cortex préfrontal dorsolatéral qui est supposé être le siège du contrôle cognitif (voir *Bases neurales de l'effort mental*). En d'autres termes, une activité permettant de réguler l'effort alloué au compromis coûts/bénéfices. Il est bien évidemment possible que ce réseau constitué du dmPFC et de l'insula antérieure assure à la fois ces deux rôles ou que des sous-parties différentes du dmPFC soit associée à différentes fonctions. Ici, nous allons chercher à voir dans quelle mesure ces deux réseaux, le réseau des valeurs et le réseau constitué du dmPFC et de l'insula antérieure, sont anti-corrélés et représentent tous deux la valeur subjective dans des sens opposés ou si, au contraire, ils reflètent des concepts cognitifs différents qui, dans certaines conditions, peuvent de fait sembler anti-corrélés, mais dans d'autres situations pourraient même être corrélés.

Dans notre première étude, nous avons voulu voir, à l'aide de l'imagerie par résonance magnétique fonctionnelle (IRMf) chez des participants sains, s'il était possible de dissocier trois concepts fondamentaux du compromis coûts/bénéfices : la valeur subjective issue du calcul bénéfices/coûts, la confiance dans ce calcul et l'effort alloué à ce calcul. En vue d'identifier des corrélats robustes de ces trois concepts, nous avons utilisé une approche peu classique dans le domaine qui consistait à voir dans quelle mesure les bases cérébrales de ces concepts cognitifs étaient robustes au travers de différentes tâches. Nous avons ainsi utilisé une tâche d'évaluation, une tâche de choix entre deux options appartenant à la même dimension, des récompenses ou des efforts et une tâche de choix nécessitant d'intégrer à la fois les bénéfices associés à une récompense à obtenir et les coûts associés à l'action nécessaire pour les obtenir. En vue de dissocier les variables décisionnelles, c'est-à-dire le calcul des coûts et des bénéfices, des variables méta-décisionnelles, c'est-à-dire les variables impliquées dans l'allocation de l'effort lié à la prise de décision en elle-même, nous avons utilisé des items purement fictifs. Ce procédé nous a en effet permis d'être sûrs que les corrélats cérébraux associés à l'estimation des récompenses et des coûts étaient purement liés à l'évaluation et pas à l'anticipation d'un effort à fournir pour obtenir la récompense. Enfin, nous avons aussi inclus des items d'effort à évaluer dans cette étude en vue de voir si les aires appartenant au réseau des valeurs étaient capables d'en estimer les coûts ou si d'autres aires, comme le dmPFC, étaient en charge du calcul de ces coûts liés à l'action.

## Études expérimentales

Ensuite, nous avons utilisé des tâches impliquant à la fois la perspective de gains et de pertes. L'aversion à la perte par rapport à un point de référence est censée être un phénomène général dans le règne animal (voir *L'économie comportementale*) mais elle a principalement été étudiée chez l'homme dans des tâches impliquant des choix monétaires. Dans cette étude, nous avons voulu voir si ce phénomène était bien observable chez des participants sains dans le cadre d'une tâche impliquant une performance d'ordre physique ou mental. Les participants de cette étude ont effectué cette tâche en IRMf pour pouvoir étudier les bases cérébrales de cette aversion à la perte, si elle avait lieu. Nous nous attendions à ce que les structures du réseau des valeurs s'activent de manière linéaire avec les enjeux monétaires signés dans nos tâches. Nous voulions aussi voir si le réseau constitué du dmPFC et de l'insula antérieure s'activerait de manière anti-corrélée au réseau des valeurs, c'est-à-dire négativement avec les enjeux monétaires signés, ou s'il s'activerait proportionnellement aux enjeux monétaires non-signés, c'est-à-dire à la quantité de ressources à investir à chaque essai.

Enfin, nous avons fait passer une tâche d'apprentissage par renforcement à des participants sains en IRMf. Ce type d'apprentissage, qu'on retrouve chez la plupart des espèces animales, pourrait se faire de manière complètement automatique sans nécessiter d'allouer de l'effort mental puisqu'il paraît observable même chez des espèces dépourvues de la capacité de contrôle cognitif. Cette étude visait, entre autres, à vérifier cette hypothèse en regardant s'il était possible d'observer une activité du réseau constitué du dmPFC et de l'insula antérieure en lien avec l'effort alloué au moment de la délibération.

## **I La valeur, la confiance et le temps de délibération au moment du compromis coûts/bénéfices ont des supports neuraux distincts dans le cortex préfrontal médian**

### **I.1 Introduction**

Toute décision implique un rapport entre les bénéfices espérés et les coûts associés à ces bénéfices. En théorie de la décision, la sélection d'une série d'actions est vue comme la recherche de maximisation d'une fonction d'utilité qui se traduit par une valeur nette où les bénéfices espérés sont décomptés par les coûts attendus. En neurosciences de la décision, deux théories s'opposent à ce sujet. L'une de ces théories propose qu'il existe deux systèmes opposants, séparés anatomiquement dans le cerveau (Knutson et al., 2007; Pessiglione and Delgado, 2015). L'un de ces deux systèmes s'activerait ainsi proportionnellement aux gains

## Études expérimentales

potentiels et motiverait l'individu à agir en vue de ces gains. L'autre système s'activerait proportionnellement à tout ce qui dévalue la valeur des gains, qu'il s'agisse de la pénibilité des actions requises à l'obtention des gains, l'incertitude d'obtention des gains ou le délai à attendre en vue des gains. La seconde théorie propose qu'il existe un système qui intègre aussi bien les bénéfices que les coûts (Kable and Glimcher, 2007; Seaman et al., 2018). Cette intégration pourrait cependant se faire sur la base d'informations issues d'aires cérébrales distinctes chargée respectivement de calculer les bénéfices et les coûts associés à une action de manière indépendante (Pessiglione et al., 2018).

Une question essentielle en neurosciences de la décision est de savoir comment le cerveau effectue un compromis entre les ressources à investir dans l'action, c'est-à-dire l'effort, et la valeur espérée des conséquences de cette action, c'est-à-dire la récompense. Le réseau des valeurs, dont le cortex préfrontal ventromédian (vmPFC) est l'une des composantes principales, a été identifié comme un réseau dont l'activité en IRMf variait linéairement avec la valeur subjective de récompenses, indépendamment de la nature de ces récompenses (Bartra et al., 2013; Levy and Glimcher, 2012). Le fait que le vmPFC encode la valeur subjective associée à des récompenses fait consensus dans le domaine. Cependant, le fait que la valeur subjective, qui est encodée dans le vmPFC prenne ou pas en compte le coût des efforts nécessaires à l'obtention d'une récompense est sujet à controverse (Schneider and Koenigs, 2017). En particulier, l'idée même que le vmPFC puisse encoder la valeur associée à une action plutôt qu'à un item fait débat (Gläscher et al., 2009; Padoa-Schioppa, 2011; Rangel and Hare, 2010). Certaines études ont pu montrer que l'activité du vmPFC baissait avec le coût envisagé des efforts nécessaires à l'obtention d'une récompense (Arulpragasam et al., 2018; Seaman et al., 2018). D'autres ont identifié un réseau opposé qui signalerait le coût associé à un effort et dont le siège serait dans le cortex préfrontal dorsomédian (dmPFC), aussi régulièrement appelé le cortex cingulaire antérieure dorsal (dACC), ainsi que dans l'insula antérieure (Kurniawan et al., 2013; Skvortsova et al., 2014).

Une des explications possibles pour ces divergences est la diversité des tâches comportementales employées. Dans cette étude, nous avons cherché non pas à isoler les bases cérébrales liées au compromis récompense/effort au sein d'une seule tâche, mais au contraire à voir dans quelle mesure nous arrivions à identifier des corrélats cérébraux stables à travers différentes tâches et différentes conditions dans le but d'atteindre des conclusions plus robustes et généralisables au travers de multiples paradigmes. Nous avons ainsi employé différentes tâches basées sur les préférences, autrement dit les valeurs subjectives, des individus. Ces tâches

## Études expérimentales

nous ont permis d'étudier : 1) la valeur subjective d'une option représentant soit une récompense soit un effort à l'aide d'une tâche d'évaluation, 2) la comparaison entre deux options (2O) issues d'une même dimension (1D), récompense ou effort, à l'aide d'une tâche de choix 1D-2O et 3) l'intégration de deux dimensions (2D) impliquant de pondérer une récompense par rapport au coût d'un effort, au sein d'une option dans une tâche de choix 1O-2D.

Une autre possibilité dans les différences entre études est la confusion possible avec des variables méta-décisionnelles. En effet, effectuer un calcul coûts/bénéfices est un processus qui pourrait, en soi, impliquer un coût en termes d'effort mental. Chacun sait que certaines décisions sont prises rapidement, alors que d'autres prennent beaucoup plus de temps. Ces décisions plus lentes sont aussi souvent perçues comme plus pénibles. Une explication de ce coût subjectif est basée sur l'idée que des décisions plus difficiles impliquent un effort mental sous la forme de contrôle cognitif. Cette idée a été formalisée par l'idée que l'investissement de ressources neurales, sous forme de contrôle cognitif, faisait en soi l'objet d'un compromis coûts/bénéfices (Shenhav et al., 2013). Si on applique cette idée à la prise de décision, le coût s'exprime au travers du temps et des ressources neurales allouées à la décision. La récompense, elle, s'exprime en termes de confiance obtenue à l'issue de ce calcul (Lee and Daunizeau, 2019). Ainsi, dans cette étude, nous n'avons pas simplement étudié les variables directement associées au compromis coûts/bénéfices lié à une prise de décision, mais aussi les variables « méta-décisionnelles » qui régulaient le compromis coûts/bénéfices : la confiance subjective dans la réponse et le temps de délibération. Dans notre étude, toutes les récompenses et les efforts proposés étaient purement fictifs. Nous pouvions ainsi plus facilement différencier l'estimation du coût associé à l'option en cours d'évaluation du coût dépensé au cours de la prise de décision en elle-même.

Nos résultats suggèrent une triple dissociation au sein du cortex préfrontal médian, qui est stable au travers des tâches de préférences et des conditions. La valeur subjective est ainsi représentée dans le vmPFC indépendamment de la dimension des options considérées (récompenses ou efforts), alors que les variables méta-décisionnelles, la confiance et le temps de délibération, sont représentées dans le mPFC et dans le dmPFC, respectivement.



## I.2 Méthodes

### I.2.a Participants

Nous avons mené une première expérience en IRMf avec une cohorte de 15 participants. Cette étude a par la suite été complétée avec une deuxième cohorte de 25 participants. L'ensemble de ces 40 participants a effectué exactement les mêmes tâches. Les participants ont été recrutés en ligne via la plateforme du RISC (Relais d'Information en Sciences de la Cognition, <https://www.risc.cnrs.fr/>). Les critères d'inclusion étaient d'être droitier, de parler couramment le français, d'avoir entre 20 et 39 ans, de ne pas porter de tatouage, de piercing ou tout objet métallique, de ne pas être enceinte (vérifié avec un test de grossesse effectué avant l'expérience) et de ne pas avoir de problèmes de vue. Nous avons aussi évalué les antécédents médicaux, en particulier neurologiques ou psychiatriques, et la consommation éventuelle de substances psychotropes des participants afin de les exclure le cas échéant. Enfin, les participants devaient aussi indiquer leurs traitements en cours afin de les exclure au cas où leurs traitements pourraient avoir un impact sur leur comportement et/ou leur activité neurale. Un participant a dû être exclu car il n'avait clairement pas compris les consignes. Au total, les données comportementales présentées ici sont donc basées sur 39 participants sains (22 femmes / 17 hommes) ayant  $25.4 \pm 4.1$  ans (moyenne  $\pm$  écart-type). Pour les données d'IRMf, nous avons exclu un participant en plus car il bougeait trop pendant la tâche (mouvement supérieur à 3 millimètres). Il a cependant été inclus dans les études de pupille et de comportement. Les données d'IRMf sont donc basées sur 38 participants au total. Pour l'analyse du diamètre pupillaire, nous avons aussi dû exclure 7 à 9 participants, selon la tâche étudiée, pour qui les données enregistrées étaient inutilisables (voir *Analyse du diamètre pupillaire*).

Tous les participants ont signé une fiche de consentement. Les participants de la première cohorte ont été payés un montant fixe de 60€ et ceux de la deuxième cohorte 75€. Cette différence de montant n'avait pas de motivation en terme expérimental, mais était simplement liée au fait qu'il s'agissait de deux protocoles expérimentaux distincts associés à des budgets différents. Le premier protocole ( $n = 15$ ) avait pour but de comparer différentes séquences d'IRMf, une séquence classique EPI (*Echo Planar Imaging*), une séquence EPI avec un facteur d'accélération multi-bandes et une séquence EPI avec un facteur d'accélération multi-bandes et une acquisition avec deux temps d'échos distincts. L'analyse principale de cette étude n'inclut qu'un tiers des données acquises chez ces participants avec la séquence classique EPI. Les participants du deuxième protocole ( $n = 25$ ) ont tous été scannés avec la séquence classique EPI

### I.2.b Tâches comportementales

Toutes les tâches ont été programmées dans Matlab 2012 (The MathWorks, Inc., USA) via la toolbox Psychtoolbox (<http://psychtoolbox.org/>). Les participants avaient à leur disposition un boîtier de réponse avec 4 boutons (FORP 932, Current Designs Inc, Philadelphia, USA) posé sur leur torse, sous leur main droite, pour donner leurs réponses. Les stimuli que nous avons défini pour l'ensemble de nos tâches étaient constitués de 144 items de récompenses et de 72 items d'efforts. Ces 144 items de récompense étaient divisés en 72 récompenses principalement constituées d'objets du quotidien et en 72 récompenses de nature alimentaire. La moitié des items de récompenses de chaque catégorie était présentée uniquement sous la forme de texte (items  $R_t$ ) et l'autre moitié était présentée sous la forme d'un texte accompagné d'une image (items  $R_{ti}$ ). Les images ont été ajustées de sorte à avoir à peu près les mêmes dimensions à l'écran, mais la luminance d'une image à l'autre pouvait varier. Les items d'efforts étaient divisés en 36 efforts de nature physique et 36 efforts de nature mentale. Tous les efforts étaient uniquement présentés sous la forme de texte à l'écran (items  $E_t$ ). Pour prendre en compte les variations de luminance de tous nos stimuli dans nos analyses, nous avons calculé la luminance de chaque stimulus à l'aide de la formule standard  $0.299 \cdot rouge + 0.587 \cdot vert + 0.114 \cdot bleu$  (<http://www.w3.org/TR/AERT#color-contrast>) et de la commande `Screen('GetImage')` de Psychtoolbox qui nous a permis d'estimer la composition en rouge, vert et bleu de chaque stimulus. Avant de démarrer chaque tâche pour la première fois, les participants effectuaient quelques essais, non inclus dans l'analyse des données, afin de se familiariser avec la tâche, le type d'items qu'ils auraient à évaluer et l'utilisation du boîtier de réponse dans la tâche. Les participants étaient informés dès le début de l'expérience que leurs décisions étaient purement virtuelles et que ni les récompenses ne seraient obtenues, ni les efforts ne seraient à effectuer à la suite de l'expérience.

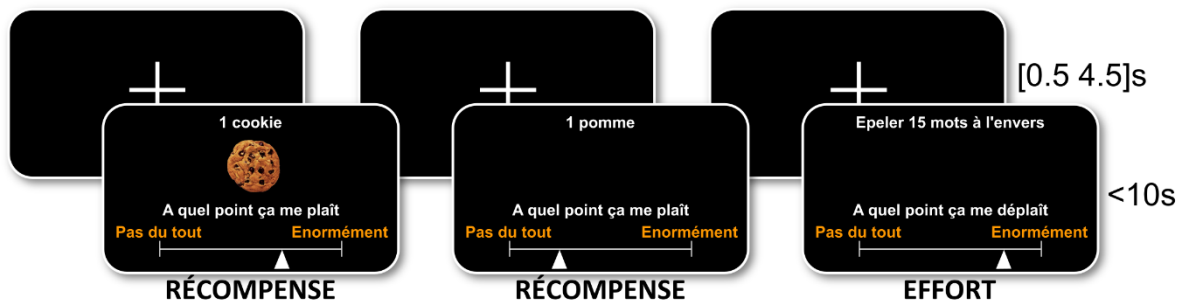
#### Tâche d'évaluation

Tous les participants ont commencé par une tâche d'évaluation. Dans cette tâche, ils devaient évaluer tous les stimuli qui leur étaient présentés (voir **Figure 25**). Chacune des trois sessions en IRMf de cette tâche était divisée en trois blocs. Un bloc était constitué de 24 items  $R_{ti}$ , de 24 items  $R_t$  ou de 24 items  $E_t$ . L'ordre des trois blocs était contrebalancé à travers les participants au sein de chaque étude (première et deuxième cohorte IRMf) mais, pour un participant donné, le même ordre était maintenu pour les trois sessions. Les items à évaluer au sein d'un bloc étaient présentés un par un au centre de l'écran. Les participants devaient donner un jugement de valeur pour chacun des items à l'aide d'un curseur à déplacer sur une échelle.

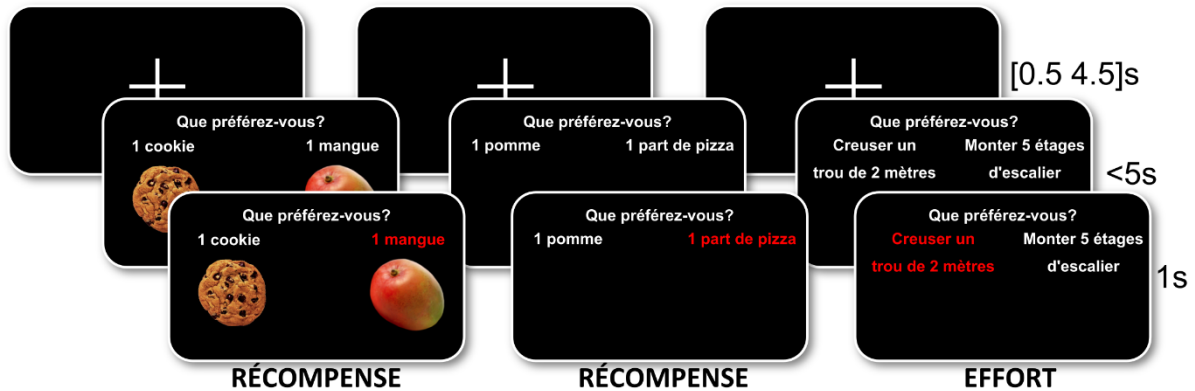
## Études expérimentales

Le curseur apparaissait aléatoirement entre 25 et 75% de l'échelle au début de chaque essai en vue de limiter tout effet d'ancrage sur la valeur subjective exprimée par les participants (Kahneman, 2012; Tversky and Kahneman, 1974). Les participants pouvaient déplacer le curseur à l'aide des deux premiers boutons du boîtier de réponse : le premier correspondait à déplacer le curseur vers la gauche et le second à déplacer le curseur vers la droite. Un troisième bouton permettait aux participants de valider la position du curseur et ainsi de mettre fin à l'essai en cours. L'échelle était divisée en 100 points de sorte à ce que les évaluations des participants pour chaque stimulus étaient traduites en un score compris entre 1 et 100. Cependant, la graduation de l'échelle n'était pas indiquée sur l'écran donnant l'impression que l'échelle était continue. Les extrêmes de l'échelle étaient accompagnés du label « Pas du tout » à gauche et « Énormément » à droite. Pour les récompenses, les participants devaient dire à quel point cela leur aurait plu d'obtenir l'item proposé à l'écran. Une évaluation à l'extrême gauche de l'échelle revenait à dire que l'item ne les intéressait pas, alors qu'une évaluation à l'extrême droite de l'échelle correspondait à un item très appétitif. Pour les efforts, les participants devaient indiquer à quel point cela leur déplairait d'avoir à effectuer l'effort proposé à l'écran. Une évaluation à l'extrême gauche de l'échelle signifiait ainsi que l'effort ne paraissait pas du tout pénible, alors qu'une évaluation à l'extrême droite de l'échelle impliquait qu'effectuer l'effort proposé à l'écran paraissait extrêmement déplaisant. Dans le cas des items  $R_{ii}$ , nous demandions aux participants d'évaluer l'item représenté par l'image plutôt que la récompense indiquée par le texte en général. Afin de motiver les participants à ne pas passer trop de temps sur chaque essai, un message était affiché à l'écran au bout de 10 secondes. Ce message affichait les mots « Trop lent ! » à l'écran. Ensuite les participants étaient tout de même invités à répondre à l'essai en cours pour pouvoir avoir une évaluation pour chaque item, mais ces essais étaient exclus des analyses.

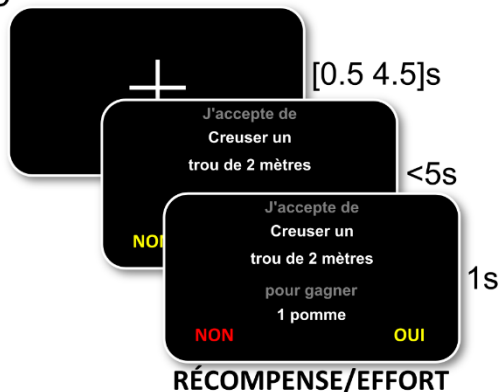
A) Tâche d'évaluation



B) Tâche de choix 1D-2O



C) Tâche de choix 1O-2D



**Figure 25 : Tâches expérimentales.** Des exemples d'essais sont illustrés pour chaque tâche et chaque condition. Chaque tâche est illustrée à l'aide d'un ensemble de captures d'écran montrant le déroulé d'un essai du haut vers le bas. La durée de chaque période est indiquée à droite des images. La durée de la croix de fixation était aléatoire à chaque essai dans chaque tâche et comprise entre 0.5 et 4.5 secondes. La durée de la période de réponse dépendait du temps de réponse des participants à chaque essai. Nous imposons une limite de 10s pour la tâche d'évaluation et de 5s pour les tâches de choix après quoi, si les participants n'avaient toujours pas répondu, le message « Trop lent ! » était affiché à l'écran afin de pousser les participants à répondre plus rapidement, puis l'essai redémarrait.

**A) Tâche d'évaluation.** Dans cette tâche, les participants devaient, à l'aide d'un curseur, indiquer à quel point cela leur plairait d'obtenir la récompense affichée à l'écran pour les blocs présentant des récompenses sous forme de texte seul (milieu) ou avec un texte et une image (gauche) ou à quel point cela leur déplairait de devoir exécuter l'effort affichée à l'écran pour les blocs présentant des efforts sous forme de texte seul (droite). Pour ce faire, les participants pouvaient déplacer le curseur qui apparaissait à l'écran aléatoirement entre 25 et 75% de l'échelle vers la gauche ou vers la droite à l'aide de deux boutons dédiés dans le boîtier de

## Études expérimentales

réponse. Lorsqu'ils étaient satisfaits de la position du curseur, ils devaient alors valider cette position à l'aide d'un troisième bouton.

**B] Tâche de choix 1D-2O.** Dans cette tâche, les participants devaient à chaque fois choisir entre 2 options (2O) issues de la même dimension (1D) laquelle ils préféreraient obtenir dans le cas hypothétique où on leur proposerait les 2 options présentées. Les choix se faisaient soit entre deux récompenses, soit entre deux efforts parmi celles et ceux qu'ils avaient évalué lors de la première tâche. Les items étaient là aussi regroupés en 3 blocs selon la dimension (récompense/effort) et selon le mode de présentation (récompenses accompagnées ou non d'une image représentative). Les participants sélectionnaient l'une des deux options à l'aide de deux boutons correspondants à la localisation spatiale (gauche/droite) de chaque option. L'option choisie s'affichait alors en rouge puis l'essai suivant démarrait.

**C] Tâche de choix 1O-2D.** Dans cette tâche, les participants devaient à chaque fois indiquer s'ils seraient prêts ou non à exécuter l'effort affiché à l'écran pour obtenir la récompense affichée à l'écran. Cette tâche impliquait ainsi deux dimensions (2D), la récompense et l'effort, à intégrer au sein d'une seule option (1O) à accepter ou à rejeter. Les participants sélectionnaient s'ils seraient prêts à accepter (« OUI ») ou s'ils rejetteraient (« NON ») la combinaison de récompense et d'effort proposé à l'écran à l'aide de deux boutons correspondants à la localisation spatiale (gauche/droite) des options « OUI » et « NON ». La localisation des options « OUI » et « NON » alternait à chaque essai. L'option choisie s'affichait alors en rouge puis l'essai suivant démarrait.

- **Tâche de choix 1D-2O**

Après les trois sessions de la tâche d'évaluation, les participants enchaînaient systématiquement avec trois sessions d'une tâche de choix binaires leur demandant, à chaque essai, d'exprimer leur préférence au sein d'une dimension (1D) (récompense ou effort) entre 2 options (2O) présentées à gauche et à droite de l'écran. Les options présentées avaient toutes évaluées au cours de la tâche d'évaluation. Comme pour la tâche d'évaluation, la tâche était divisée en trois blocs (voir **Figure 25**) correspondant à 24 choix entre des items  $R_{ti}$ , 24 choix entre des items  $R_t$  ou 24 choix entre des items  $E_t$ . L'ordre des trois blocs pour un sujet donné était le même que pour la tâche d'évaluation et il était ainsi, de même, contrebalancé à travers les participants. Pour un choix donné, les options affichées étaient toujours issues de la même dimension (effort vs effort ; récompense vs récompense), du même mode de présentation ( $R_{ti}$  vs  $R_{ti}$  ;  $R_t$  vs  $R_t$  ;  $E_t$  vs  $E_t$ ) et de la même catégorie (physique vs physique ; mental vs mental ; alimentaire vs alimentaire ; objet vs objet). Chaque item évalué lors de la première tâche était présenté deux fois au cours de cette tâche en fonction de deux types d'appariement, effectués indépendamment pour chaque catégorie de chaque bloc. L'un des deux appariements impliquait que l'évaluation moyenne des deux options était constante, mais que la distance des évaluations entre les deux options variait, de sorte à ce qu'il y ait à la fois des choix très faciles où la distance des évaluations entre les deux options du choix était maximale et d'autres plus difficiles où les

évaluations étaient proches. Le deuxième appariement impliquait de regrouper les items en deux groupes selon qu'ils étaient associés à une évaluation supérieure ou inférieure à la médiane. Ensuite, la distance entre les évaluations des deux options variait au sein de chaque groupe tout en maintenant la moyenne des évaluations des deux options constante au sein de chaque groupe. Les essais étaient ensuite brassés aléatoirement, au sein de chaque bloc. Dans le cas des récompenses, les participants devaient choisir la récompense qu'ils préféreraient obtenir. Dans le cas des efforts, ils devaient choisir l'effort qu'ils préféreraient devoir effectuer. Pour répondre, ils devaient utiliser le boîtier de réponse. Les deux boutons correspondant à déplacer le curseur vers la gauche ou vers la droite dans la tâche d'évaluation, servaient à sélectionner l'option de gauche ou l'option de droite. Une fois que le choix était effectué, l'option choisie apparaissait en rouge, ce qui permettait aux participants de vérifier que leur réponse avait bien été enregistrée correctement (voir **Figure 25**). Afin de motiver les participants à ne pas passer trop de temps sur chaque essai, un message était affiché à l'écran au bout de 5 secondes. Ce message affichait les mots « Trop lent ! » à l'écran puis le même essai reprenait. Ces essais étaient exclus des analyses.

- **Tâche de choix 1O-2D**

Après avoir effectué les trois sessions de la tâche d'évaluation et les trois sessions de la tâche de choix binaires, les participants devaient effectuer trois sessions d'une tâche leur demandant d'effectuer des compromis entre une récompense et un effort au cours de trois sessions, constituée chacune de 48 essais. Dans cette tâche, les participants devaient décider, à chaque essai, d'accepter ou de refuser une option (1O) qui leur était proposée. Cette option était constituée de deux dimensions (2D) à intégrer par le participant pour décider si, oui ou non, il serait prêt à exécuter l'item d'effort  $E_t$  proposé pour obtenir l'item de récompense  $R_t$  associé (voir **Figure 25**). En d'autres termes, les récompenses présentées avec du texte et une image n'étaient plus utilisées pour cette tâche. Chaque item  $E_t$  et  $R_t$  était présenté deux fois dans cette tâche selon deux modes d'appariement qui brassaient les items indépendamment de leur catégorie d'appartenance (alimentaire, objet, physique, mental). Le premier mode d'appariement regroupait les récompenses les plus appétitives avec les efforts les plus pénibles ce qui permettait de varier les options présentées tout en contrôlant la difficulté du choix, et le deuxième mode d'appariement associait les récompenses les plus agréables avec les efforts les moins pénibles en vue de varier la difficulté des choix. Les options définies ont là aussi été réparties aléatoirement au travers des essais. Les participants répondaient à l'aide des boutons du boîtier de réponse correspondant aux positions gauche/droite de l'écran, en acceptant l'offre

## Études expérimentales

proposée en cliquant du côté où se trouvait l'option « OUI » ou en la refusant en cliquant du côté où se trouvait l'option « NON ». D'un essai à l'autre, la position des options « OUI » et « NON » alternait. Une fois que le participant avait appuyé sur l'un des deux boutons réponses (gauche/droite), l'option choisie (OUI/NON) apparaissait en rouge pour confirmer le choix effectué au participant puis l'essai se terminait (voir **Figure 25**). Afin de motiver les participants à ne pas passer trop de temps sur chaque essai, un message était affiché à l'écran au bout de 5 secondes. Ce message affichait les mots « Trop lent ! » à l'écran puis le même essai reprenait. Ces essais étaient exclus des analyses.

### I.2.c Analyse comportementale

Toutes les analyses ont été effectuées sous Matlab 2017a (The MathWorks, Inc., USA). Les résultats comportementaux sont issus de l'ensemble des participants ayant effectué correctement la tâche, c'est-à-dire 39 participants. A l'issue de la tâche d'évaluation, chaque item était associé à une note comprise entre 1 et 100. Pour l'ensemble des analyses (choix et temps de réponse), nous avons zscoré, par bloc, ces notes à travers les essais et les sessions afin de les normaliser à travers les individus. Le but de cette normalisation était, entre autres, de pallier au fait que chaque individu aurait pu utiliser l'échelle de manière légèrement différente bien que la représentation neurale et l'effet comportemental sous-jacent soit identique.

- **Analyse des choix dans la tâche de choix 1D-2O**

Nous avons d'abord regroupé l'ensemble des essais d'un même bloc à travers les 3 sessions. Les choix effectués dans la tâche 1D-2O ont ensuite été modélisés séparément pour chaque bloc avec une régression logistique de la forme :

$$p(\textit{gauche}) = \frac{1}{1 + e^{\beta_0 + \beta_1 \cdot (E_{\textit{gauche}} - E_{\textit{droit}})}}$$

(Équation 1)

Où  $p(\textit{gauche})$  est la probabilité de choisir l'option située à gauche de l'écran à chaque essai,  $\beta_0$  et  $\beta_1$  sont des constantes individuelles estimées par le modèle,  $E_{\textit{gauche}}$  et  $E_{\textit{droit}}$  correspondent à l'évaluation donnée par le participant lors de la tâche d'évaluation pour l'option affichée à gauche et à droite, respectivement.

Ensuite, nous avons moyenné les betas issus du modèle pour chaque bloc à travers les participants et nous avons testé leur significativité à l'aide d'un t.test en les comparant à zéro.

- **Analyse des choix dans la tâche de choix 2D-1O**

Nous avons d'abord regroupé l'ensemble des essais des 3 sessions où la tâche a été effectuée. Les choix effectués dans la tâche de choix 1O-2D ont ensuite été modélisés avec une régression logistique de la forme :

$$p(\text{accepter}) = \frac{1}{1 + e^{\beta_0 + \beta_R \cdot E_R + \beta_E \cdot E_E}}$$

(Équation 2)

Où  $p(\text{accepter})$  correspond à la probabilité d'accepter l'offre présentée à l'écran à chaque essai,  $\beta_0$ ,  $\beta_R$ ,  $\beta_E$  sont des constantes estimées par le modèle,  $E_R$  et  $E_E$  correspondent aux évaluations pour, respectivement, les récompenses et les efforts.

Ensuite, nous avons moyenné les betas issus du modèle pour chaque bloc à travers les participants et nous avons testé leur significativité à l'aide d'un t.test en les comparant à zéro.

- **Analyse des temps de délibération**

Le temps de délibération (DT) dans chacune de nos tâches a été défini comme le temps entre l'apparition du stimulus à l'écran et le premier appui d'un bouton pour répondre à la tâche. Les variations du DT ont d'abord été modélisées avec un modèle linéaire en regroupant toutes les tâches et toutes les sessions ensemble en vue d'identifier des régresseurs de nuisance qui expliqueraient le DT à inclure dans nos analyses. Nous avons modélisé le DT avec une constante par bloc pour chaque tâche, puis nous avons inclus des régresseurs pouvant expliquer le DT comme le temps de délibération de l'essai précédent, le numéro d'essai dans chaque bloc, le nombre de mots de chaque stimulus, la luminance, la durée de présentation de la croix de fixation précédant le début de chaque essai et enfin aussi nos facteurs d'intérêts, la valeur du stimulus et la confiance séparément pour chaque tâche (voir **Tableau 1**). Ayant identifié que le nombre de mots ( $\beta = 0.066 \pm 0.004$ ,  $p = 710-20$ ), la luminance ( $\beta = 0.022 \pm 0.006$ ,  $p = 610-4$ ) et la durée de présentation de la croix de fixation précédant chaque essai ( $\beta = 0.054 \pm 0.011$ ,  $p = 810-6$ ), mais pas la durée du temps de délibération de l'essai précédent ( $\beta = -0.004 \pm 0.007$ ,  $p = 0.618$ ), ni le numéro d'essai ( $\beta = -0.001 \pm 0.001$ ,  $p = 0.125$ ), expliquaient le temps de délibération, nous avons décidé d'inclure le nombre de mots, la luminance et la durée de présentation de la croix de fixation dans les analyses ultérieures comme des régresseurs de non-intérêt.

Ensuite, nous avons modélisé le DT avec un modèle linéaire séparément pour chaque bloc de chaque tâche. Nous avons d'abord regroupé l'ensemble des essais d'un même bloc à



## Études expérimentales

travers les 3 sessions de chaque tâche. Pour cette analyse, nous avons fait en sorte que l'évaluation, autant dans les blocs de récompense que dans les blocs d'effort, soit orientée dans le même sens avec la valeur la plus basse correspondant à la récompense la moins désirable ou l'effort le plus aversif, et la valeur la plus élevée correspondant à la récompense la plus désirable ou à l'effort le moins aversif. Dans chaque bloc de chaque tâche, nous avons modélisé le temps de délibération comme suit :

$$DT = \beta_{s1} + \beta_{s2} + \beta_{s3} + \beta_{Lum} \cdot Lum + \beta_{croix} \cdot Croix + \beta_{mots} \cdot N_{mots} + \beta_{Val} \cdot Val + \beta_{Conf} \cdot Conf$$

Ce modèle incluait une constante par session ( $\beta_{s1}$ ,  $\beta_{s2}$ ,  $\beta_{s3}$ ), une constante  $\beta_{Lum}$  exprimant la sensibilité à la luminance ( $Lum$ ) des stimuli (voir *Pré-traitement de la pupille* pour la manière de définir la luminance), une constante  $\beta_{croix}$  exprimant la sensibilité à la durée de présentation de la croix de fixation précédant le début de chaque essai ( $Croix$ ), une constante  $\beta_{mots}$  permettant de prendre en compte le temps de lecture lié au nombre de mots ( $N_{mots}$ ) de chaque stimulus, une constante  $\beta_{Val}$  permettant de prendre en compte un effet de la valeur subjective  $Val$  sur le temps de délibération et enfin une constante  $\beta_{Conf}$  afin de vérifier le résultat généralement admis que le temps de délibération était d'autant plus rapide que la confiance subjective était élevée. Pour chaque tâche, le régresseur correspondant à  $Val$  et à  $Conf$  est détaillé dans le tableau ci-dessous (voir **Tableau 1**). Dans toutes les tâches,  $Val$  représentait la valeur des stimuli à l'écran.  $Conf$  nous servait de proxy pour la confiance des participants dans leur décision, c'est-à-dire la probabilité subjective d'avoir choisi la meilleure option. Dans les deux tâches de choix, la probabilité était issue du modèle de régression logistique expliquant les choix.

	<b>Tâche d'évaluation</b>	<b>Tâche de choix 1D-2O</b>	<b>Tâche de choix 1O-2D</b>
<b>Val</b>	$E$	$E_{gauche} + E_{droit}$	$\beta_0 + \beta_R * E_R + \beta_E * E_E$
<b>Conf</b>	$(E \text{ centré})^2$	$(p(\text{gauche}) \text{ centré})^2$	$(p(\text{accepter}) \text{ centré})^2$

**Tableau 1 : Détails des variables Val et Conf par tâche.** Dans la tâche d'évaluation,  $Val$  équivalait à l'évaluation ( $E$ ) fournie par les participants à chaque essai zscorée à travers les essais et les sessions du bloc modélisé,  $Conf$  correspondait à cette même variable mise au carré. Dans la tâche de choix 1D-2O,  $Val$  correspondait à la somme des évaluations des deux options présentées et  $Conf$  correspondait à la probabilité de choisir l'option de gauche mise au carré après l'avoir centrée en lui soustrayant la moyenne, à travers les essais, de la probabilité d'avoir choisi l'option de gauche. Dans la tâche de choix 1O-2D,  $Val$  correspondait à la valeur nette déduite du modèle permettant de modéliser les choix dans cette tâche (voir *Analyse des choix dans la tâche de choix 2D-1O*),  $Conf$  correspondait à la probabilité de choisir d'accepter l'option proposée mise au carré après l'avoir centrée en lui soustrayant la moyenne, à travers les essais, de la probabilité d'avoir choisi d'accepter l'offre proposée. Dans chacune des trois tâches, les évaluations  $E$  utilisées et liées à chaque item étaient zscorées au travers de tous les items par bloc. Pour les blocs impliquant des efforts, l'évaluation était dans le même sens que l'échelle c'est-à-dire qu'une valeur plus élevée représentait un effort plus pénible.

## Études expérimentales

Ensuite, nous avons moyenné les betas issus du modèle linéaire pour chaque bloc à travers les participants et nous avons testé leur significativité à l'aide d'un t.test en les comparant à zéro.

### I.2.d IRM

- **Acquisition des données d'IRM anatomique et fonctionnelle**

Les données d'IRM ont été acquises au Centre de Neuro-Imagerie de Recherche (CENIR, Institut du Cerveau, ICM, Paris) avec un scanner 3 Teslas de modèle Siemens Magnetom Prisma et avec une antenne de tête Siemens à 64 canaux.

Les images d'IRM anatomique T1 ont été obtenues avant de démarrer l'expérience. Les paramètres d'acquisition de ces images anatomiques sont détaillés dans le tableau ci-dessous (voir **Tableau 2**). Ces images anatomiques ont ensuite été co-enregistrées avec la moyenne des images EPI, segmentées, normalisées dans l'espace MNI et moyennées à travers les participants afin de servir à l'identification des clusters fonctionnels des cartes d'activation de groupe.

Temps de répétition (TR) (secondes)	Temps d'écho (millisecondes)	Angle de bascule (°C)	Nombre de coupes	Épaisseur de coupes (millimètres)	Taille des voxels (mm <sup>3</sup> )
2.3	4.18	9	176	1	1

**Tableau 2 : Paramètres d'acquisition pour l'IRM anatomique.**

Les images d'IRM fonctionnelle T2 ont été obtenues via 9 sessions. Pour la première cohorte (n = 15), les participants ont effectué 3 types de séquences d'IRMf. Les paramètres utilisés pour chacune de ces séquences sont détaillés ci-dessous (voir **Tableau 3**). Chaque tâche étant effectuée en 3 sessions, nous avons employé une séquence différente pour chaque session pour cette cohorte. Cette étude visait à comparer les différentes séquences entre elles. Plus spécifiquement, nous cherchions à voir s'il était possible d'améliorer la sensibilité à l'activité du vmPFC en lien avec la valeur subjective en utilisant différentes séquences. On pourra trouver le résumé de ces résultats en annexe (voir Appendice : *Enregistrement de l'activité du vmPFC à l'aide de différentes séquences d'IRMf*). Les résultats montrés ici prennent uniquement en compte les données de cette cohorte ayant été acquises avec la séquence A. Pour la deuxième cohorte (n = 25), nous avons employé uniquement la séquence A (voir **Tableau 3**) pour l'ensemble des 9 sessions. La durée de chaque essai, et donc de chaque session, était liée à la

## Études expérimentales

vitesse de réponse des participants, nous n'avons donc pas prédéfini le nombre de volumes à acquérir pour chaque session. L'acquisition des volumes était stoppée lorsque les participants avaient complété l'ensemble des essais d'une session.

	Temps de répétition (TR) (secondes)	Temps d'écho (TE) (millisecondes)	Angle de bascule (°C)	Nombre de coupes	Épaisseur de coupe (millimètres)	Facteur d'accélération multi-bandes
Séquence A = <i>echo-planar imaging</i> (EPI)	2.01	25	78	37	2.5	non
Séquence B = multi-bandes	1.20	25	66	44	2.5	oui
Séquence C = multi-échos	1.28	11.00	69	44	2.5	oui
		29.89				

**Tableau 3 : Paramètres d'acquisition pour chaque séquence utilisée en IRM fonctionnelle.** La première cohorte de participants (N=15) a effectué une session de chacune des tâches avec chaque séquence. La deuxième cohorte de participants (N=25) a effectué toutes les sessions avec la séquence A.

- **Pré-traitement des données**

Les données d'IRMf ont été pré-traitées et analysées avec la toolbox SPM12 (Wellcome Trust Center for NeuroImaging, Institute Of Neurology, Londres, Royaume-Uni) dans Matlab 2017a (The MathWorks, Inc., USA). Les images fonctionnelles ont d'abord été réalignées, puis elles ont été normalisées dans l'espace MNI (*Montreal Neurological Institute*) selon la même transformation que l'image anatomique. Enfin, les images fonctionnelles ont été lissées avec un noyau gaussien de FWHM (*Full Width Half-Maximum*) de 8mm.

- **Analyse des données d'IRMf : les modèles utilisés**

Après le prétraitement, nous avons analysé les données à l'aide d'un modèle linéaire généralisé (*general linear model*, GLM) dans SPM12 au premier niveau pour chaque sujet, puis nous avons regardé la significativité des paramètres de notre GLM au second niveau, c'est-à-dire au niveau du groupe de participants. Pour chaque sujet et chaque session, six paramètres de mouvement étaient directement estimés par SPM12 au moment de l'étape de réalignement

## Études expérimentales

du pré-traitement. Ces paramètres de mouvement étaient ensuite rajoutés au sein du GLM en tant que régresseurs de non-intérêt.

Dans notre GLM principal (**GLM1**), nous avons modélisé, séparément pour chaque bloc de chaque session de chaque tâche, le moment où les stimuli apparaissaient à l'écran ainsi que la dérivée temporelle au premier ordre de cet événement à l'aide d'une fonction delta. Cette fonction delta était ensuite modulée par des régresseurs paramétriques. Nous avons d'abord inclus les 3 régresseurs paramétriques de non-intérêt qui expliquaient le temps de délibération dans l'ordre suivant : 1) luminance, 2) nombre de mots des stimuli et 3) la durée de la croix de fixation précédant chaque essai. Ensuite, nous avons rajouté nos variables d'intérêt avec, dans l'ordre, 4) Val, 5) Conf et 6) DT. La formule utilisée pour les variables Val et Conf était la même que celle utilisée pour modéliser le temps de délibération (voir **Tableau 1**). Chacun de ces régresseurs était zscoré. Les régresseurs paramétriques étaient orthogonalisés par SPM en fonction de l'ordre dans lequel ils étaient rentrés, c'est-à-dire que chaque régresseur était orthogonalisé par rapport aux régresseurs précédents. Le temps de délibération inclus dans ce GLM était donc orthogonalisé par rapport aux 5 régresseurs précédents qui corrélaient d'ailleurs tous avec le temps de délibération. Nous avons ensuite cherché à observer les corrélats cérébraux de Val, Conf et DT au second niveau à l'aide d'un t.test contre zéro, après avoir moyenné chacun de ces régresseurs au travers des blocs de chaque tâche et au travers de nos trois tâches. Pour cette moyenne, le régresseur Val correspondant aux blocs d'efforts dans la tâche d'évaluation et dans la tâche de choix 1D-2O était orienté de sorte à être orienté dans la même direction que les récompenses, c'est-à-dire de l'effort le plus pénible vers l'effort le moins pénible.

Nous avons effectué un GLM additionnel, le **GLM2**, dont le but était d'extraire l'activité de nos régions d'intérêt pour chaque essai. Dans cette optique, ce modèle n'incluait aucun modulateur paramétrique mais chaque essai y était modélisé séparément. Nous pouvions ainsi obtenir une moyenne de l'activité d'une région d'intérêt donnée pour chaque essai en vue de voir dans quelle mesure cette activité était ou non corrélée au diamètre pupillaire (voir *Analyse du diamètre pupillaire*).

- **Régions d'intérêt**

Nous avons observé les corrélats neuraux de nos variables d'intérêt (Val, Conf et DT) sur des cartes seuillées à  $p < 0.05$  corrigée pour les comparaisons multiples au niveau des voxels et avons sélectionné le cluster principal localisé dans le cortex préfrontal médian pour chacune

## Études expérimentales

de ces cartes, nous permettant ainsi de définir trois régions : le cortex préfrontal ventromédian (vmPFC) associé à Val, le cortex préfrontal médian (mPFC) associé à Conf et le cortex préfrontal dorsomédian (dmPFC) associé à DT. Nous avons ensuite extrait les coefficients de régression correspondants à chacun de ces clusters pour différents contrastes d'intérêt. Nous avons aussi utilisé ces cartes pour extraire l'activité moyenne, essai par essai, à l'aide du GLM2, de chacun de ces clusters en vue de notre analyse du diamètre pupillaire.

- **Méta-analyse**

Les cartes de la méta-analyse ont été extraites via la plateforme en ligne Neurosynth (<https://neurosynth.org/>). Nous avons extrait 3 cartes correspondants aux mots « *value* » pour voir les corrélats cérébraux de la valeur subjective, « *confidence* » pour la confiance subjective et « *effort* » pour l'effort. Les cartes correspondant aux mots « *value* », « *confidence* » et « *effort* » étaient, respectivement, basées sur 470, 79 et 204 études incluses dans la banque de données de Neurosynth. Nous avons utilisé le test d'uniformité en ligne de Neurosynth qui permet d'afficher les aires cérébrales systématiquement présentes avec le mot clé utilisé. Pour chaque carte, nous avons extraits les clusters survivants, après une correction pour les comparaisons multiples au niveau des clusters pour un taux de faux positifs (*false discovery rate*) inférieur à 0.01 d'après la correction en ligne de Neurosynth. Nous avons ensuite binarisé les cartes afin que tous les clusters survivant à ce seuil soit affichés avec la même couleur pour chacune des cartes.

### I.2.e Données oculaires

- **Acquisition des données oculaires**

La position du regard et le diamètre pupillaire de l'œil droit ont été enregistrés via l'appareil EyeLink 1000 plus (SR Research). Les données ont été acquises à une fréquence de 1000Hz. L'appareil a été calibré avant de démarrer l'expérience, une fois que le sujet était installé dans le scanner.

- **Pré-traitement de la pupille**

Les données extraites ont été analysées dans Matlab 2017a (The MathWorks, Inc., USA). Du fait du clignement des yeux ou d'une perte de signal par l'appareil, le signal ne peut pas être intégralement enregistré en continu. Une interpolation cubique a été effectuée afin de compenser pour ces périodes temporelles où le signal n'était pas enregistré. Toute tâche où, pour un participant donné, plus de 50% des données n'ont pas pu être enregistrées, a été rejetée des analyses. Ensuite, les données de diamètre pupillaire ont été filtrées par un filtre passe-

## Études expérimentales

bande (de  $\frac{1}{128}$ Hz à 1Hz). Enfin, le diamètre pupillaire a été zscoré sur l'ensemble de chaque session pour chaque participant. Les données n'ayant pas pu être acquises pour certaines sessions de certains participants ou n'étant pas d'assez bonne qualité, nous avons dû exclure certains participants des analyses du diamètre pupillaire. En maximisant le nombre de participants pour chaque tâche, nous avons pu inclure 30 participants pour la tâche d'évaluation, 29 pour la tâche de choix 1D-2O et 28 pour la tâche de choix 1O-2D.

### • Analyse du diamètre pupillaire

Le diamètre pupillaire a été analysé pour chaque tâche dans Matlab 2017a (The MathWorks, Inc., USA). Après le pré-traitement, nous avons extrait, pour chaque tâche et pour chaque sujet le diamètre pupillaire à chaque essai. Nous avons regroupé les données de deux façons : 1) par rapport au moment d'apparition des stimuli à l'écran et 2) par rapport au moment de la réponse, considérée comme le moment où les participants commencent à bouger le curseur pour la tâche d'évaluation et le moment du choix pour les deux autres tâches. Pour le premier cas, nous avons ainsi extrait l'activité de 500 millisecondes avant l'apparition des stimuli jusqu'à 3.5 secondes après leur apparition. Pour le second cas, nous avons extrait l'activité de 2.5 secondes avant la réponse jusqu'à 1.5 secondes après la réponse. La même fenêtre d'extraction a été utilisée pour chacune des tâches. De plus, quelle que soit l'extraction effectuée, nous avons moyenné le diamètre pupillaire pendant les 200 millisecondes précédant l'apparition des stimuli avant chaque essai et nous avons soustrait cette valeur au diamètre pupillaire pour chaque essai afin de corriger pour la valeur basale du diamètre pupillaire.

Une fois l'activité ainsi extraite, nous avons effectué un GLM pour chaque tâche pour chaque point temporel. Ce GLM incluait d'abord une constante pour chaque bloc dans la tâche d'évaluation et dans la tâche de choix 1D-2O et une constante globale pour la tâche de choix 1O-2D. Puis nous avons inclus nos trois variables de non-intérêt : 1) la luminance, 2) le nombre de mots associés aux stimuli et 3) la durée de présentation de la croix de fixation avant le début de l'essai. Ensuite, nous avons ajouté nos trois variables d'intérêt : 4) Val, 5) Conf et 6) DT. Une fois toutes ces variables rentrées, nous avons rajouté 3 variables en vue de voir si l'activité de nos régions d'intérêt, sélectionnées sur la base de leur corrélation avec Val, Conf et DT (voir *Méta-analyse*

*Les cartes de la méta-analyse ont été extraites via la plateforme en ligne Neurosynth (<https://neurosynth.org/>). Nous avons extrait 3 cartes correspondants aux mots « value » pour voir les corrélats cérébraux de la valeur subjective, « confidence » pour la confiance subjective*

## Études expérimentales

et « effort » pour l'effort. Les cartes correspondant aux mots « value », « confidence » et « effort » étaient, respectivement, basées sur 470, 79 et 204 études incluses dans la banque de données de Neurosynth. Nous avons utilisé le test d'uniformité en ligne de Neurosynth qui permet d'afficher les aires cérébrales systématiquement présentes avec le mot clé utilisé. Pour chaque carte, nous avons extraits les clusters survivants, après une correction pour les comparaisons multiples au niveau des clusters pour un taux de faux positifs (*false discovery rate*) inférieur à 0.01 d'après la correction en ligne de Neurosynth. Nous avons ensuite binarisé les cartes afin que tous les clusters survivant à ce seuil soit affichés avec la même couleur pour chacune des cartes.

) était corrélée au diamètre pupillaire au travers des 3 tâches étudiées. Nous avons ainsi rajouté l'activité 7) du cluster dans le cortex préfrontal ventromédian (vmPFC) lié à Val d'après le GLM1 (voir **Figure 27**), 8) du cluster dans le cortex préfrontal médian (mPFC) lié à Conf d'après le GLM1 (voir **Figure 27**) et 9) du cluster dans le cortex préfrontal dorsomédian (dmPFC) lié au DT d'après le GLM1 (voir **Figure 27**). L'activité, essai par essai, de chacune de ces régions d'intérêt a été extraite à l'aide du GLM2 (voir *Analyse des données d'IRMf : les modèles utilisés*). Chaque variable était zscorée. Ensuite, les régresseurs étaient tous orthogonalisés de manière sérielle dans l'ordre indiqué précédemment. Ceci implique que les régresseurs liés à l'activité neurale, étant placés à la fin, étaient donc orthogonalisés par rapport à Val, à Conf et au DT. Une fois le GLM effectué pour chacun des points temporels ayant été extraits pour un sujet donné, nous avons moyenné la série temporelle de beta lié à chacun des régresseurs du GLM à travers les participants. Puis nous avons voulu corriger pour les comparaisons multiples afin de corriger pour le nombre de tests ayant été effectué sur l'ensemble de la série temporelle analysée. Pour ce faire, pour une série de beta donnée correspondant à l'un des régresseurs du GLM, nous avons appliqué la théorie des champs aléatoires à l'aide de la fonction RFT\_GLM\_contrast.m de la VBA toolbox (Daunizeau et al., 2014). Cette fonction nous a permis d'identifier, à l'aide d'un test F, des clusters temporels où les betas étaient significativement différents de zéro après correction pour les comparaisons multiples. Pour l'affichage des données, les séries de betas ont été lissées sur une fenêtre de 100 millisecondes.

## I.3 Résultats

### I.3.a Le comportement

- **Tâche d'évaluation**

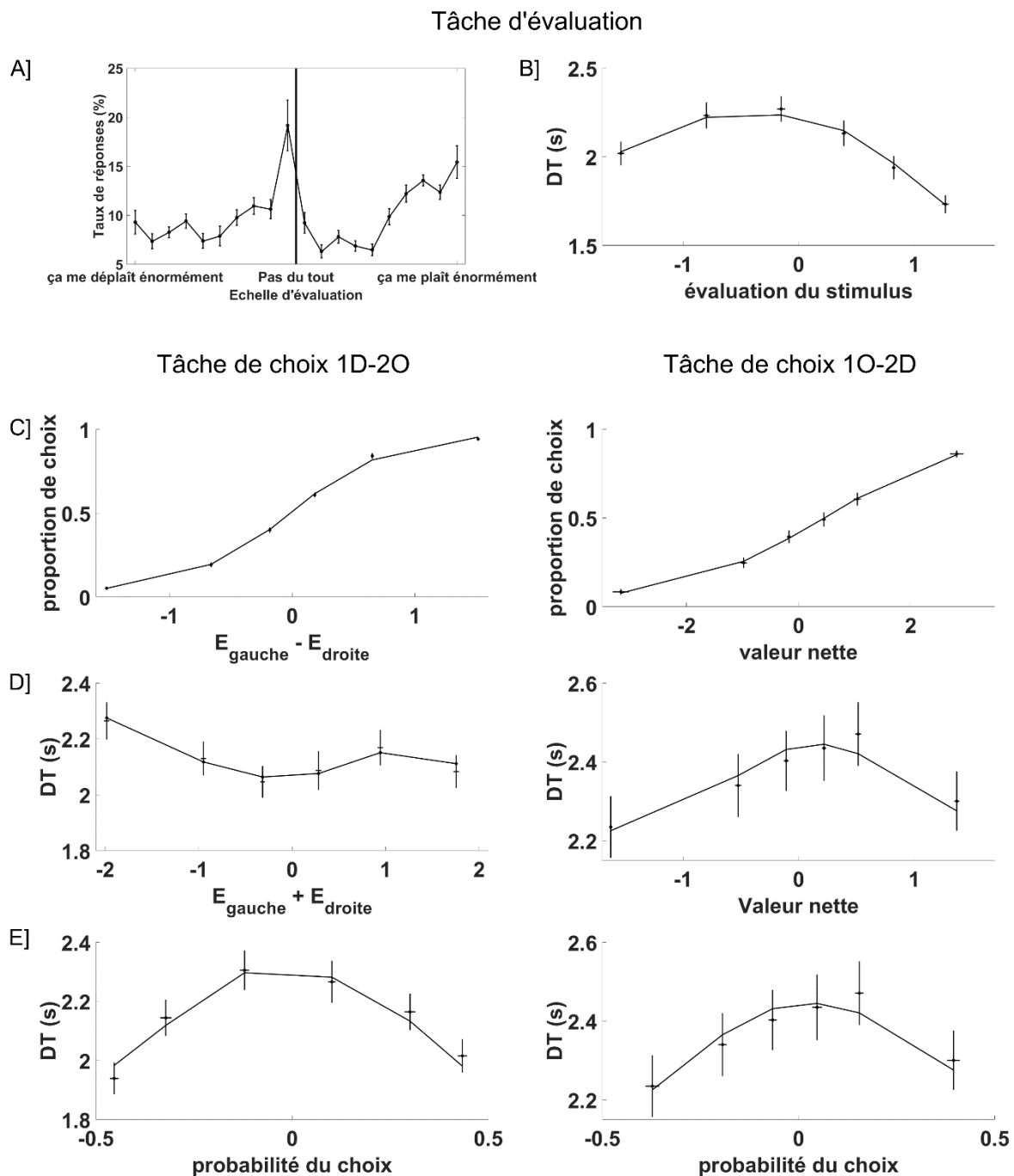
Les participants ( $n = 39$  au total) ont d'abord effectué une tâche d'évaluation (voir **Figure 25**). Cette tâche était divisée en trois blocs équitablement répartis au sein de 3 sessions d'IRMf. Chaque bloc contenait 72 items à évaluer individuellement : des items de récompenses présentés avec un texte et une image, des items de récompenses présentés avec un texte uniquement et des items d'efforts présentés avec un texte uniquement. Nous avons choisi d'avoir deux modes de présentation distincts pour les récompenses parce que nous voulions vérifier la généralité du processus d'évaluation avec des entrées requérant plus ou moins d'imagination sur la base d'une étude précédente qui a montré que (Lebreton et al., 2013). L'ordre des blocs était constant chez un participant donné mais il était contrebalancé à travers les participants. Dans le cas des récompenses, nous demandions aux participants de nous indiquer à quel point ils seraient satisfaits si on leur donnait la récompense affichée à l'écran à l'instant. Pour les efforts, ils devaient de même indiquer à quel point cela leur déplairait s'ils avaient dû exécuter l'effort affiché à l'écran au moment où on leur demandait d'effectuer leur évaluation. Nous avons mélangé des récompenses alimentaires et des récompenses non-alimentaires, des objets principalement, et aussi des efforts physiques et mentaux dans nos différents items. Il n'y avait pas de chiffre sur l'échelle mais des labels aux extrémités pour indiquer aux participants le sens de l'échelle. Les évaluations se faisaient de manière pseudo-continue sur cette échelle. L'extrême gauche de l'échelle correspondait systématiquement à un point d'indifférence (« ça ne me (dé)plairait pas du tout ») alors que l'autre extrémité correspondait à une appréciation extrême, dans le cas des récompenses, et à une aversion extrême, dans le cas des efforts.

Les évaluations zscorées ont été utilisées comme un proxy pour la valeur des stimuli (Val) dans cette tâche. L'évaluation zscorée mise au carré a été utilisée comme proxy de la confiance subjective (Conf) à la suite de plusieurs autres études ayant montré que cette mesure est très similaire à une évaluation de confiance exprimée (De Martino et al., 2017; Lebreton et al., 2015; Lopez-Persem et al., 2020). Dans cette tâche, le temps de délibération (DT) a été défini comme le temps mis entre l'apparition d'un stimulus à l'écran et le premier appui du participant en vue de bouger le curseur sur l'échelle.



## Études expérimentales

Nous avons effectué une régression linéaire du DT avec les régresseurs Val et Conf en plus de régresseurs de non-intérêt, comme la luminance, le nombre de mots et la durée d'affichage de la croix de fixation précédant la croix de fixation. Nous avons pu voir que, dans tous les blocs, DT variait significativement avec Val ( $R_{ti}: \beta_{Val} = -0.208 \pm 0.023$  (moyenne  $\pm$  erreur type de la moyenne à travers les participants),  $p = 4 \cdot 10^{-11}$ ;  $R_t: \beta_{Val} = -0.169 \pm 0.019$ ,  $p = 6 \cdot 10^{-11}$ ;  $E_t: \beta_{Val} = 0.261 \pm 0.028$ ,  $p = 2 \cdot 10^{-11}$ ) et avec Conf ( $R_{ti}: \beta_{Conf} = -0.174 \pm 0.025$ ,  $p = 3 \cdot 10^{-8}$ ;  $R_t: \beta_{Conf} = -0.188 \pm 0.028$ ,  $p = 7 \cdot 10^{-8}$ ;  $E_t: \beta_{Conf} = -0.130 \pm 0.040$ ;  $p = 0.0024$ ). En résumé, les participants répondaient plus rapidement lorsque l'item évalué était plus appétitif dans le cas des récompenses ou moins aversif dans le cas des efforts. Ils répondaient aussi plus rapidement lorsque leur confiance subjective, déduite des essais où ils répondaient aux extrêmes de l'échelle, était plus élevée.



**Figure 26 : Comportement dans les tâches basées sur la valeur.**

A] Répartition des évaluations. L'échelle des évaluations a été divisée en 10 bins de taille équivalente pour les récompenses (partie droite du graphique) et pour les efforts (partie gauche du graphique). Pour les efforts, les participants devaient répondre à la question « A quel point ça me déplaît » sur une échelle continue dont l'extrême gauche sur l'écran était affublée de l'expression « Pas du tout » et l'extrême droite sur l'écran de l'expression « Enormément ». Pour les récompenses, les participants devaient répondre à la question « A quel point ça me plaît » sur la même échelle continue dont les extrêmes étaient affublés des mêmes expressions. Pour que l'affichage soit plus cohérent, la figure représente les évaluations classées des efforts les plus pénibles à l'extrême gauche vers les récompenses les plus agréables à l'extrême droite de la figure.

## Études expérimentales

B] Temps de délibération (DT) dans la tâche d'évaluation en fonction de l'évaluation du stimulus centrée. Une valeur positive représente ainsi une récompense plus appréciée, ou un effort moins pénible. 0 est la moyenne à travers les essais.

C] Figure de gauche : Proportion de choix de l'option de gauche en fonction de la différence entre les évaluations des deux options ( $E_{\text{gauche}} - E_{\text{droite}}$ ) dans la tâche de choix 1D-2O. Figure de droite : Proportion de choix où l'offre proposée à l'écran a été acceptée dans la tâche de choix 1O-2D en fonction de la valeur nette ( $\beta_0 + \beta_R \cdot E_R + \beta_E \cdot E_E$ ). Dans le cas des efforts, les évaluations ont été inversées de sorte à être orientées dans le même sens que les récompenses ( $E_{\text{gauche}} - E_{\text{droite}} > 0$  implique que l'effort de droite a été évalué comme plus pénible) avant de moyenner les données des trois blocs ensemble.

D] Temps de délibération (DT) en fonction de la somme des évaluations des stimuli dans la tâche de choix 1D-2O à gauche et de la valeur nette dans la tâche de choix 1O-2D à droite. Pour la tâche de choix 1D-2O, la somme des évaluations a été définie dans le sens allant des récompenses les moins appétitives vers les récompenses les plus appétitives et des efforts les plus pénibles vers les efforts les moins pénibles.

E] Temps de délibération (DT) en fonction de la probabilité de choix. Dans la tâche de choix 1D-2O, à gauche, la probabilité de choix est l'output de la fonction softmax pour les choix de l'option de gauche. Cette probabilité a été centrée à travers les essais de sorte que 0 correspond à la moyenne à travers les essais de cette variable. Dans la tâche de choix 1O-2D à droite, la probabilité de choix est l'output de la fonction softmax pour les choix d'accepter l'offre proposée à chaque essai, centrée à travers les essais de sorte que 0 correspond à la moyenne de cette variable à travers les essais.

Les barres représentent la moyenne ainsi que l'erreur type de la moyenne à travers les 39 participants. Dans la figure (A), les bins ont été établis en découpant l'échelle d'évaluation en 10 segments et la courbe relie simplement les moyennes du taux de réponse pour chaque bin. Dans les autres figures, les données ont été réparties en 6 bins contenant le même nombre d'essais pour chaque participant et les courbes représentent le fit des données pour les bins correspondants.

- **Tâches de choix**

Après la tâche d'évaluation, les participants ont effectué une série de choix binaires, d'abord des choix 1D-2O, puis des choix 1O-2D. La tâche d'évaluation précédait systématiquement les tâches de choix car nous utilisons les évaluations fournies par les participants lors de la première tâche afin de contrôler la difficulté des choix sur la base de la différence entre les évaluations des options présentées. D'abord les participants ont effectué la tâche de choix 1D-2O divisée en 3 blocs (voir **Figure 25**). Chaque bloc était là aussi associé à une seule dimension (1D), récompense ou effort, parmi laquelle deux options (2O) étaient présentées à chaque essai parmi les items évalués lors de la première tâche. Les participants devaient alors choisir laquelle des deux options ils préféreraient obtenir, dans le cas des récompenses, et quel effort ils préféreraient devoir effectuer, dans le cas des efforts. Nous avons fait en sorte que les deux items qui étaient présentés au sein de chaque choix appartiennent non

## Études expérimentales

seulement à la même dimension, mais aussi à la même sous-catégorie (récompense alimentaire ou non-alimentaire, effort physique ou mental) afin d'éviter que la délibération soit court-circuitée par une préférence générale pour une catégorie par rapport à une autre. Le mode de présentation, avec un texte seul ou avec un texte accompagné d'une image, était aussi le même pour les deux options. Chaque option était présentée deux fois afin d'avoir le même nombre total d'essais que dans la tâche d'évaluation au sein de chaque bloc. Chaque bloc contenait ainsi 72 choix à effectuer répartis équitablement sur les 3 sessions de scanner. L'ordre des blocs pour un sujet donné était le même que dans la tâche d'évaluation. Enfin, dans la tâche de choix 1O-2D, les participants devaient indiquer s'ils étaient prêts à exercer un effort en vue d'obtenir une récompense (voir **Figure 25**). Seuls les items présentés sous forme de texte étaient utilisés dans cette tâche. Cette tâche était constituée de 144 choix au total répartis équitablement en 3 sessions dans le scanner pour chaque participant.

Dans la tâche de choix 1D-2O, nous avons voulu vérifier que les choix étaient bien prédits par la différence entre les évaluations attribuées à chaque option lors de la tâche d'évaluation. Nous avons utilisé une régression logistique pour modéliser les choix des participants en fonction de la différence de valeurs entre les deux options. Nous avons pu voir que la différence de valeurs était bien un prédicteur significatif des choix (voir **Figure 26**) dans chacun des blocs ( $R_{ii}$ :  $\beta_{DV} = 3.383 \pm 0.273$ ,  $p = 7 \cdot 10^{-15}$ ;  $R_i$ :  $\beta_{DV} = 2.669 \pm 0.155$ ,  $p = 2 \cdot 10^{-19}$ ;  $E_i$ :  $\beta_{DV} = 2.278 \pm 0.157$ ,  $p = 4 \cdot 10^{-17}$ ). Dans la tâche de choix 1O-2D, nous avons voulu vérifier que les choix effectués par les participants prenaient bien en compte les 2 dimensions associées à chaque option. Nous avons ainsi effectué une régression logistique des choix dans cette tâche en fonction de l'évaluation des récompenses ( $E_R$ ) et de l'évaluation des efforts ( $E_E$ ) associées à chaque option. Nous avons vu ainsi que les choix étaient bien impactés aussi bien par la valeur des récompenses ( $\beta_R = 1.502 \pm 0.085$ ,  $p = 6 \cdot 10^{-20}$ ) que par la valeur des efforts ( $\beta_E = -1.117 \pm 0.079$ ,  $p = 1 \cdot 10^{-16}$ ). Dans cette tâche, les participants ont, en moyenne, accepté  $44.71 \pm 2.64$  % (moyenne  $\pm$  erreur type de la moyenne à travers les participants) des offres qui leurs étaient proposées.

Dans ces tâches de choix aussi, nous avons modélisé le DT, considéré comme le temps mis entre l'apparition des stimuli à l'écran et le moment de la réponse, en fonction de nos variables de non-intérêt (luminance, nombre de mots et durée d'affichage de la croix de fixation précédant l'essai) et des variables Val et Conf (voir **Tableau 1**). Pour la tâche de choix 1D-2O, Val était définie comme la somme des évaluations des deux options affichées. Pour la tâche de choix 1O-2D, Val était définie comme la valeur nette issue de l'intégration des coûts et des

## Études expérimentales

bénéfices en une combinaison linéaire expliquant les choix. Dans les deux cas, la probabilité du choix a été modélisée à l'aide d'une régression logistique (fonction softmax de la variable décisionnelle). La confiance subjective dans la réponse (Conf) a été définie, par analogie avec la tâche d'évaluation, comme la probabilité de choix centrée mise au carré. La régression linéaire de DT en fonction de Val nous a permis de voir que, dans la tâche de choix 1D-2O, le DT était d'autant plus rapide que Val était élevée ( $R_{ti}: \beta_{Val} = -0.061 \pm 0.010, p = 3 \cdot 10^{-7}$ ;  $R_t: \beta_{Val} = -0.061 \pm 0.010, p = 3 \cdot 10^{-7}$ ;  $E_t: \beta_{Val} = -0.048 \pm 0.013, p = 8 \cdot 10^{-4}$ ) (voir **Figure 26**). Cette relation n'était cependant pas vérifiée dans le cas de la tâche de choix 1O-2D ( $\beta_{Val} = -0.033 \pm 0.024, p = 0.172$ ). De plus, DT était d'autant plus rapide que Conf était élevée aussi bien dans la tâche de choix 1D-2O ( $R_{ti}: \beta_{Conf} = -1.738 \pm 0.201, p = 2 \cdot 10^{-10}$ ;  $R_t: \beta_{Conf} = -1.975 \pm 0.184, p = 4 \cdot 10^{-13}$ ;  $E_t: \beta_{Conf} = -1.731 \pm 0.221, p = 2 \cdot 10^{-9}$ ) (voir **Figure 26**) que dans la tâche de choix 1O-2D ( $\beta_{Conf} = -1.148 \pm 0.145, p = 1 \cdot 10^{-9}$ ) (voir **Figure 26**). Dans la plupart de nos tâches, DT était ainsi d'autant plus rapide que les options étaient appétitives (ou moins aversives). De plus, dans la tâche d'évaluation comme dans les tâches de choix, DT était d'autant plus rapide que nos participants étaient confiants dans leurs décisions.

### I.3.b Le cortex préfrontal médian et les différents aspects du compromis coûts/bénéfices

Le but de l'analyse des données d'IRMf était de dissocier d'un côté les variables liées à l'évaluation au premier ordre, c'est-à-dire l'estimation de la valeur subjective associée aux récompenses et aux efforts considérés à chaque essai, et les variables liées à la méta-décision, c'est-à-dire la confiance et le temps de délibération, au travers de nos tâches d'évaluation et de choix. Nous avons vu en effet qu'une méta-analyse des études d'IRMf sur la plateforme en ligne Neurosynth reflète que la valeur, la confiance et le DT sont tous associés à l'activité du vmPFC et du dmPFC (voir **Figure 27A**). Nous avons ainsi défini un modèle linéaire général (GLM) où l'apparition des stimuli était modulée par nos trois paramètres d'intérêt : Val, Conf et DT. Les paramètres de non-intérêt qui influençaient le DT, soit la luminance, le nombre de mots et la durée d'apparition de la croix de fixation avant le début de l'essai, ont aussi été inclus en tant que modulateurs de la convolution de la fonction de réponse hémodynamique au moment de l'apparition des stimuli avant nos variables d'intérêt. Toutes les variables du GLM étaient orthogonalisées de manière sérielle de sorte que nos variables d'intérêt étaient orthogonalisées par rapport à ces variables de non-intérêt. Le DT, en particulier, était ainsi orthogonalisé non seulement par rapport à ces variables de non-intérêt, mais aussi par rapport à Val et à Conf.

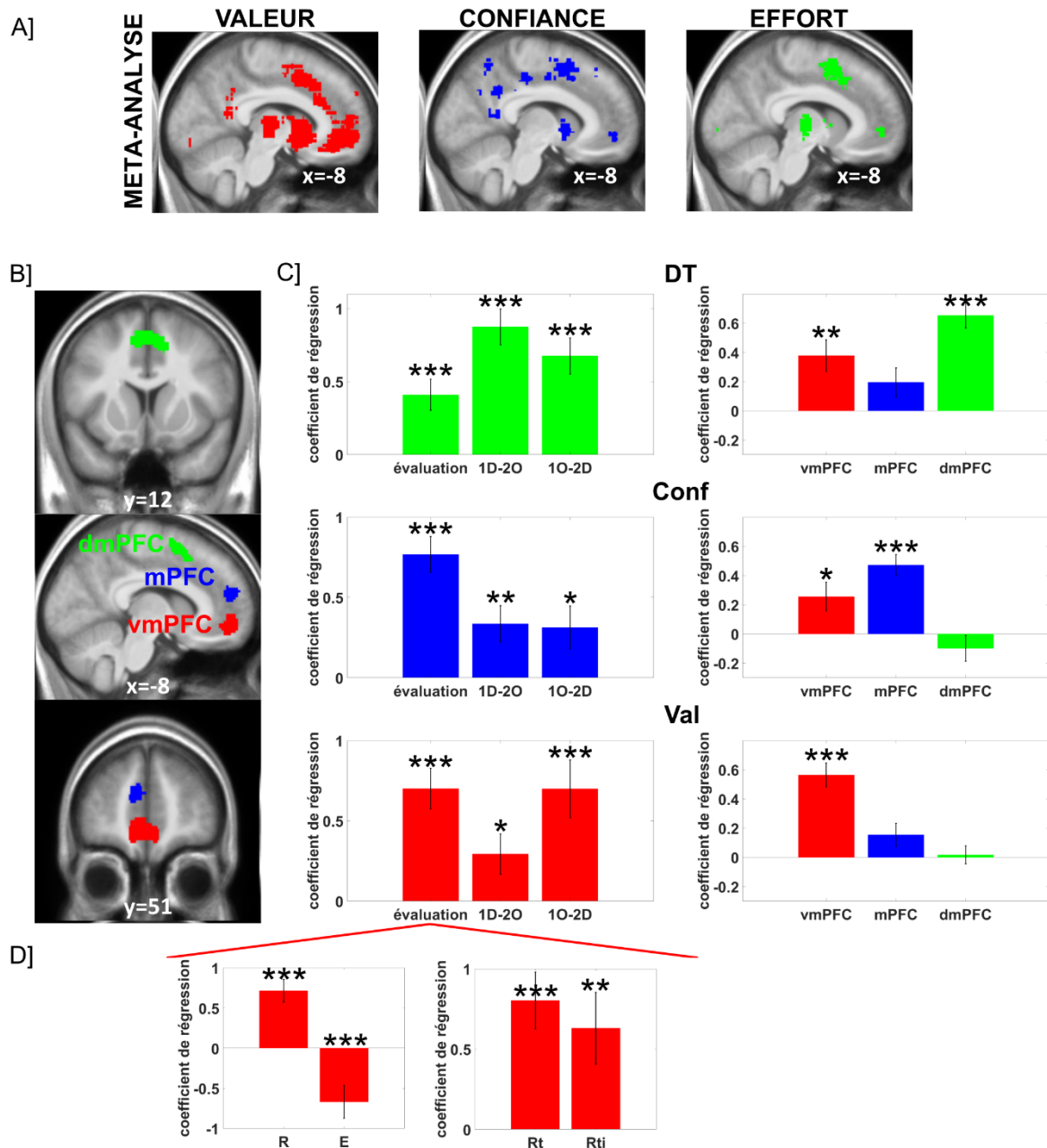
## Études expérimentales

Régresseur	Valeur			
Région	x	y	z	Nombre de voxels
Gyrus frontal médian orbital gauche	-10	48	-12	364
Gyrus frontal milieu orbital gauche	-28	36	-14	57
Gyrus lingual droit	16	-70	-6	64
Cortex cingulaire antérieure gauche	-8	38	6	16
Précuneus gauche	-6	-54	14	22
Régresseur	Confiance			
Région	x	y	z	Nombre de voxels
Gyrus frontal supérieur médian	-8	52	18	128
Gyrus temporal inférieur gauche	-46	2	-36	13
Gyrus supramarginal gauche	-62	-40	32	56
Gyrus temporal milieu gauche	-56	-26	-10	63
Gyrus temporal milieu gauche	-46	-64	12	22
Noyau caudé gauche	-12	14	-12	14
Hippocampe droit	32	-12	-14	10
Régresseur	DT			
Région	x	y	z	Nombre de voxels
Insula gauche	-30	26	4	110
Aire motrice supplémentaire gauche	10	12	48	365
Insula droite	32	26	4	95
Gyrus frontal inférieur Triangulaire gauche	-40	22	24	242
Gyrus lingual gauche	-18	-88	-10	23

**Table 1 : Tables d'activation pour chacun des régresseurs d'intérêt au travers des tâches d'évaluation et de choix avec, dans l'ordre, Val, Conf et DT.** Les tables montrent la liste des clusters identifiés pour chaque variable d'intérêt pour un seuil de p.valeur < 0.05 corrigée pour les comparaisons multiples au niveau des voxels. Les clusters de taille inférieure à 10 voxels ont été exclus. Les appellations des différentes aires ont été établies sur la base des coordonnées correspondantes dans l'atlas AAL (Automated Anatomical Labeling). Les coordonnées indiquées sont dans l'espace MNI. Les p.valeur rapportées correspondent à la p.valeur de chaque cluster après correction pour les comparaisons multiples au niveau du cluster.

## Études expérimentales

Après correction, au niveau des voxels, pour les comparaisons multiples au niveau de la famille (*Family Wise Error Rate*), nous avons identifié trois clusters significatifs dans le cortex préfrontal médian avec chacune de nos variables d'intérêt (voir **Figure 27**). Val était corrélée avec l'activité du cortex préfrontal ventromédian (vmPFC) avec un pic aux coordonnées MNI (-10 ; 48 ; -12). Conf était corrélée avec l'activité du cortex préfrontal médian (mPFC) surjacent au vmPFC avec un pic aux coordonnées MNI (-8 ; 52 ; 18) et DT était corrélé avec l'activité du cortex préfrontal dorsomédian (dmPFC) avec un pic aux coordonnées MNI (10 ; 12 ; 48). On peut noter que chacun de ces clusters d'intérêt s'accompagnait aussi d'autres aires plus ou moins attendues au vu de la corrélation testée (voir **Table 1**). Val était ainsi aussi corrélée à l'activité du cortex postérieur cingulaire (pic aux coordonnées MNI (-6 ; -54 ; 14)), du striatum ventral (deux clusters de 1 voxel chacun, d'où leur absence de la table au seuil utilisé, avec un pic aux coordonnées MNI (-12 ; 12 ; -2) à gauche et (8 ; 8 ; -4) à droite) et du cortex orbitofrontal latéral gauche (pic aux coordonnées MNI (-28 ; 36 ; -14)). Conf était aussi corrélée à l'activité du striatum ventral avec un pic aux coordonnées MNI (-12 ; 14 ; -12) à gauche. Enfin, le DT était aussi corrélé à l'activité de l'insula antérieure bilatéralement (pics aux coordonnées MNI (-30 ; 26 ; 4) à gauche et (32 ; 26 ; 4) à droite) ainsi que du cortex préfrontal dorsolatéral gauche (pic aux coordonnées MNI (-40 ; 22 ; 24)).



**Figure 27 : Résultats neuraux pour Val, Conf et DT.**

A] Méta-analyse des études d'IRMf. Les cartes statistiques (coupes sagittales en  $X = -8$ ) ont été extraites sur la base de la plateforme en ligne [Neurosynth](https://www.neurosynth.org/) avec les termes « value » (470 études), « confidence » (79 études) et « effort » (204 études). Les clusters dans le cortex préfrontal médian sont similaires au travers des 3 mots clés. Tous les clusters passant les seuils de  $p < 0.01$  (*False Discovery Rate*) avec le test d'uniformité sont montrés avec une couleur unie pour chaque mot clé (rouge pour la valeur, bleu pour la confiance et vert pour l'effort). Les cartes ont été superposées à la moyenne des 38 cerveaux de cette étude.

B] Corrélats cérébraux de la valeur (Val), la confiance (Conf) et le temps de délibération (DT) dans cette étude au sein du cortex préfrontal médian. Cette carte représente les 3 clusters dans le cortex préfrontal médian sur lesquels nous nous sommes concentrés. La coupe sagittale a été effectuée au même niveau que les images issues de Neurosynth ( $x = -8$ ) et superposée sur la moyenne des 38 cerveaux inclus dans l'étude. Chacun de ces clusters a été extrait, en lien avec



## Études expérimentales

le terme correspondant (Val en rouge, Conf en bleu et DT en vert) sur la base de la carte correspondante seuillée à  $p < 0.05$  corrigé au niveau des voxels pour les comparaisons multiples au niveau de la famille (*Family Wise Error Rate*).

C] Décomposition des contrastes par variable (Val, Conf et DT) par tâche pour la région d'intérêt correspondant à chacune des variables (le vmPFC pour Val en rouge, le mPFC pour Conf en bleu et le dmPFC pour DT en vert) à gauche. A droite, décomposition des contrastes par variable (Val, Conf et DT) par région d'intérêt. D] Décomposition des contrastes obtenus pour Val dans le vmPFC pendant la tâche d'évaluation par dimension (récompense ou effort) et par mode de présentation (récompense présentée avec ( $R_{ii}$ ) ou sans image illustrative ( $R_t$ )).

Les barres et les barres d'erreurs montrent, respectivement, la moyenne et l'erreur type de la moyenne à travers les participants. Les étoiles indiquent la significativité à la suite d'un t-test des données contre zéro : \*p.value < 0.05 ; \*\*p.value < 0.01 ; \*\*\*p.value < 0.005 ; \*\*\*\*p.value < 0.001.

Nous avons ensuite effectué des t-tests post-hoc afin d'étudier le lien entre nos variables d'intérêt et l'activité au sein de nos trois régions d'intérêt au sein du cortex préfrontal médian. Tout d'abord, nous avons voulu vérifier que l'activité au sein de chacune de ces régions n'était pas liée à une tâche spécifique. Nous avons ainsi regardé, pour chacune de nos variables d'intérêt, c'est-à-dire Val, Conf et DT, que le lien avec la région d'intérêt identifiée dans le cortex préfrontal médian était bien stable au travers des tâches (voir **Figure 27**). Nous avons pu vérifier que c'était bien le cas pour le lien entre l'activité du vmPFC et Val (tâche d'évaluation :  $\beta_{\text{Val}} = 0.701 \pm 0.126$ ,  $p = 2 \cdot 10^{-6}$  ; tâche de choix 1D-2O:  $\beta_{\text{Val}} = 0.293 \pm 0.126$ ,  $p = 0.025$  ; tâche de choix 1O-2D :  $\beta_{\text{Val}} = 0.700 \pm 0.180$ ,  $p = 4 \cdot 10^{-4}$ ), pour le lien entre l'activité du mPFC et Conf (tâche d'évaluation:  $\beta_{\text{Conf}} = 0.768 \pm 0.111$ ,  $p = 3 \cdot 10^{-8}$  ; tâche de choix 1D-2O :  $\beta_{\text{Conf}} = 0.335 \pm 0.114$ ,  $p = 0.006$  ; tâche de choix 1O-2D:  $\beta_{\text{Conf}} = 0.313 \pm 0.133$ ,  $p = 0.024$ ) et pour le lien entre l'activité du dmPFC et DT (tâche d'évaluation:  $\beta_{\text{DT}} = 0.410 \pm 0.106$ ,  $p = 4 \cdot 10^{-4}$  ; tâche de choix 1D-2O:  $\beta_{\text{DT}} = 0.875 \pm 0.123$ ,  $p = 2 \cdot 10^{-8}$  ; tâche de choix 1O-2D:  $\beta_{\text{DT}} = 0.676 \pm 0.125$ ,  $p = 4 \cdot 10^{-6}$ ). Ensuite, nous avons vérifié que chaque variable corrélait plus fortement avec la partie du cortex préfrontal médian correspondante par rapport aux deux autres clusters (voir **Figure 27**). Nous avons ainsi vu que l'activité du vmPFC était en effet significativement mieux corrélée à Val que l'activité du mPFC ( $p = 9 \cdot 10^{-9}$ ) ou que celle du dmPFC ( $p = 5 \cdot 10^{-8}$ ). De même, l'activité du mPFC était significativement mieux corrélée à Conf que celle du vmPFC ( $p = 0.004$ ) ou du dmPFC ( $p = 2 \cdot 10^{-7}$ ). Le DT, enfin, corrélait mieux avec l'activité du dmPFC qu'avec celle du vmPFC ( $p = 0.021$ ) ou celle du mPFC ( $p = 2 \cdot 10^{-4}$ ). Il faut d'ailleurs préciser que la triple dissociation n'implique pas que chacune de nos variables d'intérêt n'était encodée qu'en une seule de nos régions d'intérêt. Par exemple, l'activité du vmPFC était aussi significativement corrélée à Conf ( $\beta_{\text{Conf}} = 0.256 \pm 0.098$ ,  $p = 0.013$ ) et au DT ( $\beta_{\text{DT}} = 0.377 \pm$

## Études expérimentales

0.108,  $p = 0.001$ ). La triple dissociation que nous avons observée sur les cartes cérébrales correspondait ainsi bien à des différences significatives entre les variables computationnelles et des associations stables et significatives au travers de nos différentes tâches.

Pour vérifier que nos résultats étaient robustes, nous avons aussi effectué quelques variantes des GLM présentés ici. Nous avons notamment essayé d'utiliser des formules alternatives pour définir la valeur subjective Val dans la tâche de choix 1D-2O. Le but était de voir s'il serait possible de mieux capter l'activité du vmPFC avec ces définitions différentes de la valeur subjective pendant un choix comme celui opéré dans la tâche de choix 1D-2O. Nous avons notamment remplacé la somme des valeurs par la valeur choisie ou la différence entre la valeur de l'option choisie et la valeur de l'option non-choisie dans la tâche de choix 1D-2O mais les résultats étaient relativement équivalents (vmPFC avec la valeur de l'option choisie :  $\beta_{\text{Val}} = 0.366 \pm 0.157$ ,  $p = 0.025$  ; vmPFC avec la différence entre la valeur de l'option choisie et la valeur de l'option non-choisie :  $\beta_{\text{Val}} = 0.172 \pm 0.139$ ,  $p = 0.224$ ). Nous avons aussi essayé d'utiliser un modèle computationnel dont le principe est d'inférer les valeurs cachées qui dirige le comportement à la fois en se basant sur les notes attribuées aux différents items dans la tâche d'évaluation et aux choix effectués dans les deux autres tâches. Là encore les résultats étaient relativement équivalents (voir Appendice : *Utiliser un modèle computationnel pour inférer les valeurs cachées associées au comportement*). Pour Conf, nous avons voulu nous assurer que l'encodage plus dorsal n'était pas lié à l'orthogonalisation de Conf par rapport à Val. Nous avons ainsi testé une variante où nous avons inclus la confiance sans la précéder de la valeur dans le GLM. Cependant, le cluster est resté centré sur la même localisation spatiale relativement plus dorsale que le vmPFC avec le même pic aux coordonnées MNI (-8 ; 52 ; 18). Pour DT, nous avons voulu vérifier si la corrélation entre le dmPFC et DT était simplement due à une activation du dmPFC tant que l'essai était en cours ou s'il s'agissait bien d'une activation proportionnelle à la quantité d'efforts alloués dont le DT ne serait que l'un des reflets. Nous avons ainsi remplacé la fonction delta utilisée pour convoluer la réponse hémodynamique au moment de l'apparition des stimuli à l'écran par une fonction boxcar dont la durée était égale à DT. Nous avons alors pu observer que nos résultats montrant une corrélation du dmPFC avec le DT restaient inchangés au sein de chacune des tâches (tâche d'évaluation :  $\beta_{\text{DT}} = 0.410 \pm 0.106$ ,  $p = 4 \cdot 10^{-4}$  ; tâche de choix 1D-2O :  $\beta_{\text{DT}} = 0.875 \pm 0.123$ ,  $p = 2 \cdot 10^{-8}$  ; tâche de choix 1O-2D :  $\beta_{\text{DT}} = 0.676 \pm 0.125$ ,  $p = 4 \cdot 10^{-6}$ ). L'activation du dmPFC en lien avec le DT semble ainsi bien refléter une modulation en amplitude en lien avec la quantité de ressources allouées pour la prise de décision et non pas seulement un signal actif tant que l'essai a lieu.

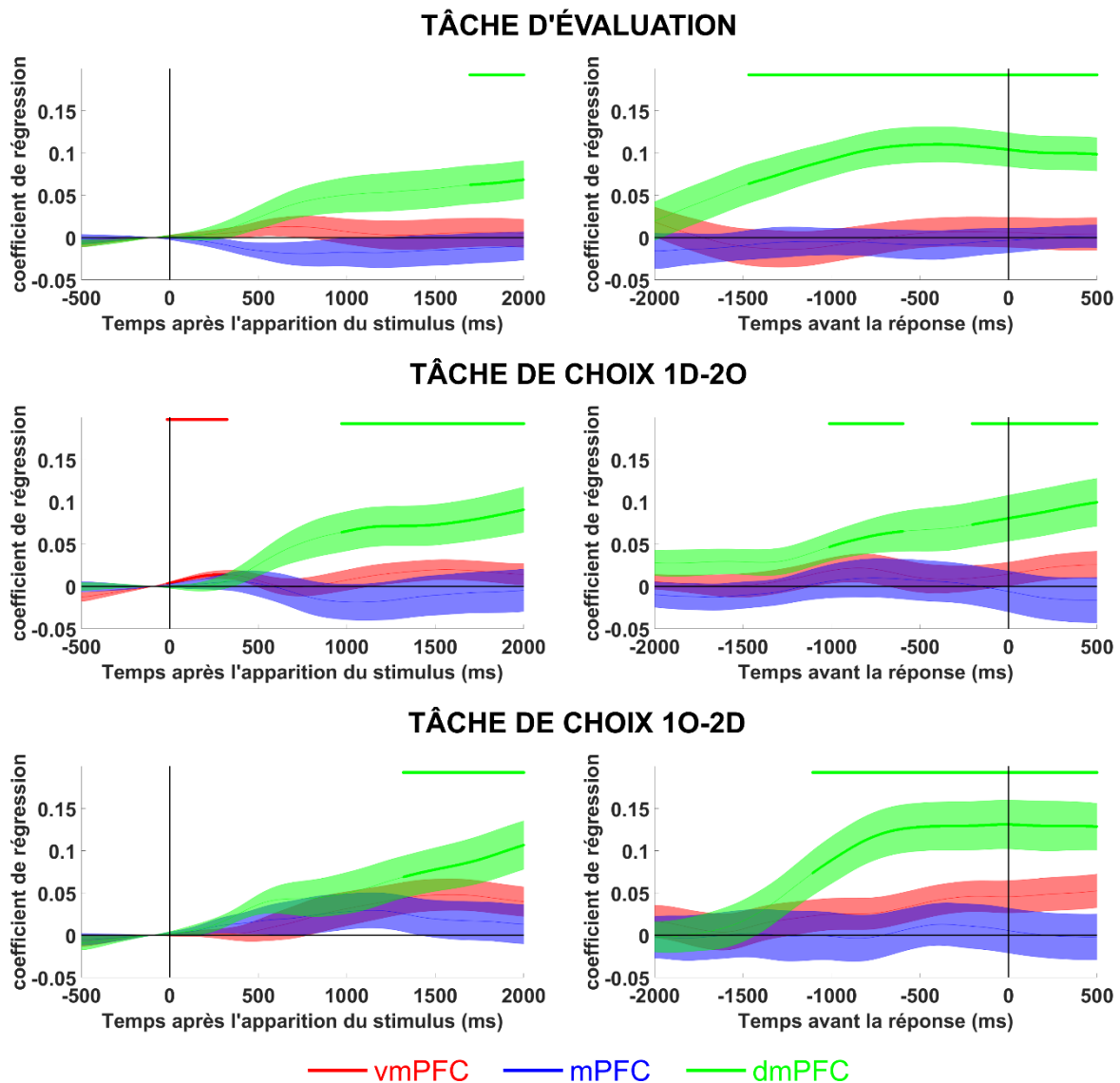
## Études expérimentales

Nous avons vu que l'encodage de la valeur subjective dans le vmPFC se généralise au travers de nos trois tâches. Nous avons ensuite voulu voir si cet encodage se généralisait aussi entre les dimensions et les modes de présentation des récompenses. Pour ce faire, nous nous sommes concentrés sur la tâche d'évaluation où il n'y avait qu'une seule option à évaluer à la fois et le lien avec l'activité neurale était ainsi plus simple à étudier. Il a été proposé que le vmPFC n'encode pas nécessairement les coûts associés à des actions et qu'une autre aire cérébrale correspondant au dmPFC identifié dans nos tâches en lien avec le DT corrèle avec cet encodage, voire corrèle négativement avec la valeur subjective en général, comme un système opposant au vmPFC. Nous avons d'abord pu observer que le dmPFC n'encodait pas positivement le coût des efforts, mais il avait en effet plutôt tendance à être d'autant moins actif que les efforts paraissaient pénibles ( $E_t: \beta_{Val} = -0.225 \pm 0.116, p = 0.060$ ). De plus, nous avons vu que le vmPFC encodait positivement la valeur des récompenses indépendamment de leur mode de présentation (voir **Figure 27**), c'est-à-dire qu'elles soient ou non accompagnées d'une image ( $R_{ti}: \beta_{Val} = 0.630 \pm 0.224, p = 0.008$ ;  $R_t: \beta_{Val} = 0.803 \pm 0.177, p = 6 \cdot 10^{-5}$ ). Enfin, le vmPFC encodait négativement la pénibilité des efforts ( $E_t: \beta_{Val} = -0.670 \pm 0.205, p = 0.002$ ) dans la tâche d'évaluation. L'association entre Val et l'activité du vmPFC était donc indépendante du mode de présentation des récompenses et elle intégrait aussi bien les bénéfices associés à des récompenses que les coûts associés à des efforts.

L'activité du dmPFC reflète donc mieux le coût dépensé, lors de la décision, qui est lié à la méta-décision « Combien d'effort et/ou de temps allouer à cette décision ? », plutôt que le coût associé à l'option en cours d'évaluation. Nous avons tiré cette conclusion du fait de la corrélation entre le dmPFC et le DT malgré le fait que le DT, dans notre GLM, était orthogonalisé par rapport à nos variables de non-intérêt ainsi que Val et Conf, c'est-à-dire que le DT représentait potentiellement l'effort alloué à la décision même après avoir pris en compte la difficulté de la décision reflétée dans Conf. Comme le DT est un proxy indirect de l'effort investi dans la tâche, nous avons aussi voulu voir si un autre proxy régulièrement associé à la notion d'effort mental permettait de vérifier ce lien entre l'effort investi dans la prise de décision et l'activité du dmPFC. Nous avons ainsi décidé d'utiliser la taille du diamètre pupillaire. L'activité au sein de chacune de nos régions d'intérêt dans le cortex préfrontal médian a ainsi été extraite pour chaque essai à l'aide du GLM2. Ensuite, nous avons effectué, pour chaque point temporel, une régression linéaire entre le diamètre pupillaire, nos variables de non-intérêt (la luminance, le nombre de mots et la durée d'affichage de la croix de fixation précédant le début de l'essai), nos variables d'intérêt (Val, Conf et DT) et, enfin, l'activité de nos régions

## Études expérimentales

d'intérêt (le vmPFC, le mPFC et le dmPFC). Toutes les variables ont été orthogonalisées de manière sérielle de sorte que l'activité du dmPFC était orthogonalisée par rapport à tous les régresseurs précédents. Une corrélation positive a été observée entre le diamètre pupillaire et le dmPFC dans chacune de nos tâches et de manière relativement spécifique par rapport au vmPFC ou au mPFC. Cette corrélation a été observée aussi bien en calant les données sur le moment d'apparition des stimuli à l'écran qu'en les calant sur le moment de réponse (voir **Figure 28**).



**Figure 28 : Corrélation entre le diamètre pupillaire et l'activité du vmPFC, du mPFC et du dmPFC dans chacune des tâches de prise de décision basée sur la valeur.** Les figures affichées montrent l'évolution temporelle des coefficients de corrélation issus d'un GLM entre le diamètre pupillaire et l'activité au sein du vmPFC (en rouge), du mPFC (en bleu) et du dmPFC (en vert) après avoir orthogonalisé chacun de ces régresseurs par rapport aux régresseurs précédents du GLM (la luminosité, le nombre de mots, la durée d'affichage de la croix de fixation, Val, Conf et DT). Dans toutes les lignes, la colonne de gauche représente les données calées sur le moment d'apparition des stimuli et la colonne de droite représente les

données calées sur le moment où les participants appuient sur une touche, c'est-à-dire lorsqu'ils ont pris leur décision. En haut, corrélations dans la tâche d'évaluation, au milieu dans la tâche de choix 1D-2O et en bas dans la tâche de choix 1O-2D. Les traits en gras au-dessus de la figure et au niveau de la moyenne représente les clusters temporels où les betas sont significatifs d'après la théorie des champs aléatoires qui permet de corriger pour les comparaisons multiples. Les données ont été extraites chez 30 participants dans la tâche d'évaluation, 29 participants dans la tâche de choix 1D-2O et 28 participants dans la tâche de choix 1O-2D.

### I.4 Discussion

Dans cette étude, nous avons cherché à identifier les bases cérébrales de variables communes à plusieurs tâches impliquant l'évaluation de différents stimuli et l'expression de préférences subjectives. Nous avons observé une triple dissociation dans le cortex préfrontal médian : la valeur subjective des stimuli, la confiance dans la réponse et le temps de délibération corrélaient, respectivement, avec le vmPFC, le mPFC et le dmPFC. Cette association entre chaque région et chaque variable était stable à travers nos tâches d'évaluation et de choix, qu'elles impliquent l'expression d'un jugement de valeur, d'une préférence entre deux options ou l'intégration de différents attributs, la récompense et l'effort. Ce résultat suggère que la valeur des récompenses et le coût des efforts sont intégrés dans la même région cérébrale, le vmPFC. Les variables méta-décisionnelles, qui régulent la prise de décision, comme la confiance dans la réponse et le temps de délibération, sont principalement représentés dans d'autres parties du cortex préfrontal médian : le mPFC et le dmPFC.

#### I.4.a Le vmPFC encode la valeur subjective de manière générique

Le réseau des valeurs, et en particulier le cortex préfrontal ventromédian (vmPFC) qui en est l'une des composantes principales a été associé à l'encodage de la valeur subjective de stimuli appétitifs, c'est-à-dire des récompenses, de manière générique (Bartra et al., 2013; Levy and Glimcher, 2012; Sescousse et al., 2013). Nous avons d'ailleurs confirmé ce résultat, chez le même groupe de participants, en IRMf, au travers de 3 tâches différentes : une tâche d'évaluation et deux tâches de choix, et ce indépendamment du mode de présentation des récompenses qui étaient affichées (voir **Figure 27**). Il a été proposé que le vmPFC n'est pas capable d'encoder la valeur subjective associée à des actions (Padoa-Schioppa, 2011; Rangel and Hare, 2010; Schneider and Koenigs, 2017). D'un autre côté, plusieurs études récentes suggèrent au contraire que le vmPFC est capable d'intégrer les coûts liés à l'obtention d'une récompense, comme les punitions ou le délai qui peuvent y être associés (Kable and Glimcher, 2007; Talmi et al., 2009; Tom et al., 2007), mais aussi le coût anticipé d'un effort qui a été choisi (Aridan et al., 2019; Hogan et al., 2019; Nagase et al., 2018). Nos résultats

## Études expérimentales

suggèrent aussi que le vmPFC encode la valeur subjective aussi bien dans le domaine des objets que dans le domaine des actions (voir **Figure 27**). Ils confirment ainsi l'idée que le vmPFC est capable de calculer la valeur nette associée à une action, en prenant à la fois en compte les bénéfices et les coûts dans son calcul de la valeur subjective. Notre paradigme nous a, de plus, permis de démontrer le fait que le vmPFC encode la valeur subjective de manière générique, c'est-à-dire indépendamment de la tâche en cours, de son mode de présentation ou du type d'item considéré, en faisant passer nos différentes tâches au même groupe de participants.

### **I.4.b Les variables méta-décisionnelles sont liées à des parties plus dorsales du cortex préfrontal médian**

Si le vmPFC était clairement corrélé à l'encodage du compromis coûts/bénéfices lié à l'évaluation, nos tâches nous ont aussi permis de dissocier cet encodage de celui de deux variables méta-décisionnelles essentielles à la prise de décision : la confiance dans la réponse et le temps de délibération.

La confiance est l'estimation subjective de la probabilité d'avoir accompli la meilleure action possible (Pouget et al., 2016). Étant donné que nous voulions tester trois tâches différentes (tâche d'évaluation, tâche de choix 1D-2O et tâche de choix 1O-2D) et plusieurs conditions (des récompenses présentées sous deux formes différentes et des efforts), mais que nous étions contraints, en termes de durée totale de l'expérience, par la durée du scan en IRMf, nous n'avons pas souhaité inclure des questions sur le degré de confiance des participants dans nos tâches. De ce fait, nous avons utilisé des substituts qui semblaient pertinents pour refléter le degré de confiance subjectif des participants. Plus précisément, dans la tâche d'évaluation nous avons utilisé les évaluations centrées et mises au carré, une mesure qui a déjà été validée plusieurs fois comme un bon reflet du degré de confiance subjectif des participants dans ce type de tâche (De Martino et al., 2017; Lebreton et al., 2015; Lopez-Persem et al., 2020). Dans les tâches de choix, nous avons utilisé la probabilité d'effectuer un choix donné dérivée de nos modèles qui est censée refléter directement le degré de confiance des participants et, par cohérence avec la tâche d'évaluation, nous avons aussi centré et mis cette mesure au carré. Dans nos trois tâches, nous avons pu voir que la mesure indirecte de la confiance que nous avons utilisée était bien corrélée avec des temps de délibération plus courts (voir **Figure 26**) suggérant que nos mesures étaient en effet valides pour représenter le degré de confiance approximatif des participants dans leur décision.

Conformément à des études antérieures, aussi bien en IRMf, qu'avec des enregistrements intracrâniens chez l'homme, et dans de multiples paradigmes, qui suggèrent que le vmPFC encode à la fois la valeur subjective et la confiance dans la décision (Chua et al., 2006; De Martino et al., 2013; Gherman and Philiastides, 2018; Lebreton et al., 2015) et que, de ce fait, il pourrait y avoir des biais entre ces deux mesures, nous avons pu observer une corrélation entre l'activité du vmPFC et la confiance (voir **Figure 26**). Cependant, l'épicentre de la corrélation entre la confiance subjective, telle que nous l'avons définie dans nos tâches, et l'activité cérébrale était situé légèrement plus dorsalement dans une zone que nous avons appelé le cortex préfrontal médian (mPFC) (voir **Figure 27**). Ce résultat est d'ailleurs en accord avec une étude de De Martino *et al* (De Martino et al., 2017), qui a montré que, dans une tâche d'évaluation, il existe deux gradients opposés dans le cortex préfrontal médian. La valeur subjective serait ainsi principalement encodée dans la partie ventrale du cortex préfrontal médian, alors que la confiance subjective serait principalement encodée dans une partie plus dorsale du cortex préfrontal médian (voir **Figure 22**). Nos résultats suggèrent en effet que la confiance dans la décision est principalement encodée dans des parties plus dorsales du cortex préfrontal médian.

Notre équipe a déjà proposé que la confiance est une variable que les participants cherchent à maximiser au cours d'une tâche (Lebreton et al., 2015). Or, lors d'une prise de décision, un moyen potentiel d'augmenter la confiance subjective dans la réponse est de prendre plus de temps pour délibérer avant de prendre une décision (Lee and Daunizeau, 2019). Ainsi, dans l'ensemble de nos tâches de prise de décision, nous avons considéré le temps de délibération (DT) comme une variable reflétant la quantité de ressources allouées à chaque décision et la sensation aversive d'effort qui va de pair. Bien sûr, ce lien dépend de la tâche considérée. Nous avons d'ailleurs que le DT était aussi bien influencé par la confiance dans la réponse, que par la valeur subjective des stimuli en cours d'évaluation. Ce lien pourrait refléter un réflexe pavlovien appétitif, comme cela a déjà été suggéré précédemment, en ce qu'un item plus appétitif a tendance à éliciter une réponse plus vigoureuse et plus rapide (Oudiette et al., 2019; Shadmehr et al., 2019). En effet, dans nos tâches, il n'y avait pas de raison de répondre plus rapidement aux stimuli plus appétitifs ou plus lentement face aux stimuli plus aversifs. Le DT était aussi généralement plus rapide lorsque la confiance était plus élevée ce qui pourrait refléter la difficulté de la décision (Kiani et al., 2014) et est cohérent avec l'idée que plus de ressources neurales sont attribuées lorsque la confiance dans la décision est faible en vue de l'augmenter. Dans nos analyses, nous avons cherché à voir les bases cérébrales du DT après

## Études expérimentales

l'avoir orthogonalisé aux différentes variables qui l'expliquaient, notamment la valeur subjective et la confiance.

Nous avons ainsi pu observer que, dans l'ensemble de nos tâches, le DT était corrélé à un cluster situé dans ce que nous avons appelé le cortex préfrontal dorsomédian (dmPFC), et cette relation était visible après avoir orthogonalisé le DT à la valeur et la confiance subjectives. Cette aire cérébrale a reçu plusieurs noms au fil des études. Elle est ainsi par exemple souvent dénommée de cortex cingulaire antérieur dorsal (dACC) (Shenhav et al., 2013), d'aire supplémentaire motrice (SMA) (Klein-Flugge et al., 2016; Kurniawan et al., 2013), d'aire pré-supplémentaire motrice (pre-SMA) (Kamiński et al., 2017; Kolling et al., 2016a) ou de cortex mid-cingulaire (Amiez et al., 2013). Au-delà de la question anecdotique de son appellation, cette aire cérébrale fait l'objet de débats intenses dans le domaine au sujet de son/ses rôle(s) précis. Elle est régulièrement identifiée en lien avec la sensation aversive d'exercer un effort (Engström et al., 2015; Kamiński et al., 2017; Pessiglione et al., 2018) et, plus spécifiquement, à la valeur attendue d'exercer du contrôle cognitif (Shenhav et al., 2013). Nous pensons que nos résultats qui montrent un lien entre le DT et l'activité du dmPFC (voir **Figure 27**) après avoir orthogonalisé le DT à différents régresseurs de non-intérêt, mais aussi à la valeur subjective et la confiance, vont dans ce sens. Le dmPFC serait ainsi le reflet des ressources investies au cours de la décision, en vue d'augmenter le niveau de confiance dans la décision (Lee and Daunizeau, 2019).

De plus, la taille du diamètre pupillaire, une autre variable régulièrement considérée comme un reflet des ressources investies dans une tâche, que ce soit dans le domaine des efforts physiques (Zénon et al., 2014) ou dans le domaine des efforts mentaux (Alnaes et al., 2014; Kahneman and Beatty, 1966; van der Wel and van Steenbergen, 2018), corrélait aussi avec l'activité du dmPFC avant de donner la réponse (voir **Figure 28**). Dans notre analyse, l'activité du dmPFC était orthogonalisée à la valeur subjective, la confiance mais aussi au DT, suggérant que l'activité de cette zone reflète en effet la détection d'un conflit (Botvinick et al., 2001) et l'allocation de ressources, sous la forme de l'activité du système nerveux autonome et/ou de l'exécution de contrôle cognitif (Beissner et al., 2013; Shenhav et al., 2013).

## **I.5 Conclusion**

En résumé, nous avons réussi à séparer les bases cérébrales de l'évaluation, de la confiance et de la délibération au sein du cortex préfrontal médian, alors que, comme montré par les cartes issues de notre méta-analyse sur Neurosynth, ces concepts n'étaient pas



## Études expérimentales

nécessairement clairement attribués à des bases cérébrales différentes dans le cortex préfrontal médian dans les études précédentes. L'une des distinctions les plus importantes de notre étude est probablement la dissociation entre l'effort considéré comme un attribut de l'option en cours d'évaluation, et l'effort, comme la quantité de ressources allouées à la prise de décision en soi. En d'autres termes, nous avons distingué d'un côté l'évaluation du coût d'un effort, associée au vmPFC, et l'effort investi dans l'évaluation, associé au dmPFC. Cette dissociation est cohérente avec l'idée que le vmPFC anticipe la valeur aversive d'un effort à effectuer, alors que le dmPFC représente l'intensité d'un effort au moment où il est exécuté et qui dépend de la difficulté de la tâche en cours (Hogan et al., 2019). Au niveau méta-décisionnel, nos résultats pourraient être interprétés dans le cadre d'un modèle d'allocation de ressources, où l'effort et/ou le temps investi dans la délibération sont censés augmenter le niveau de confiance dans la réponse, qu'il s'agisse d'une évaluation ou d'un choix (Lee and Daunizeau, 2019). Ce modèle prédit que l'augmentation de l'activité du dmPFC devrait avoir pour but d'augmenter le niveau d'activité du mPFC à la fin de l'essai. Nous ne pouvions pas tester cette prédiction ici puisque nous approchons à base de corrélation empêche d'effectuer un lien de causalité. Il est d'ailleurs aussi possible que l'activité du dmPFC ne reflète pas la quantité de contrôle cognitif à investir, mais la quantité de contrôle cognitif en train d'être investie. Elle pourrait aussi encoder la sensation aversive d'effort mental qui accompagne l'exécution de contrôle cognitif ou l'augmentation de l'activité du système nerveux autonome. Notre approche nous a en tout cas permis de montrer une association entre trois concepts cognitifs liés au compromis coûts/bénéfices qui est stable à travers nos tâches d'évaluation et de choix et qui implique d'étendre nos modèles du fonctionnement du cerveau lors d'un compromis coûts/bénéfices au-delà des modèles très étudiés qui ont été établis pour des tâches de choix binaires. Notre approche visant à observer les composantes stables à travers les tâches a cependant aussi des limites puisqu'elle ne nous permettait pas de comparer les tâches entre elles. En effet, les tâches différaient par plusieurs aspects : les tâches de choix venaient toujours après les tâches d'évaluation, chaque tâche n'impliquait pas le même nombre d'attributs et d'options à évaluer ou à comparer, etc. Cependant, nous espérons ouvrir une nouvelle approche avec la présente étude qui permet d'observer les bases cérébrales d'un même concept au travers de plusieurs tâches avec les mêmes participants.

Nous avons aussi profité de cette étude pour :

## Études expérimentales

- 1) tester la qualité de différentes séquences d'IRMf dans leur affinité pour capter l'activité du cortex préfrontal ventromédian en lien avec la valeur ;
- 2) voir si le fait d'utiliser un modèle qui a été développé dans l'équipe dont le principe est d'inférer des valeurs cachées qui dirigent le comportement dans l'ensemble des tâches (évaluation, choix 1D-2O et choix 1O-2D) changeait nos résultats basés sur les évaluations fournies par les participants dans la tâche d'évaluation ;
- 3) voir si la position du regard et la valeur associée aux différentes options était ou non en lien avec la théorie du modèle de diffusion attentionnel.

On pourra retrouver l'ensemble de ces questions, ainsi que l'article en anglais, que nous sommes sur le point de soumettre et qui correspond aux données présentées ici, dans les appendices à la fin de ce manuscrit (voir *Appendices*).

## **II L'effort, tant physique que mental, est plus motivé par l'attrait de gagner que par la peur de perdre**

### **II.1 Introduction**

Dans la première partie, nous nous sommes intéressés aux bases cérébrales du compromis coûts/bénéfices dans le cas de situations fictives impliquant des récompenses de diverses natures, comme des items alimentaires ou des objets, et des efforts fictifs de nature physique ou mentale. Nous avons ainsi pu voir que, non seulement l'activité du vmPFC, l'une des composantes principales du réseau des valeurs, augmentait linéairement avec la valeur appetitive des récompenses, mais elle diminuait aussi linéairement avec la valeur aversive du coût des efforts, et ce, au travers de différentes tâches effectuées par les mêmes participants (voir **Figure 27**). Dans cette partie, nous allons nous intéresser à l'autre composante principale du réseau des valeurs : le striatum ventral. L'activité du striatum ventral a en effet été corrélée à l'anticipation d'une récompense monétaire de valeur plus élevée depuis les débuts de l'IRMf chez l'homme (Breiter et al., 1996, 2001; Knutson et al., 2001a, 2001b). Ce lien entre l'activité du striatum ventral et l'anticipation d'une récompense de valeur subjective plus élevée a depuis été confirmé par plusieurs méta-analyses (Bartra et al., 2013; Clithero and Rangel, 2014; Sescousse et al., 2013).

De plus, les ganglions de la base, un réseau cérébral dont le striatum fait partie, jouent un rôle fondamental dans l'intensité et la rapidité d'exécution d'un mouvement (Dudman and Krakauer, 2016; Turner and Desmurget, 2010; Turner et al., 2003; Yttri and Dudman, 2016).

## Études expérimentales

Dans le cas d'un effort physique, comme d'un effort mental, des récompenses plus élevées motivent aussi une performance plus élevée que ce soit en termes d'intensité d'un effort physique ou de rapidité d'exécution d'un effort mental. En IRMf, cela correspond aussi à des situations où l'activité du striatum est plus élevée (Pessiglione et al., 2007; Schmidt et al., 2012). Il a d'ailleurs été prouvé qu'une lésion des ganglions de la base chez l'homme faisait perdre toute capacité de moduler l'intensité d'un effort physique en fonction de la valeur des enjeux monétaires (Schmidt et al., 2008). En résumé, non seulement le striatum réagit proportionnellement à la valeur subjective de la récompense en jeu, mais son activité pourrait déterminer l'intensité de l'effort physique, ou mental, effectué en vue d'obtenir cette récompense.

Qu'en est-il dans le domaine des pertes ? D'un côté, l'activité du réseau des valeurs diminue avec la perspective de punitions plus importantes, comme la perspective d'une perte monétaire conséquente (Park et al., 2011; Shapiro and Grafton, 2020; Talmi et al., 2009; Tom et al., 2007). D'un autre côté, l'activité du striatum ventral a été associée à la fois à l'encodage de la valeur subjective signée et de la valeur subjective non-signée (Bartra et al., 2013). De plus, dans des tâches de choix probabilistes, il a été montré qu'il existe une aversion à la perte (Kahneman and Tversky, 1984). Cette aversion s'exprime notamment par le fait que l'équivalent de la perspective de subir une perte d'un certain montant serait de 1.5 à 2.5 fois plus important pour une perspective d'obtenir un gain (Kahneman, 2012). Daniel Kahneman propose que ce type de biais soit ancré dans l'évolution. Les animaux auraient des chances de survie accrues en traitant les menaces, c'est-à-dire les punitions potentielles, de manière plus urgente que les opportunités en termes de récompenses éventuelles. Ce type de biais a cependant été très peu exploré dans des tâches impliquant d'effectuer un effort physique ou mental en fonction des enjeux chez l'homme.

Le dmPFC a été identifié comme une aire cérébrale faisant partie d'un réseau neural d'alarme (Eisenberger and Lieberman, 2004). Son activité augmente en effet dans des situations associées à un signal de saillance plus important (Litt et al., 2011; Seeley et al., 2007). L'activité du dmPFC a été associée au calcul de la valeur attendue d'allouer du contrôle cognitif à une tâche en cours (Shenhav et al., 2013). Il pourrait ainsi détecter des situations saillantes et « décider » d'allouer une certaine quantité de contrôle cognitif à la tâche en cours en vue de gérer la situation au mieux. Plus les enjeux monétaires non-signés d'une tâche seraient importants, plus on pourrait s'attendre à ce que son activité augmente en vue d'allouer plus de contrôle cognitif. D'un autre côté, l'activité du dmPFC augmente à mesure que la valeur

## Études expérimentales

subjective signée diminue (Bartra et al., 2013). Il se pourrait ainsi qu'il soit plus sensible dans l'anticipation de pertes que dans l'anticipation de gains. L'interaction de ce système de contrôle cognitif avec une éventuelle aversion à la perte n'est d'ailleurs pas clairement définie. Est-ce que des situations impliquant des pertes impliquent l'allocation d'une quantité plus importante de contrôle cognitif ?

Ainsi, nous avons adapté le paradigme de Schmidt *et al.* qui impliquait de demander à des participants d'effectuer un effort physique ou mental en vue d'obtenir des gains monétaires proportionnellement à la performance effectuée (Schmidt et al., 2012), en introduisant des pertes à éviter, en plus des gains à obtenir. Cette étude visait à :

- 1) Voir s'il existait une aversion à la perte dans des tâches impliquant d'effectuer une performance physique ou mentale et non pas seulement de faire un choix ;
- 2) Etudier le comportement du striatum et des aires associées au contrôle cognitif en lien avec les enjeux signés ou non-signés et vérifier comment ces différentes aires cérébrales réagissent en rapport avec une éventuelle aversion à la perte.

## II.2 Méthodes

### II.2.a Participants

Pour cette étude en IRMf, nous avons recruté 24 participants (13 femmes / 11 hommes) de  $25.9 \pm 3.7$  ans (moyenne  $\pm$  écart-type). Les participants ont été recrutés en ligne via la plateforme du RISC (Relais d'Information en Sciences de la Cognition, <https://www.risc.cnrs.fr/>). Les critères d'inclusion étaient d'être droitier, de parler couramment le français, d'avoir entre 20 et 39 ans, de ne pas porter de tatouage, de piercing ou tout objet métallique, de ne pas être enceinte (vérifié avec un test de grossesse effectué avant l'expérience) et de ne pas avoir de problèmes de vue. Nous avons aussi évalué les antécédents médicaux, en particulier neurologiques ou psychiatriques, et la consommation éventuelle de substances psychotropes des participants afin de les exclure le cas échéant. Enfin, les participants devaient aussi indiquer leurs traitements en cours afin de les exclure au cas où leurs traitements pourraient avoir un impact sur leur comportement et/ou leur activité neurale.

Nous avons exclu un participant qui a eu des performances déplorables dans la tâche d'effort physique. Il a en effet obtenu un décompte final négatif à la fin de chaque session de la tâche d'effort physique.

## Études expérimentales

Pour les données d'IRMf, nous avons exclu 3 participants car ils bougeaient trop (mouvement supérieur à 5 mm). Ces participants ont cependant été inclus dans les études de comportement. Les données d'IRMf sont donc basées sur 20 participants au total (11 femmes / 9 hommes).

Tous les participants ont signé une fiche de consentement. Les instructions indiquaient aux participants qu'ils obtiendraient un montant fixe de 50€ pour avoir participé à l'étude et qu'ils pourraient obtenir un bonus monétaire allant de 0 à 25€ en fonction de leurs performances cumulées au cours des différentes sessions de chacune des tâches effectuées dans le scanner. En pratique, tous les participants ont été payés un montant fixe de 75€ quelques mois après l'expérience.

### II.2.b Tâches

- **Points communs aux deux tâches**

Les deux tâches ont été programmées dans la toolbox Psychtoolbox (<http://psychtoolbox.org/>) au sein du logiciel Matlab 2012 (The MathWorks, Inc., USA). Les participants ont reçu des instructions pour chaque tâche à l'écrit et à l'oral. Le principe des deux tâches tel qu'il était indiqué aux participants était de maximiser leurs gains. Les deux tâches ont été programmées afin de minimiser autant que possible les différences entre elles qui n'étaient pas liées à la nature de l'effort (effort physique ou effort mental) à effectuer dans chaque tâche. Les deux tâches effectuées par les participants correspondent à des variantes de tâches qui ont largement été étudiées par le passé dans notre équipe (Meyniel et al., 2013; Pessiglione et al., 2007; Schmidt et al., 2008, 2009, 2012).

Dans les deux cas, un essai démarrait par la présentation d'une croix de fixation pour une durée fixe de 500 millisecondes (ms). Ensuite, le montant en jeu pour l'essai en cours apparaissait sur l'écran (voir **Figure 29**) pour une durée variant aléatoirement entre 1 et 4 secondes (s). Cet enjeu était présenté à l'écran sous la forme d'une pièce de monnaie ou d'un billet en haut à gauche de l'écran. Si une croix rouge se superposait à l'image du montant en jeu accompagnée du texte « A PERDRE », le participant savait qu'il jouait pour ne pas perdre de l'argent à cet essai. Plus sa performance serait élevée, moins il allait perdre d'argent proportionnellement au montant en jeu. Dans le cas contraire, dans les essais sans croix rouge et avec le texte « A GAGNER », le participant savait, d'après les instructions, que plus sa performance serait élevée, plus il obtiendrait une proportion importante du montant en jeu à cet essai. Ensuite, les participants voyaient une échelle graduée au milieu de l'écran qui leur

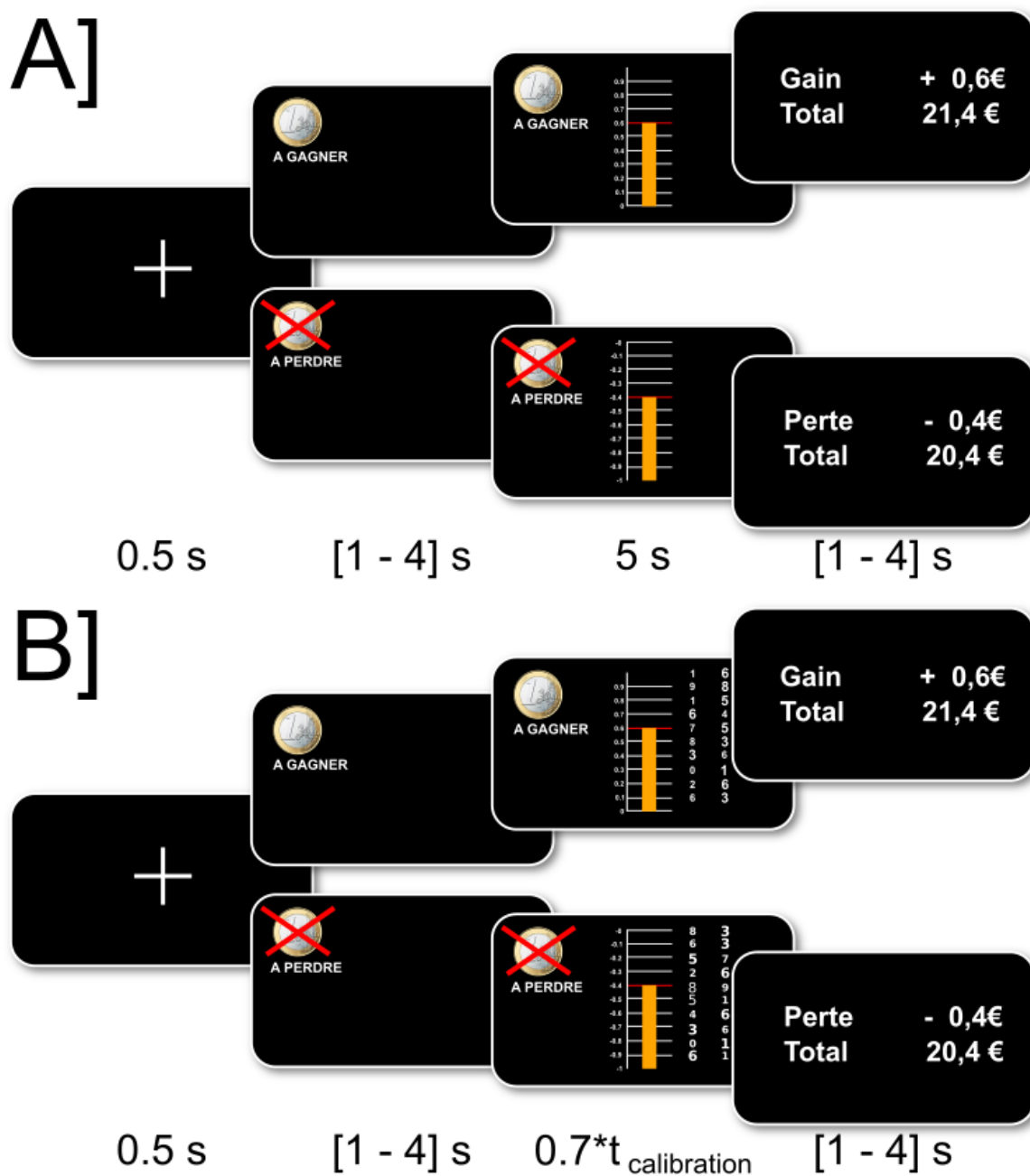
## Études expérimentales

indiquait leur niveau de performance maximale atteint pour l'essai en cours. Ce niveau de performance était indiqué par une barre orange surplombée d'un trait rouge. Le niveau de performance affiché était continuellement actualisé au cours de l'essai en fonction de la performance du participant. Le haut de l'échelle correspondait à obtenir 100% du gain présenté, ou à ne rien perdre dans le cas des pertes. Le bas de l'échelle correspondait à ne rien obtenir dans le cas des gains ou à perdre l'intégralité du montant présenté dans le cas des pertes. Chaque barre de l'échelle correspondait à un pas de 10% sur l'échelle. Le montant correspondant à chaque barre de l'échelle était, de plus, indiqué à gauche de l'échelle. Une fois que l'échelle apparaissait à l'écran, les participants pouvaient commencer à effectuer la tâche pour l'essai en cours. Cette période était toujours d'une durée fixe au cours d'une session. Enfin, les participants recevaient un feedback. Ce feedback consistait en deux nombres (voir **Figure 29**). Le premier nombre indiquait aux participants le montant obtenu, ou perdu le cas échéant, lors de l'essai en cours. Le deuxième nombre informait les participants du montant total accumulé au cours des essais et des sessions de la tâche en cours après l'avoir actualisé en fonction de l'essai en cours. Ce montant était indépendant pour chaque tâche. Lors de la deuxième session d'une des deux tâches, les participants démarraient ainsi avec le cumul des gains obtenus lors de la première session de la tâche en question. Le feedback était présenté pendant une durée variant aléatoirement entre 1 et 4 secondes.

Les instructions informaient les participants qu'ils auraient à effectuer 2 sessions de chacune des tâches, mais pas du nombre d'essais correspondant à chaque session. Chaque session contenait 60 essais pour chacune des tâches. Sans que les participants en soient informés, les essais, pour une session donnée, étaient organisés en 5 mini-blocs. L'organisation des essais en mini-blocs n'était en effet pas spécifiée dans les instructions et aucune séparation concrète ne permettait de voir qu'un mini-bloc était achevé. Chaque mini-bloc était constitué de 12 essais correspondants aux 12 montants possibles dans la tâche : perdre 20.00€, 5.00€, 1.00€, 0.50€, 0.20€ ou 0.01€ ou gagner 0.01€, 0.20€, 0.50€, 1.00€, 5.00€ ou 20.00€. Ces 12 niveaux d'enjeux étaient présentés de manière aléatoire au sein de chaque mini-bloc. L'ordre de passation des tâches était fixé pour un sujet donné. Autrement dit, si la tâche d'effort physique précédait la tâche d'effort mental pour la première session, il en était de même lors de la seconde session. L'ordre des tâches était cependant contrebalancé à travers les participants. Nous avons dit aux participants que leur rémunération finale consistait en un paiement fixe de 50€ pour avoir participé à l'expérience et à un bonus monétaire allant de 0€ à 25€. Il était

## Études expérimentales

indiqué que ce bonus monétaire dépendrait de leurs performances au sein de chacune des 3 tâches.



**Figure 29 : Tâche d'effort physique (A) et d'effort mental (B).** Un essai démarrait toujours par 0.5 secondes (s) d'affichage d'une croix de fixation au milieu de l'écran. Ensuite, le montant en jeu pour l'essai en cours apparaissait à l'écran sous la forme d'une pièce de monnaie (0.01€, 0.20€, 0.50€ ou 1€) ou d'un billet (5€ ou 20€). Le texte en dessous leur indiquait s'ils jouaient pour gagner le montant en jeu (« A GAGNER ») ou pour ne pas le perdre (« A PERDRE »). Pour souligner la différence entre ces deux conditions, nous avons aussi rajouté une croix rouge superposée au montant en jeu. Cette période d'affichage des enjeux avait une durée aléatoire allant de 1 à 4 secondes. Ensuite, une échelle apparaissait au milieu de l'écran leur indiquant qu'ils pouvaient commencer à effectuer la tâche. Une barre orange ainsi qu'un trait rouge leur indiquait le niveau de performance atteint jusqu'à maintenant au cours de l'essai. A] Pour la

## Études expérimentales

tâche d'effort physique, la performance dépendait de l'intensité maximale de force effectuée au cours d'une période fixe de 5s. B] Pour la tâche d'effort mental, la performance dépendait du nombre de paires de Stroop numérique correctement résolues au cours du temps imparti. La barre augmentait d'un cran pour chaque paire résolue. D'un essai à l'autre, les paires présentées étaient toujours les mêmes. Cependant, l'ordre de présentation des paires sur l'échelle variait d'un essai sur l'autre. La moitié des paires était congruente. Dans ce cas, le chiffre le plus grand numériquement était aussi le plus grand en taille d'affichage. L'autre moitié des paires était incongruente. Dans ce cas, le chiffre le plus grand numériquement avait une taille plus petite sur l'écran que l'autre chiffre de la même paire. Les paires variaient aussi par la distance entre les chiffres au sein d'une même paire. L'écart allait de 1 à 5 entre 2 chiffres d'une même paire. Le temps de cette période pour une session donnée dépendait du temps minimum effectué lors de la calibration qui précédait chaque session. Enfin, pour les deux tâches, un feedback permettait au sujet de savoir quelle proportion du montant en jeu il avait obtenu, pour les gains, ou perdu, pour les pertes. Au cours de cette période, le montant accumulé pour chacune des tâches était aussi mis à jour et présenté à l'écran pour permettre aux participants de suivre leur performance. Ce feedback était affiché pour une durée aléatoire allant de 1 à 4 secondes.

- **Spécificités de la tâche d'effort physique**

Pour mesurer la force effectuée par les participants dans la tâche d'effort physique, nous avons utilisé une poignée de force IRM-compatible construite au laboratoire. Cette poignée a déjà été utilisée lors d'études précédentes (Meyniel et al., 2013; Schmidt et al., 2009). Elle était composée de 2 cylindres en plastique au sein desquels se trouvait un tube en plastique. Serrer la poignée de force et, ainsi les 2 cylindres qu'elle contenait, revenait à compresser le tube d'air qui y était inséré. Le tube en question était relié à une salle hors de l'IRM où la pression de l'air était convertie en voltage à l'aide d'un transducteur. Toute compression de la poignée de force avait ainsi pour conséquence la production d'un signal linéaire, sous forme de voltage. Ce signal était proportionnel à la force exercée sur la poignée de force. Ce signal était ensuite envoyé à l'ordinateur de mesure via un conditionneur de signal (CED 1401; Cambridge Electronic Design). Ce signal était enfin lu par Matlab (The MathWorks, Inc., USA) pour les besoins de la tâche.

Contrairement aux poignées de force qui permettent de mesurer la force exercée directement en Newtons, notre poignée de force ne permettait d'obtenir qu'une mesure relative. Ces poignées de force fonctionnent en effet avec des matériaux incompatibles avec l'IRM. Une calibration était donc nécessaire afin que le gain utilisé soit le mieux adapté aux capacités de chaque individu. Pour ce faire, avant chaque session de la tâche d'effort physique, nous demandions aux participants de serrer la poignée de force avec divers degrés de force afin d'adapter le gain du transducteur à leurs capacités. Une fois le gain réglé, nous leur demandions d'essayer de serrer la poignée de force avec leur force maximale sur une durée de 5 secondes. Cette durée correspondait à la durée pendant laquelle ils auraient à effectuer un effort au cours



## Études expérimentales

d'un essai de la tâche. Cela nous permettait d'obtenir la valeur de la contraction maximale volontaire (*maximum voluntary contraction force*, MVC) pour un sujet donné. Cette calibration était opérée au cours de 3 essais successifs non-rémunérés avec leur main droite. Pour motiver les participants à serrer de leur mieux, bien que ces essais ne soient pas rémunérés, nous affichions les valeurs en volts correspondant à la force exercée à chaque instant en vue de les pousser à atteindre le score le plus élevé possible. De plus, à la fin de chaque essai, un message indiquant le record de force effectuée à travers les 3 essais était affiché, toujours en volts, afin de les pousser à battre leur record d'un essai sur l'autre et, ainsi, à effectuer leur vraie force maximale. Le maximum atteint au cours des 3 essais de calibration était ensuite utilisé comme MVC pour la session. Cependant, dans certains cas, les participants parvenaient néanmoins à effectuer une force plus importante au cours de la tâche, ce qui indiquait qu'ils n'avaient pas effectué leur vraie force maximale au cours de la calibration. Lorsque cela se produisait, nous remplaçons la MVC initiale par cette nouvelle valeur pour les essais suivants. La MVC, pour un essai  $i$  donné ( $i > 1$ ), était calculée ainsi :

$MVC(i) = \max(MVC(i-1), PF(i-1) - bF(i-1))$  où  $PF(i)$  est le pic de force maximal exercé lors de l'essai précédent et  $bF(i-1)$  est la moyenne de la force en volts au cours des 500 millisecondes de l'affichage de la croix de fixation à l'écran, juste avant le début de l'essai précédent, c'est-à-dire la baseline de force (bF). Les participants n'étaient pas informés de cette mise à jour de la MVC dans les instructions.

L'échelle était calibrée de sorte à ce que, pour un essai  $i$  donné, effectuer une force égale à  $MVC(i)$  revenait à atteindre 75% du haut de l'échelle. Le but de cette procédure était de pousser les participants à prendre en compte les enjeux au moment de faire leur effort plutôt que d'essayer d'atteindre le haut de l'échelle à chaque essai. Si la tâche avait été trop facile, nous avons peur en effet qu'ils ne proportionnent pas l'effort fourni en fonction des enjeux. Les participants n'étaient cependant pas informés de la manière dont la correspondance entre la force effectuée et l'échelle observée était effectuée. Le haut de l'échelle,  $F_{max_{top}}$ , était ainsi défini selon la formule :

$$F_{max_{top}}(i) = \frac{MVC(i)}{\frac{75}{100}}$$

Seul le pic de force, c'est-à-dire l'intensité maximale au cours d'un essai, et non sa durée était pris en compte pour le calcul de la performance. Les participants en étaient informés dans les instructions et pouvaient de toute façon l'observer visuellement comme la barre orangée se

## Études expérimentales

mettait à jour en direct pendant qu'ils effectuaient la tâche (voir **Figure 29**). La hauteur de la barre orangée à l'écran à un instant  $t$  de l'essai, et donc la performance au cours d'un essai, variait selon la formule :

$$P(i, t) = \begin{cases} \frac{PF(i, t) - bF(i)}{Fmax_{top}(i)}, & \text{si } PF(i, t) - bF(i) \leq Fmax_{top}(i) \\ 1, & \text{si } PF(i, t) - bF(i) > Fmax_{top}(i) \end{cases} \quad \text{où } P(i, t) \text{ est la performance au}$$

moment  $t$  au cours de l'essai  $i$ ,  $PF(i, t)$  le pic de force maximal exercé jusqu'au moment  $t$  de l'essai  $i$ ,  $bF(i)$  la moyenne de la force en volts au cours des 500 millisecondes pendant l'affichage de la croix de fixation à l'écran juste avant le début de l'essai, c'est-à-dire la baseline de force ( $bF$ ), et  $Fmax_{top}(i)$  correspond à la force à effectuer pour atteindre le haut de l'échelle. Nous avons pris en compte la baseline de force dans nos mesures de force au cas où il y ait une déviation du signal traduisant la force exercée en voltage au cours du temps.

Seuls les efforts effectués pendant que l'échelle apparaissait à l'écran étaient pris en compte dans la mesure de la force effectuée. Si les participants serraient la poignée de force avant ou après que l'échelle apparaisse l'effort effectué n'était pas pris en compte dans la mesure de la performance. Ils avaient 5 secondes pour serrer la poignée lorsque l'échelle apparaissait à l'écran.

Avant d'effectuer la tâche dans le scanner, les participants ont pu se familiariser avec la tâche au cours de 12 essais (1 par niveau d'enjeu) à l'aide d'un ordinateur portable et d'un appareil similaire à celui utilisé dans le scanner pour mesurer la force exercée.

Dans le scanner, les participants ont effectué 2 sessions contenant 60 essais de cette tâche. Avant la première session de cette tâche, les participants effectuaient systématiquement une session d'une tâche d'apprentissage (voir *L'apprentissage par renforcement implique de l'effort mental*) et parfois aussi une session de la tâche d'effort mental selon l'ordre de passation des tâches pour chaque sujet. Entre les 2 sessions de la tâche d'effort physique, les participants effectuaient systématiquement une session de la tâche d'apprentissage et une session de la tâche d'effort mental. Ces deux tâches requérant l'utilisation d'un boîtier de réponse avec la même main que celle utilisée pour la tâche d'effort physique, l'expérimentateur rentrait dans la salle à chaque fois qu'une transition était nécessaire afin de remplacer le boîtier de réponse par la poignée de force et réciproquement. Cela nous permettait d'éviter que le sujet bouge entre 2 sessions pour faire l'échange entre le boîtier de réponse et la poignée de force. Nous voulions nous assurer que les participants utilisent leur main droite pour effectuer toutes les tâches

puisque'ils étaient droitiers et aussi afin d'éviter toute différence de latéralité entre les deux tâches.

- **Spécificités de la tâche d'effort mental**

L'échelle d'effort dans la tâche d'effort mental a été définie pour ressembler autant que possible à celle utilisée dans la tâche d'effort physique. Chaque graduation en traits blancs de l'échelle correspondait à 10% du montant en jeu et était associée à une paire de chiffres. La hauteur de la barre orangée sur l'échelle indiquait la performance atteinte, en direct, au cours de l'essai et correspondait au nombre de paires résolues. Dès qu'une paire était correctement résolue, la barre montait ainsi d'un cran sur l'échelle. Résoudre une paire de Stroop numérique correspondait à choisir le chiffre le plus haut numériquement au sein d'une paire. Les chiffres variaient cependant en termes de valeur numérique, entre 0 et 9, et en termes de taille de police Arial. Un des deux chiffres d'une paire apparaissait ainsi en taille 24 et l'autre en taille 40. Au sein d'une paire donnée, un des deux chiffres était toujours supérieur à l'autre numériquement. Il y avait donc toujours une réponse correcte et une réponse incorrecte pour chaque paire de chiffres (voir **Figure 29**). Pour la moitié des paires, le chiffre le plus grand numériquement était aussi le chiffre qui apparaissait avec la taille la plus grande (paires congruentes). Pour l'autre moitié des paires, le chiffre le plus grand numériquement au sein d'une paire était celui qui était affiché avec la taille la plus petite (paires incongruentes). La différence entre les 2 chiffres d'une paire variait de 1 à 5 selon les paires.

Chaque essai contenait 10 paires de Stroop numérique à résoudre. Ces paires étaient toujours les mêmes. Cependant l'ordre d'affichage des paires du bas vers le haut de l'échelle et la position, au sein de chaque paire, de chaque chiffre entre la gauche et la droite variaient aléatoirement d'un essai à l'autre. Les participants ne pouvaient ainsi pas anticiper l'ordre des paires qu'ils allaient rencontrer ni la localisation de la bonne réponse pour chaque paire avant que l'échelle d'effort ainsi que les paires associées apparaissent à l'écran.

Avant chaque session, les participants devaient effectuer 3 essais de calibration. Un essai de calibration consistait à résoudre 10 paires de Stroop numérique le plus vite possible sans rémunération monétaire. Cependant, pour les motiver, nous avons affiché leur record temporel au cours des essais de calibration précédents. Là non plus, nous ne leur disions pas que ces essais seraient déterminants pour la suite de la tâche. Le temps minimum pour résoudre ces 10 paires était ensuite utilisé comme temps de référence pour l'ensemble de la session (*t<sub>calibration</sub>*). La durée de la période pendant laquelle les participants pouvaient effectuer un effort mental

## Études expérimentales

était ainsi définie selon la formule :  $T_{essai} = \frac{70}{100} * t_{calibration}$  secondes. Ce temps était fixe pour une session donnée d'un participant donné. Il pouvait cependant varier d'une session à l'autre d'un même sujet et aussi d'un sujet à l'autre en fonction de la calibration.

Au sein d'une paire, lorsque le participant sélectionnait la mauvaise réponse, les deux chiffres de la paire concernée étaient affichés en rouge pendant une durée de  $0.1 * T_{essai}$ . Pendant cette période, toute réponse du sujet était ignorée. Toute mauvaise réponse était ainsi pénalisée par un malus temporel. Le but de ce malus temporel était de décourager les participants d'adopter une stratégie où ils répondraient le plus vite possible mais aléatoirement pour chaque paire en vue de maximiser leurs gains. Nous voulions ainsi les forcer à maximiser aussi bien leur vitesse que leur précision pour chaque réponse.

Comme pour la tâche d'effort physique, avant d'effectuer la tâche dans le scanner, les participants ont pu se familiariser avec la tâche au cours de 12 essais à l'aide d'un ordinateur portable. Chacun de ces 12 essais était associé à l'un des 12 niveaux d'enjeux auxquels ils seraient confrontés dans les différents essais de la tâche. Pour répondre sur l'ordinateur portable, ils pouvaient utiliser les flèches gauche/droite du clavier qui servaient d'équivalent aux boutons gauche/droite du boîtier de réponse.

### II.2.c Analyse comportementale

Toutes nos analyses ont été effectuées avec le logiciel Matlab 2017a (The MathWorks, Inc., USA).

- **Tests de l'efficacité de la tâche de Stroop numérique**

Nous avons d'abord voulu voir si les paires incongruentes induisaient bien plus d'effort cognitif que les paires congruentes dans la tâche de Stroop numérique. Pour cela, nous avons d'abord regardé le nombre moyen de paires de chaque type correctement résolues par les participants à travers les sessions. Nous avons comparé si ce nombre différait entre les deux types de paires à l'aide d'un t.test apparié.

Ensuite, nous avons regardé le nombre moyen d'erreurs commises pour chaque essai pour chaque type de paire pour chaque participant. Une erreur était comptabilisée lorsque le participant n'avait pas donné la réponse correcte du premier coup pour une paire donnée. Si le participant avait donné plusieurs fois la même réponse incorrecte pour une paire donnée, nous ne comptabilisions cela que comme une seule erreur. Nous avons ensuite effectué un t.test apparié pour voir si le nombre de paires résolues après avoir commis une erreur était plus élevé

## Études expérimentales

dans le cas des paires incongruentes que dans le cas des paires congruentes. Nous avons ensuite fait le même test en remplaçant le nombre de paires résolues avec au moins une erreur par les temps de réaction. Le temps de réaction, pour une paire donnée autre que la première d'un essai, était calculé comme la différence entre le moment de la première réponse pour cette paire et le moment de la dernière réponse pour la paire précédente. Pour la première paire, il s'agissait simplement de la différence entre le moment de la première réponse pour cette paire et le moment où les paires apparaissaient à l'écran. Au sein de chaque type de paire, nous avons aussi cherché à voir s'il existait une différence dans les temps de réaction entre les essais où une réponse incorrecte a été donnée par rapport aux essais résolus sans aucune erreur.

Enfin, nous avons aussi regardé si la distance entre les chiffres au sein d'une paire donnée avait un impact sur le fait d'effectuer ou non une erreur pour chaque paire et sur le temps de réaction. Nous avons pour cela testé la corrélation linéaire entre le nombre moyen d'erreurs

### • **Modèle linéaire de la performance**

Nous avons effectué un premier modèle linéaire pour expliquer la performance, en termes de pourcentage du haut de la barre atteint, dans les deux tâches. Pour cela, nous avons utilisé un modèle bayésien à l'aide la VBA toolbox (Daunizeau et al., 2014). Le modèle a été inversé en utilisant l'approche variationnelle bayésienne avec une approximation de Laplace. Nous avons ainsi modélisé la performance selon l'équation :

$$Perf = \beta_0 + \beta_{|E|} * |E| + \beta_E * E + \beta_{NE} * NE$$

(Équation 3)

$\beta_0$  est une constante permettant d'exprimer la performance basale des participants à tous les essais,  $\beta_{|E|}$  est une constante exprimant la sensibilité aux enjeux,  $|E|$  correspond aux enjeux monétaires exprimés en euros en valeur absolue,  $\beta_E$  est une constante permettant d'exprimer la sensibilité aux montants en jeu en fonction de leur valence,  $E$  représente les enjeux monétaires exprimés en euros signés,  $\beta_{NE}$  est une constante exprimant une sensibilité au temps passé à faire la tâche et  $NE$  est le numéro d'essai. Nous avons mis un *a priori* avec une moyenne à zéro et un écart-type de 1 pour chacun de ces paramètres. Les paramètres étaient libres. Nous avons regroupé l'ensemble des données de chaque participant à travers les deux sessions à l'aide de l'option « multisession » de la VBA toolbox. Cette option permet d'ajuster la variance du bruit pour chaque session. Pour un participant donné, nous avons ainsi modélisé les deux sessions d'une même tâche avec le même set de paramètres, mais les deux tâches étaient modélisées

## Études expérimentales

indépendamment. La VBA\_toolbox nous a ainsi permis d'obtenir, pour un participant donné, le set de paramètres  $\beta_0$ ,  $\beta_{|E|}$ ,  $\beta_E$  et  $\beta_{NE}$  qui permet de maximiser l'énergie libre du modèle selon l'approche variationnelle bayésienne. Nous avons ainsi obtenu un set de paramètres pour chaque participant pour chaque tâche.

Par la suite, nous avons effectué des t.tests sur chacun de ces paramètres pour voir s'ils étaient significativement différents de zéros à travers tous les participants.

- **Modèle computationnel de la performance en fonction d'une ressource X**

Nous avons effectué un second modèle de la performance dans les deux tâches à l'aide de la VBA\_toolbox (Daunizeau et al., 2014). Le modèle a aussi été inversé en utilisant l'approche variationnelle bayésienne avec une approximation de Laplace. Ce second modèle était une adaptation d'un modèle qui a déjà été publié pour modéliser la performance dans des tâches d'effort physique (Le Bouc and Pessiglione, 2018; Le Bouc et al., 2016; Pessiglione et al., 2018). Ce modèle partait du principe qu'à chaque essai les participants cherchaient à maximiser un rapport bénéfices/coûts. En effet, d'un côté, les participants cherchaient à maximiser leurs profits en obtenant ou en évitant de perdre le plus d'argent que possible à chaque essai. De l'autre côté, nous considérons que nos deux tâches impliquent un coût que nous avons exprimé à l'aide d'une variable  $X$ . Ce coût peut aussi bien correspondre au coût d'effectuer une action physique qu'au coût d'engager des ressources au niveau neural. Les participants devaient ainsi chercher à minimiser cette variable subjective de coût  $X$  à chaque essai tout en maximisant leurs gains monétaires. Nous avons ainsi défini que la valeur attendue (*expected value – EV*) à un essai  $i$  donné dépendait d'un côté d'un terme de bénéfices  $B_i$  et de l'autre d'un terme de coût  $C_i$  selon la formule :

$$EV(X_i) = B_i * P(X_i) - C_i * X_i$$

(Équation 4)

$X_i$  est la ressource dépensée à l'essai  $i$ ,  $P(X_i)$  correspond à la performance espérée, en termes de hauteur sur le thermomètre entre 0 (bas de l'échelle) et 1 (haut de l'échelle) en fonction de de la quantité  $X_i$  de ressources allouées pour l'essai. Le terme de bénéfice  $B_i$  se décomposait selon la formule :

$$B_i = 1 + k_{|E|} * |E|_i + k_E * E_i$$

(Équation 5)

## Études expérimentales

$k_{|E|}$  représente la sensibilité aux enjeux,  $|E|/i$  représente l'enjeu en euros pour l'essai  $i$  en valeur absolue,  $k_E$  est la sensibilité représente la sensibilité aux enjeux en fonction de leur valence et  $E_i$  représente l'enjeu en euros pour l'essai  $i$  en valeur signée. Le terme de coût  $C_i$  se décomposait selon la formule :

$$C_i = k_C * (1 + k_{NE} * NE_i)$$

(Équation 6)

$k_C$  est une constante représentant la sensibilité aux coûts,  $k_{NE}$  est une constante représentant l'augmentation du coût au fur et à mesure des essais et  $NE_i$  est le numéro de l'essai  $i$  dans la session. Nous avons posé que la performance variait en fonction de la quantité de ressources  $X$  selon la formule :

$$P(X) = \frac{X}{k_X + X}$$

(Équation 7)

où  $k_X$  est une constante représentant le montant de ressources nécessaires pour atteindre le haut de l'échelle que nous contraignons à être positive. En d'autres termes, cette équation implique que la performance s'approche du haut de l'échelle lorsque les ressources  $X$  investies sont maximales et, à l'inverse, qu'elle est nulle lorsqu'aucune ressource n'est investie.

Pour trouver la quantité optimale de ressources  $X$  à allouer à chaque essai, nous sommes partis du principe que le cerveau chercherait, à chaque essai, la quantité de ressources  $X$  pour laquelle la dérivée  $\frac{dEV}{dX}$  de la valeur attendue  $EV$  en fonction de  $X$  serait égale à zéro. Cette équation n'avait qu'une seule solution cohérente avec le fait que la quantité de ressources  $X$  à allouer soit positive. Pour chaque essai, la quantité optimale de ressources à allouer, d'après notre modèle, pouvait ainsi s'écrire selon la formule :

$$X_i = k_X * \left( \sqrt{\frac{B_i}{k_X * C_i}} - 1 \right)$$

(Équation 8)

De cette (Équation 8), découle l'équation de la performance optimale à chaque essai  $i$  suivante:

$$P_i = 1 - \sqrt{k_X * \frac{C_i}{B_i}}$$

(Équation 9)

A part  $k_X$  que nous avons forcé à être positif, les autres paramètres étaient libres. Cependant, nous avons forcé le terme  $B_i$  et le terme  $C_i$  à être positifs puisque, si ce n'était pas le cas, la formule de (Équation 9) n'aurait pas de sens. La VBA\_toolbox nous a ainsi permis d'obtenir, pour un participant donné, le set de paramètres  $k_{|E|}$ ,  $k_E$ ,  $k_C$ ,  $k_{NE}$  et  $k_X$  qui fait correspondre au mieux cette performance optimale qui serait calculée par le cerveau, d'après notre modèle, à la performance réellement effectuée à chaque essai. Ce set de paramètres est obtenu de sorte à maximiser l'énergie libre du modèle selon l'approche variationnelle bayésienne. Pour chaque participant, nous avons modélisé les données des deux sessions d'une tâche avec les mêmes paramètres. Nous avons utilisé l'option « multisession » de la VBA toolbox pour prendre en compte les différences éventuelles entre les 2 sessions au niveau du bruit des données. Nous avons modélisé les données indépendamment pour chaque tâche. Nous avons ainsi obtenu un set de paramètres pour chaque participant pour chaque tâche.

- **Comparaison des modèles**

Pour vérifier que notre second modèle computationnel basé sur l'idée d'une ressource  $X$  à maximiser était plus pertinent pour expliquer les données que le première modèle linéaire un peu plus simpliste, nous avons effectué une comparaison de modèles. Pour ce faire nous avons utilisé une approche bayésienne en utilisant la fonction VBA\_groupBMC de la VBA toolbox<sup>24</sup>. Cette fonction permet d'estimer la fréquence respective de chaque modèle sur la base de l'énergie libre qui y est associée, ainsi que la probabilité d'excédance de chaque modèle. La probabilité d'excédance consiste à estimer la probabilité qu'un modèle donné soit plus fréquent que les autres modèles du set de modèles qui sont comparés. Le seuil utilisé nous permettait de limiter le nombre de faux positifs à 5%.

- **Analyse des temps de réaction**

Le temps de réaction n'était instrumental que dans la tâche d'effort mental, mais il pourrait refléter des mécanismes neuraux plus généraux qui faciliteraient une approche plus rapide face à des gains et plus lente face à des pertes. Nous avons ainsi effectué un GLM pour

---

<sup>24</sup> <http://mbb-team.github.io/VBA-toolbox/wiki/BMS-for-group-studies/#rfx-bms>



## Études expérimentales

voir si le temps de réaction dans chacune de nos tâches était impacté par les variables affectant la performance. Dans le cas de la tâche d'effort physique, nous avons défini le temps de réaction comme la différence entre le moment d'apparition de l'échelle d'effort à l'écran et le moment où la force mesurée dans Matlab en volts était supérieure à 1% de la force nécessaire pour atteindre le haut de l'échelle. Nous avons choisi ce seuil parce qu'il semblait suffisant pour éliminer les oscillations liées au bruit de la mesure par rapport à la force réellement exercée par les participants. Dans le cas de la tâche d'effort mental, nous avons pris en compte le temps de réaction pour la première paire, c'est-à-dire la différence entre le moment de la première réponse donnée au cours d'un essai et le moment d'affichage de l'échelle d'effort à l'écran.

Pour ce faire, nous avons défini un GLM dans Matlab où la variable à expliquer était le temps de réaction dans les deux sessions de chaque tâche. Les variables explicatives incluaient une constante par session pour prendre en compte les différences moyennes éventuelles d'une session à l'autre, les enjeux monétaires en euros signés, les enjeux monétaires en euros non-signés et le numéro d'essai.

Nous avons effectué ce GLM indépendamment pour chaque participant dans chacune des tâches. Ensuite, pour vérifier si nos variables avaient un impact significatif sur le temps de réaction nous avons testé, avec un t.test, si l'ensemble des betas des participants pour chacune des variables étaient significativement différents de zéro.

### II.2.d IRM

- **Acquisition**

Les données d'IRM ont été acquises au Centre de Neuro-Imagerie de Recherche (CENIR, Institut du Cerveau, ICM, Paris) avec un scanner 3 Teslas de modèle Siemens Magnetom Prisma et avec une antenne de tête Siemens à 64 canaux.

Les images d'IRM anatomique ont été obtenues avant de démarrer l'expérience. Les paramètres d'acquisition de ces images anatomiques sont détaillés dans le tableau ci-dessous (voir **Tableau 4**).

Temps de répétition (TR) (secondes)	Temps d'écho (millisecondes)	Angle de bascule (°C)	Nombre de coupes	Epaisseur de coupes (millimètres)	Taille des voxels (mm <sup>3</sup> )
2.3	4.18	9	176	1	1

**Tableau 4 : Paramètres d’acquisition de l’IRM anatomique pour les participants de la deuxième et de la troisième étude.**

Les images d’IRM fonctionnelle ont été obtenues au cours de 7 sessions d’acquisition. La durée d’un essai dans la tâche d’effort mental dépendait du temps calibré au début de chaque session. De ce fait, le nombre de volumes par session et par sujet variait d’une session et d’un sujet à l’autre.

Les paramètres utilisés pour l’acquisition des données d’imagerie fonctionnelle sont détaillés dans le tableau ci-dessous.

Temps de répétition (TR) (secondes)	Temps d’écho (millisecondes)	Angle de bascule (°C)	Nombre de coupes	Épaisseur de coupe (millimètres)	Facteur d’accélération multi-bandes
1.10	25	60	54	2.0	3

**Tableau 5 : Paramètres d’acquisition de l’IRM fonctionnelle pour la tâche d’effort physique, la tâche d’effort mental et la tâche d’apprentissage.**

- **Pré-traitement**

Les données d’IRMf ont été prétraitées et analysées avec la toolbox SPM12 (Wellcome Trust Center for NeuroImaging, Institute Of Neurology, Londres, Royaume-Uni) dans Matlab 2017a (The MathWorks, Inc., USA). Les images fonctionnelles ont d’abord été réalignées. Par la suite, elles ont été co-enregistrées avec l’image anatomique de chaque sujet. Enfin, elles ont été normalisées dans l’espace MNI (*Montreal Neurological Institute*) et lissées avec un facteur de 8mm.

- **Analyse des données d’IRMf : les modèles utilisés**

Après le prétraitement, nous avons analysé les données à l’aide de modèles linéaires généralisés (*general linear model*, GLM) dans SPM12 au premier niveau pour chaque sujet. Pour chaque sujet et chaque session, 6 paramètres de mouvement étaient directement estimés par SPM12 et rajoutés au sein du GLM. Ensuite, nous avons effectué une analyse au second niveau en regardant la significativité des paramètres issus du premier niveau au niveau du groupe.

Nous avons effectué un premier GLM, le GLM1, similaire entre les deux tâches, qui incluait les variables expliquant la performance. Dans ce GLM, nous avons modélisé séparément les 3 étapes de chaque essai. Nous avons modélisé le moment d’apparition des

## Études expérimentales

enjeux et le moment d'apparition du feedback avec une fonction delta. Le moment d'apparition de l'échelle d'effort était modélisé à l'aide d'une fonction boxcar qui englobait toute la période où il était possible d'effectuer un effort. La fonction delta modélisant l'apparition des enjeux était aussi modulée par les 3 régresseurs paramétriques qui expliquaient la performance : 1) les enjeux monétaires signés, 2) les enjeux monétaires non-signés et 3) le numéro d'essai.

Nous avons effectué un second GLM, le **GLM2**, qui était similaire au GLM1. La seule différence était basée sur le fait que nous avons remplacé les modulateurs paramétriques au moment de l'apparition des enjeux à l'écran par la variable X représentant la quantité de ressources allouée à chaque essai d'après notre modèle computationnel (voir *Modèle computationnel de la performance en fonction d'une ressource X*).

- **Régions d'intérêt**

Nous avons utilisé 3 régions d'intérêt basée sur la plateforme en ligne Neurosynth. Nous avons d'un côté employé une région d'intérêt dans le striatum ventral liée au mot « *incentive* » (enjeux) dans Neurosynth. De l'autre, nous avons défini deux régions d'intérêt, une région dans le cortex préfrontal dorsomédian (dmPFC) et une région bilatérale dans le cortex préfrontal dorsolatéral (dlPFC), en lien avec le terme « *cognitive control* » (contrôle cognitif). Chacune de ces cartes a été extraite sur la base du test uniforme de Neurosynth et nous avons extrait les clusters après avoir filtré les cartes pour ne garder que les voxels survivant avec une t.value > 12 dans chaque cas.

## II.3 Résultats

### II.3.a Résultats comportementaux

- **Contrôle de l'effet Stroop dans la tâche d'effort mental**

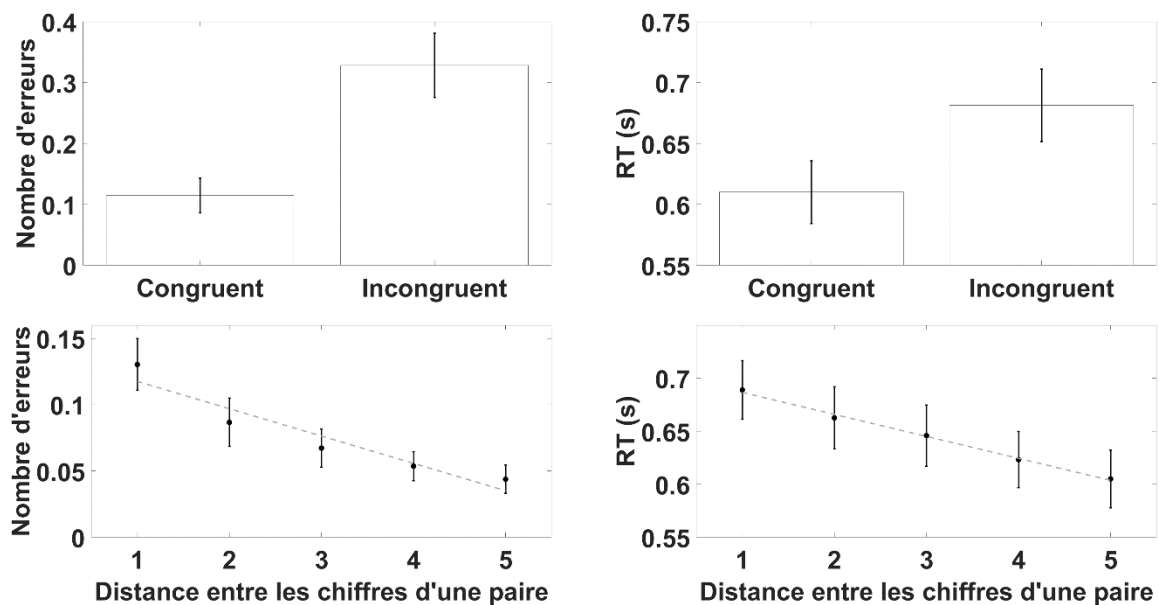
En premier lieu, nous avons vérifié que la tâche de Stroop numérique donnait bien des résultats en accord avec la littérature, suggérant qu'elle impliquait d'employer du contrôle cognitif.

Les participants ont, en moyenne, résolu 3.0 paires congruentes par essai contre 2.9 paires incongruentes par essai. Cette différence est certes minime, mais elle était significativement différente (p.value = 0.006). Les participants ont aussi commis plus d'erreurs (p.value =  $1 \cdot 10^{-6}$ ) lorsque les paires à résoudre étaient incongruentes ( $n = 0.33 \pm 0.05$  paires incongruentes résolues avec au moins une erreur par essai, moyenne  $\pm$  erreur type de la moyenne) que lorsqu'elles étaient congruentes ( $n = 0.11 \pm 0.03$  paires avec au moins une erreur

## Études expérimentales

par essai) (voir **Figure 30**). Les participants ont, en moyenne, mis  $0.61 \pm 0.03$  secondes (moyenne  $\pm$  erreur type de la moyenne) pour répondre aux paires congruentes contre  $0.68 \pm 0.03$  secondes pour répondre aux paires incongruentes (voir **Figure 30**). Leur temps de réaction moyen était ainsi plus lent pour les paires incongruentes que pour les paires congruentes ( $p.value = 5 \cdot 10^{-8}$ ).

La distance entre les deux chiffres d'une paire avait aussi un impact sur le nombre d'erreurs commises et sur le temps de réaction (voir **Figure 30**). Les participants commettaient en effet de moins en moins d'erreurs à mesure que la distance entre les chiffres augmentait ( $b = -0.021 \pm 0.003$  ;  $p.value = 1 \cdot 10^{-6}$ ). Leur temps de réaction était aussi de plus en plus rapide à mesure que la distance entre les chiffres augmentait ( $b = -0.021 \pm 0.002$  ;  $p.value = 2 \cdot 10^{-11}$ ).



**Figure 30 : Différents niveaux de difficulté en fonction du type de paire dans la tâche d'effort mental.** En haut, le nombre moyen d'erreurs par essai, à gauche, et le temps de réaction moyen, à droite, est représenté en fonction du type de paires. En bas, les figures représentent le nombre moyen d'erreurs par essai, à gauche, et le temps de réaction moyen, à droite, en fonction de la différence numérique entre les deux chiffres de chaque paire. Sur les deux figures du bas, les points avec les barres d'erreur représentent la moyenne ainsi que l'erreur type de la moyenne à travers les participants du nombre d'erreurs par essai à gauche et du temps de réaction à droite après avoir séparé les données en fonction de la distance entre les chiffres de chaque paire. Les traits en pointillés représentent la moyenne au travers des participants des données estimées par le GLM pour chaque niveau de distance entre les chiffres d'une paire. Toutes les figures sont basées sur les 23 participants ayant correctement effectué les tâches au niveau comportemental.

Notre tâche d'effort mental nécessitait de résoudre, au plus vite, les 10 paires présentées à chaque essai. Les paires, au sein d'un essai, variaient en difficulté principalement du fait de leur nature congruente ou incongruente et de la distance entre les deux chiffres qui les

## Études expérimentales

constituaient. Même si les paires étaient fixes à travers les essais, les participants ne pouvaient pas anticiper leur ordre d'apparition. Au cours d'un essai, ils devaient donc constamment adapter leur niveau d'effort mental pour répondre aux difficultés de chaque paire rencontrée et inhiber les biais automatiques dans la réponse donnée.

- **Ce qui détermine la performance et le temps de réaction**

Dans nos tâches, les participants étaient confrontés à des essais où ils jouaient pour augmenter leurs gains et à d'autres essais où ils jouaient pour ne pas perdre une partie des gains déjà accumulés. L'augmentation des gains ou la réduction des pertes impliquaient cependant exactement le même type d'effort : serrer une poignée de force pour la tâche d'effort physique et résoudre un nombre maximal de paires de Stroop numérique dans le temps imparti pour la tâche d'effort mental. Les participants pouvaient adapter leur niveau de performance en fonction de la valeur absolue des enjeux monétaires en vue de ne pas dépenser inutilement des ressources lorsque les montants étaient faibles. Nous avons ainsi décidé de modéliser la performance en fonction des enjeux monétaires non-signés. Aussi bien dans la tâche d'effort physique ( $\beta_{|E|} = 0.0118 \pm 0.0012$  (moyenne  $\pm$  erreur type de la moyenne à travers les participants) ;  $p.value = 2 \cdot 10^{-9}$ ), que dans la tâche d'effort mental ( $\beta_{|E|} = 0.0019 \pm 5 \cdot 10^{-4}$  ;  $p.value = 0.0012$ ), nous avons alors constaté que la performance augmentait en effet en fonction des enjeux monétaires non-signés, c'est-à-dire que les participants adaptaient leur niveau d'effort en fonction des enjeux (voir **Figure 33**).

De plus, l'aversion à la perte suppose une asymétrie entre le traitement d'une perspective de gains et le traitement d'une perspective de pertes. Nous avons ainsi inclus, dans notre modèle linéaire de la performance, en plus des enjeux non-signés, une variable correspondant aux enjeux signés en vue de capter toute asymétrie éventuelle qui serait liée à un traitement différent entre les gains et les pertes du fait d'un impact linéaire des enjeux signés sur la performance. Nous avons alors constaté que, contrairement à ce qu'une aversion à la perte aurait prédit, la performance des participants augmentait linéairement avec les enjeux monétaires signés ( $\beta_E = 8 \cdot 10^{-4} \pm 2 \cdot 10^{-4}$  ;  $p.value = 0.0013$ ) dans la tâche d'effort physique et dans la tâche d'effort mental ( $\beta_E = 8 \cdot 10^{-4} \pm 3 \cdot 10^{-4}$  ;  $p.value = 0.030$ ).

Enfin, ce type de tâche est réputé entraîner une certaine fatigue, que ce soit au niveau musculaire ou au niveau mental. Nous avons ainsi décidé de rajouter une variable correspondant au numéro d'essai dans notre modèle linéaire en vue de capter toute baisse de la performance qui pourrait être liée à de la fatigue. Nous avons alors observé en effet que la performance

## Études expérimentales

chutait au fur et à mesure des essais (voir **Figure 33**) dans la tâche d'effort physique ( $\beta_{NE} = -0.0013 \pm 3 \cdot 10^{-4}$ ; p.value =  $6 \cdot 10^{-4}$ ) et dans la tâche d'effort mental ( $\beta_{NE} = -7 \cdot 10^{-4} \pm 2 \cdot 10^{-4}$ ; p.value = 0.0035).

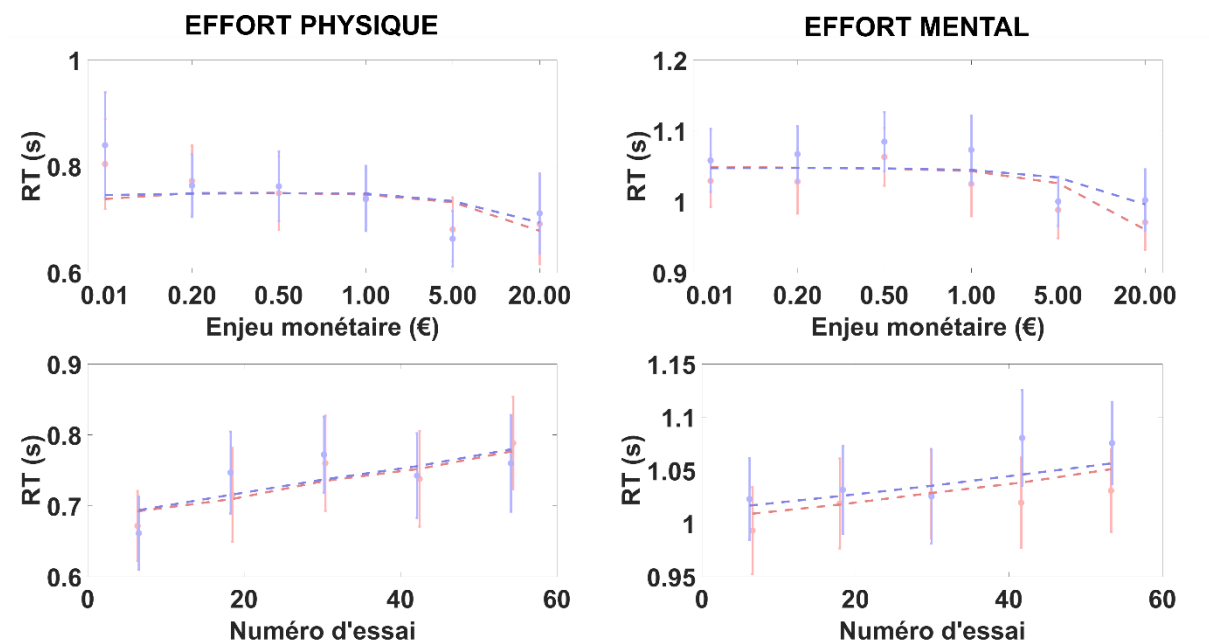
Le  $R^2$  du modèle qui permet de voir dans quelle mesure le modèle s'ajuste bien aux données était de  $0.373 \pm 0.031$  (moyenne  $\pm$  erreur type de la moyenne à travers les participants) dans la tâche d'effort physique et de  $0.061 \pm 0.006$  dans la tâche d'effort mental. Il faut noter que la performance avait une variance nettement plus faible dans la tâche d'effort mental par rapport à la tâche d'effort physique (voir **Figure 33**) ce qui explique, au moins en partie, pourquoi les facteurs testés avaient une incidence plus faible dans la tâche d'effort mental en comparaison à la tâche d'effort physique et aussi pourquoi les  $R^2$  sont aussi différents entre ces deux tâches. En résumé, dans les deux tâches, la performance augmentait aussi bien avec les enjeux signés qu'avec les enjeux non-signés et elle baissait avec le temps.

Nous avons constaté que, contrairement à ce que l'aversion à la perte prédit, la performance avait plutôt tendance à augmenter dans le domaine des gains que dans le domaine des pertes. Nous avons pensé que cet effet pourrait être lié à un effet pavlovien consistant à avoir une réponse de vigueur et de vitesse plus importantes en réponse à des gains appétitifs par rapport à des pertes aversives. Pour vérifier si un effet de cette nature était impliqué dans nos tâches, nous avons regardé si le temps de réaction dans chacune des tâches variait avec les variables que nous avons identifiées en lien avec la performance, en particulier avec les enjeux monétaires signés. Pour la tâche d'effort physique, nous avons considéré le temps de réaction comme la différence entre le moment d'apparition de l'échelle d'effort à l'écran et le moment où le participant commençait à serrer la poignée de force. Pour la tâche d'effort mental, nous nous sommes concentrés sur le temps mis pour répondre à la première paire de chaque essai, à partir de l'apparition des paires à l'écran. Nous avons alors constaté que le temps de réaction était significativement plus rapide lorsque les enjeux non-signés étaient plus élevés dans la tâche d'effort mental ( $\beta = -0.0035 \pm 9 \cdot 10^{-4}$ ; p.value =  $5 \cdot 10^{-4}$ ). Cette corrélation n'était pas significative dans la tâche d'effort physique où les temps de réaction suivaient cependant la même tendance ( $\beta = -0.0032 \pm 0.0029$ ; p.value = 0.276). De plus, nous avons constaté que le temps de réaction dans la tâche d'effort mental affichait une tendance non-significative à diminuer linéairement avec les enjeux monétaires signés ( $\beta = -9 \cdot 10^{-4} \pm 0.009$ ; p.value = 0.083). La tendance était similaire dans la tâche d'effort physique, mais elle n'était pas significative là non plus ( $\beta = -4 \cdot 10^{-4} \pm 5 \cdot 10^{-4}$ ; p.value = 0.370). Enfin, au fur et à mesure des essais, les participants répondaient de plus en plus lentement (voir **Figure 31**) dans la tâche d'effort mental

## Études expérimentales

( $\beta = 8 \cdot 10^{-4} \pm 4 \cdot 10^{-4}$ ; p.value = 0.044) et dans la tâche d'effort physique ( $\beta = 0.0018 \pm 7 \cdot 10^{-4}$ ; p.value = 0.021).

En résumé, on voit que la plupart des variables impactant la performance n'avait un effet significatif sur les temps de réaction que dans le cas de la tâche d'effort mental. Il faut d'ailleurs préciser que le temps de réaction avait un rôle instrumental dans la tâche d'effort mental. Répondre plus vite et correctement à la première paire d'un essai permet en effet, *a priori*, d'avoir une meilleure performance globale puisqu'il reste alors plus de temps pour répondre aux paires suivantes. A l'inverse, dans la tâche d'effort physique, seule la force maximale mais pas la vitesse de réaction importait pour la performance. Nous n'avons donc pas pu observer d'effet significatif des enjeux monétaires signés, qui auraient pu refléter un effet pavlovien, sur les temps de réaction dans aucune de nos deux tâches. Par contre, il semblerait que le numéro d'essai qui pourrait refléter une augmentation de la fatigue ou un désintérêt croissant des participants pour la tâche en cours avait un impact dans les deux tâches. Les temps de réaction avaient en effet tendance à être de plus en plus lent à mesure que les participants progressaient dans la tâche.

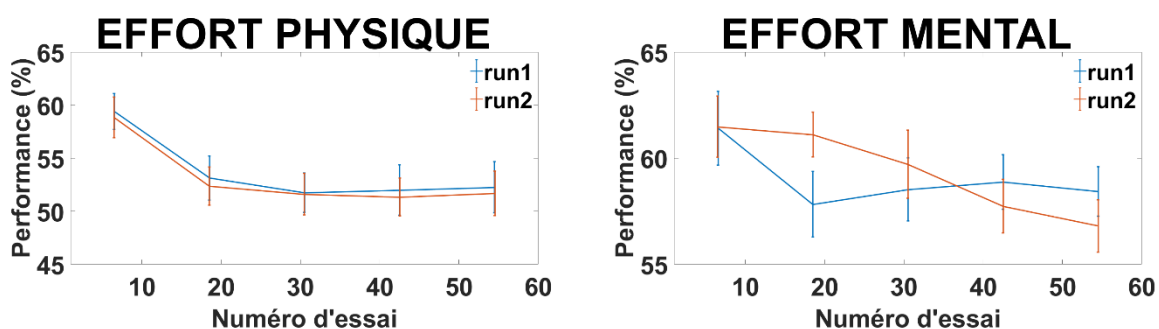


**Figure 31 : Temps de réaction en fonction des enjeux monétaires (en haut) et du numéro d'essai (en bas) dans la tâche d'effort physique (à gauche) et dans la tâche d'effort mental (à droite).** Les données ont été séparées en fonction du type d'essai. En rouge, on peut voir les courbes correspondant aux essais où les participants jouaient pour gagner une partie du montant en jeu. En bleu, les données correspondent aux essais où les participants jouaient pour ne pas perdre une partie du montant en jeu sur leur total accumulé au fil des essais. Les points ainsi que les barres d'erreur représentent la moyenne ainsi que l'erreur type de la moyenne à travers les participants de la performance pour chaque enjeu monétaire (en haut) et pour chaque bin moyennant 5 essais en fonction du numéro d'essai (en bas). Les traits en pointillés représentent

## Études expérimentales

la moyenne correspondante à travers les participants de l'estimation par notre GLM du temps de réaction pour chacune des conditions. Les données des figures sont basées sur les 23 participants inclus dans l'étude. Notez que l'échelle des abscisses, pour les enjeux monétaires, est définie de sorte à ce que la distance entre chaque montant soit uniforme et pas en fonction de la valeur réelle.

Voyant que la performance et les temps de réaction étaient impactés par le numéro d'essai, c'est-à-dire le temps passé à faire la tâche, nous avons pensé qu'il pourrait s'agir là du reflet d'une baisse de motivation lié soit 1) à une augmentation de la fatigue au cours de la tâche soit 2) à une baisse de l'aspect appétitif (ou aversif) relatif des enjeux au cours de la tâche une fois qu'une certaine somme a déjà été accumulée. Ce deuxième point refléterait en quelque sorte une évolution dynamique du point de référence des participants au cours de la tâche. Le cas échéant, un décalage du point de référence qui sert à définir l'aversion à la perte pourrait en partie expliquer pourquoi nous ne constatons pas d'aversion à la perte dans nos tâches, mais plutôt un biais dans le sens inverse (performance plus élevée pour des gains équivalents que pour des pertes). Pour tester ces deux hypothèses, nous avons voulu vérifier si cette baisse de performance et l'augmentation du temps de réaction étaient liées au fait que les participants accumulaient des gains au fur et à mesure du temps, ou s'il s'agissait plutôt d'un effet lié à la fatigue au cours d'une session. Nous avons ainsi décidé de regarder la performance en fonction du numéro d'essai séparément dans chaque session (voir **Figure 32**). Comme on peut le voir sur la figure, la performance démarrait toujours au même niveau au début de chaque session et baissait au cours de chacune des sessions suggérant que la baisse de la performance et l'augmentation des temps de réaction au cours du temps relevait plus d'un effet de fatigue au sein d'une session que d'un changement du point de référence des participants.



**Figure 32 : Performance en fonction du numéro d'essai au sein de chaque session** de la tâche d'effort physique, à gauche, et de la tâche d'effort mental, à droite. La figure représente la moyenne ainsi que l'erreur type de la moyenne à travers les participants de la performance divisée en 5 bins en fonction du numéro d'essai pour chaque session. Les traits relient les points entre eux mais ne correspondent pas à un fit de la performance. A gauche, on peut voir les résultats pour la tâche d'effort physique et à droite de même pour la tâche d'effort mental. Les



données de la première session sont représentées en bleu et celles de la deuxième session en rouge. Les données des figures sont basées sur les 23 participants inclus dans l'étude.

- **Le modèle computationnel basé sur l'allocation d'une ressource X**

L'effort correspond à une sensation aversive qui accompagne l'exécution d'une action musculaire ou d'une action requérant l'utilisation du système neural exécutif. Cette sensation va de pair avec l'idée de la dépense de ressources physiologiques (ATP, glucides, lipides) et éventuellement aussi de l'accumulation de déchets (lactate, protéine amyloïde beta) à éliminer. Si cette sensation d'effort n'existait pas, nos participants auraient pu chercher à effectuer une performance maximale à chaque essai indépendamment des enjeux. Or, comme nous l'avons vu, leur performance était modulée par le montant des récompenses ou des punitions en jeu. C'est donc qu'ils cherchaient à optimiser l'effort alloué à chaque essai. Pour modéliser la quantité de ressources allouée par nos participants à chaque essai, sans chercher aucunement à définir la nature de cette ressource, nous avons ainsi défini un modèle computationnel basé sur l'idée de l'optimisation d'une ressource X. L'idée de ce modèle était que les participants cherchaient à la fois d'un côté à maximiser leurs gains monétaires et, de l'autre, à minimiser la dépense de cette ressource X inférée par le modèle mais que nous ne pouvions mesurer ne sachant pas quelle est sa nature précise.

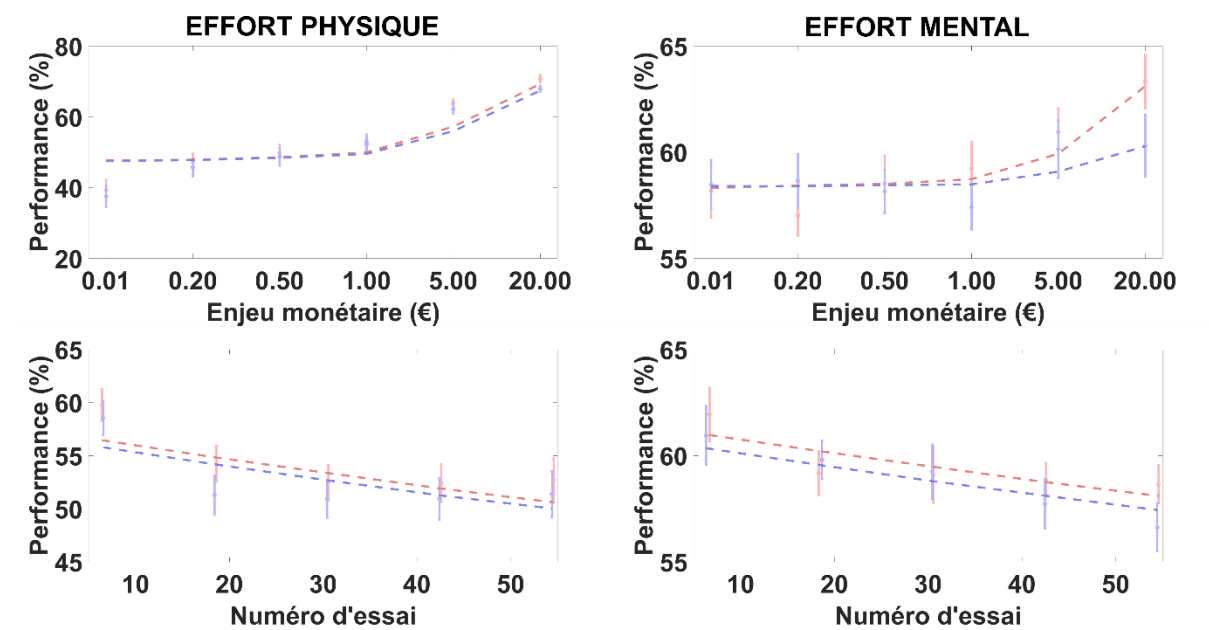
Dans les deux tâches effectuées par nos participants, notre comparaison de modèle nous a permis de voir que le modèle incluant la notion d'une ressource X à maximiser avait une probabilité de dépassement (*exceedance probability*) nettement plus importante ( $p = 1$  dans les deux tâches) que le modèle linéaire ( $p = 0$ ). De même, la fréquence du modèle linéaire n'était que de 2% dans chacune des deux tâches contre 98% pour le second modèle. Ce modèle permet donc non seulement d'inférer la quantité de ressources X qui ne peut être directement mesurée, mais aussi de mieux rendre compte des données.

Dans ce modèle, la sensibilité moyenne aux enjeux monétaires était de  $k_{|E|} = 1.842 \pm 0.190$  (moyenne  $\pm$  erreur type de la moyenne à travers les participants) dans la tâche d'effort physique et de  $k_{|E|} = 0.240 \pm 0.052$  dans la tâche d'effort mental. La sensibilité aux enjeux monétaires signés était de  $k_E = 0.166 \pm 0.044$  dans la tâche d'effort physique et de  $k_E = 0.089 \pm 0.038$  dans la tâche d'effort mental. Dans le terme de coût, la sensibilité moyenne aux coûts était de  $k_C = 0.225 \pm 0.017$  dans la tâche d'effort physique et de  $k_C = 0.155 \pm 0.008$  dans la tâche d'effort mental. La sensibilité au numéro d'essai était de  $k_{NE} = 0.431 \pm 0.094$  dans la tâche d'effort physique et de  $k_{NE} = 0.240 \pm 0.062$  dans la tâche d'effort mental. Enfin, le terme  $k_X$  était de  $1.059 \pm 0.010$  dans la tâche d'effort physique et de  $1.026 \pm 0.003$  dans la tâche d'effort

## Études expérimentales

mental. Le  $R^2$  du modèle qui permet de voir dans quelle mesure le modèle s'ajuste bien aux données était de  $0.416 \pm 0.033$  (moyenne  $\pm$  erreur type de la moyenne à travers les participants) dans la tâche d'effort physique et de  $0.061 \pm 0.006$  dans la tâche d'effort mental.

En résumé, la performance augmentait avec les enjeux, elle était sensible au fait de jouer pour gagner ou pour ne pas perdre de l'argent et elle diminuait avec le numéro d'essai (voir **Figure 33**).



**Figure 33 : Performance en fonction des enjeux monétaires (en haut) et du numéro d'essai (en bas) dans la tâche d'effort physique (à gauche) et dans la tâche d'effort mental (à droite).** Les données ont été séparées en fonction du type d'essai. En rouge, on peut voir les courbes correspondant aux essais où les participants jouaient pour gagner une partie du montant en jeu. En bleu, les données correspondent aux essais où les participants jouaient pour ne pas perdre une partie du montant en jeu sur leur total accumulé au fil des essais. Les points ainsi que les barres d'erreur représentent la moyenne ainsi que l'erreur type de la moyenne à travers les participants de la performance pour chaque enjeu monétaire (en haut) et pour chaque bin moyennant 5 essais en fonction du numéro d'essai (en bas). Les traits en pointillés représentent la moyenne correspondante à travers les participants de l'estimation par notre second modèle computationnel de la performance pour chacune des conditions. Les données des figures sont basées sur les 23 participants inclus dans l'étude. Notez que l'échelle des abscisses, pour les enjeux monétaires, est définie de sorte à ce que la distance entre chaque montant soit uniforme et pas en fonction de la valeur réelle.

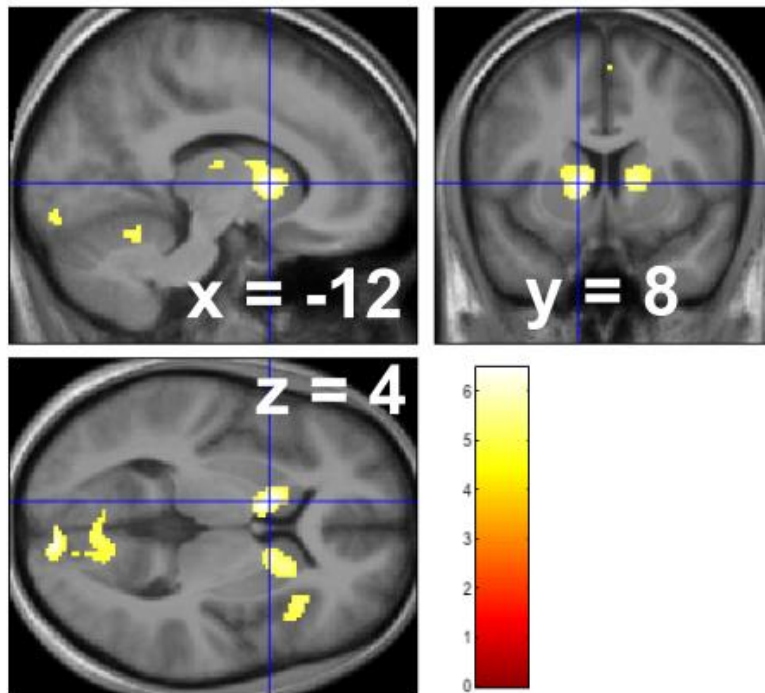
### II.3.b Résultats neuraux

Ayant identifié les variables qui expliquaient la performance, nous avons voulu voir si nous pouvions identifier des corrélats cérébraux pour chacune de ces variables. Nous nous sommes concentrés sur ce qui était commun aux deux tâches, c'est-à-dire les variables modulant la performance indépendamment du type d'effort qui devait être exécuté. Aussi, nous avons

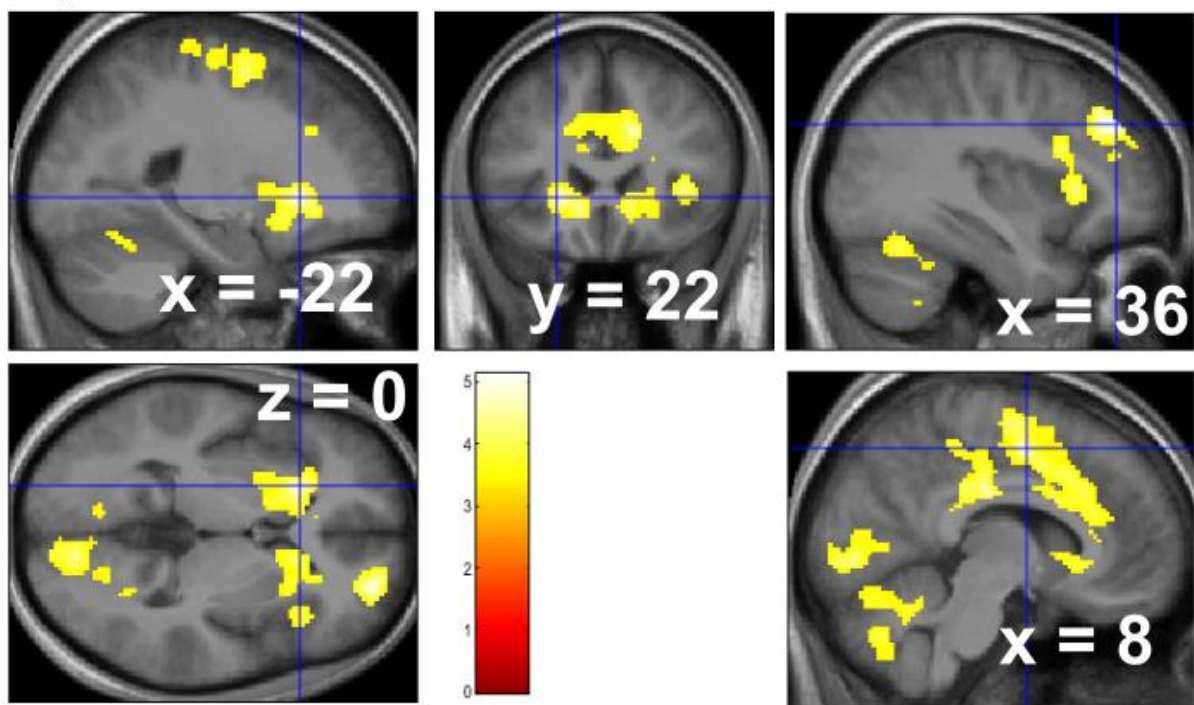
## Études expérimentales

systematiquement utilisé des conjonctions entre nos deux tâches pour être sûrs que nous ne voyions que ce qui était commun aux deux tâches. La seule carte permettant d'observer des corrélats cérébraux significatifs après correction pour les comparaisons multiples, que ce soit au niveau des voxels ou au niveau des clusters, en observant des corrélats positifs ou négatifs de chacune des variables était la variable représentant les enjeux monétaires non-signés. Cette carte révélait un cluster bilatéral au niveau du striatum (voir **Figure 34**). En d'autres termes, l'activité du striatum, au moment de l'apparition des enjeux, augmentait avec l'importance des enjeux, indépendamment de la valence de ceux-ci. On peut aussi noter des clusters dans des aires ayant été associée avec l'anticipation des coûts d'un effort avec, entre autres, un cluster au niveau de l'insula antérieure droite avec un pic aux coordonnées MNI (34 ; 22 ; 6) ainsi que trois clusters autour du dmPFC et de l'aire supplémentaire motrice avec des pics aux coordonnées MNI (2 ; 18 ; 36), (4 ; -2 ; 60) et (8 ; -22 ; 36).

## A) |Enjeux monétaires|



## B) Ressource X



**Figure 34 : Résultats neuraux.** Les cartes représentent le résultat d'une conjonction entre la tâche d'effort physique et la tâche d'effort mental. Les cartes sont superposées à la moyenne des scans anatomiques des 20 cerveaux inclus dans l'étude après le pré-traitement, en particulier après normalisation dans l'espace MNI. Les barres colorées indiquent la correspondance entre les couleurs de la carte et les valeurs T pour chaque voxel. A) Carte représentant les corrélats cérébraux des enjeux monétaires non-signés dans le GLM1. La conjonction représentée sur la

## Études expérimentales

carte est seuillée à  $p < 0.05$  corrigée au niveau des voxels pour les comparaisons multiples au niveau de la famille. La carte est zoomée sur le cluster situé dans le striatum gauche au niveau du pic d'activation (-12 ; 8 ; 4). On peut aussi voir son symétrique du côté droit dont le pic est situé aux coordonnées (14 ; 10 ; 6). B] Carte représentant les corrélats cérébraux de la ressource X dans le GLM2. La conjonction représentée sur la carte est seuillée à  $p < 0.05$  corrigée au niveau des clusters pour les comparaisons multiples au niveau de la famille. La carte montre trois images différentes : à gauche, on peut observer une corrélation avec le striatum légèrement plus antérieure à l'autre carte avec un zoom sur le pic d'activation situé aux coordonnées MNI (-22 ; 22 ; 0) ; à droite, la carte montre deux zones incluses dans le même cluster qui englobe à la fois le cortex préfrontal dorsolatéral droit, visible en haut, avec un pic aux coordonnées (36 ; 40 ; 34) et le cortex préfrontal dorsomédian, visible en bas. À  $p < 0.05$  corrigée au niveau des voxels pour les comparaisons multiples, seul un cluster situé spécifiquement dans le cortex préfrontal dorsolatéral droit (pic en 36 ; 40 ; 34) survit.

Ensuite, nous avons voulu voir s'il était possible d'identifier des aires cérébrales positivement corrélées avec la quantité optimale de ressources X allouées à chaque essai d'après notre modèle computationnel. Nous espérons ainsi éventuellement découvrir des aires cérébrales nous permettant de déduire indirectement la nature de la/des ressource(s) qui étaient optimisées dans nos tâches. Cette carte a permis d'observer quatre grands clusters (voir **Table 2**) qui incluaient notamment le cortex préfrontal dorsomédian, le cortex préfrontal dorsolatéral, l'insula antérieure et le striatum (voir **Figure 34**). Il est intéressant de noter qu'en utilisant un seuil corrigé au niveau des voxels au lieu des clusters, on pouvait alors voir que le seul cluster survivant à ce seuil était situé dans le cortex dorsolatéral préfrontal, une aire régulièrement associée à l'utilisation du contrôle cognitif.

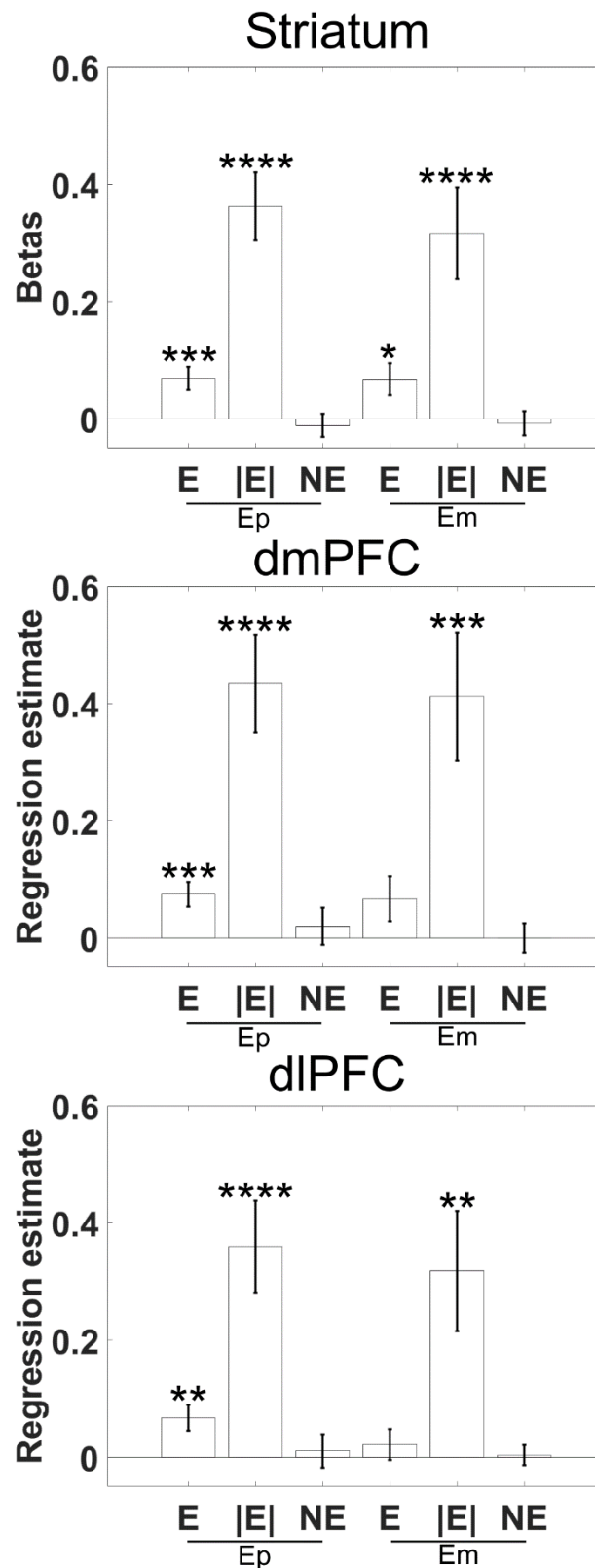
Nom (d'après l'atlas AAL)	x	y	z	Nombre de voxels
Gyrus Frontal Milieu droit	36	40	34	6348
Putamen gauche	-22	22	0	764
Gyrus lingual droit	10	-80	-2	4139
Gyrus précentral gauche	-30	-12	64	1167

**Table 2 : Table d'activation en lien avec la quantité de ressources allouées X.** La table correspond aux clusters issus de la conjonction entre la tâche d'effort physique et la tâche d'effort mental pour le contraste correspondant à la ressource X seuillée à  $p < 0.05$  corrigée au niveau des clusters pour les comparaisons multiples au niveau de la famille (*Family Wise Error Rate*). Les appellations des différentes aires ont été établies sur la base des coordonnées correspondantes dans l'atlas AAL (Automated Anatomical Labeling). Elles sont données à titre indicatif bien que, étant donné la taille des clusters, plusieurs autres aires anatomiques étaient englobées par chacun de ces clusters.

Enfin, nous avons voulu préciser les résultats observés à l'aide des cartes. En premier lieu, nous avons voulu vérifier si le striatum réagissait aussi aux autres variables impactant la performance, en particulier les enjeux monétaires signés puisqu'il a déjà été associé à la fois à

## Études expérimentales

la valeur subjective signée et non-signée. En second lieu, nous voulions aussi vérifier si les aires cérébrales associées à l'exercice de contrôle cognitif, le dmPFC et le dlPFC, réagissaient aussi aux différentes variables impactant la performance ou si leur activité n'était principalement liée qu'à la saillance des enjeux monétaires non-signés. Nous avons ainsi défini trois régions d'intérêt : le striatum a été défini sur la base du mot « incentive » (enjeux) et le dmPFC et le dlPFC ont été définis sur la base du mot « cognitive control » (contrôle cognitif) à l'aide de la plateforme en ligne Neurosynth. Nous avons alors constaté que l'activité du striatum corrélait positivement avec les enjeux monétaires non-signés dans chacune de nos tâches comme nous l'avons déjà vu avec notre conjonction (tâche d'effort physique (Ep) :  $\beta_{|E|} = 0.362 \pm 0.058$ ,  $p = 6 \cdot 10^{-6}$ ; tâche d'effort mental (Em) :  $\beta_{|E|} = 0.316 \pm 0.078$ ,  $p = 0.0007$ ). Son activité augmentait aussi avec les enjeux monétaires signés (voir **Figure 35**) dans chacune de nos tâches (Ep :  $\beta_E = 0.069 \pm 0.020$ ,  $p = 0.002$  ; Em :  $\beta_E = 0.068 \pm 0.027$ ,  $p = 0.022$ ). Cependant, son activité n'était pas significativement impactée par le numéro d'essai (Ep :  $\beta_{NE} = -0.011 \pm 0.020$ ,  $p = 0.571$  ; Em :  $\beta_{NE} = -0.008 \pm 0.021$ ,  $p = 0.709$ ) ce qui suggère une fois de plus que la baisse de performance n'était pas liée à une réduction de l'aspect appétitif des récompenses monétaires envisagées mais plutôt à un effet de fatigue. De plus, nous avons vu que les aires liées au contrôle cognitif voyaient aussi leur activité augmenter avec les enjeux monétaires non-signés (voir **Figure 35**). L'activité du dmPFC augmentait en effet avec les enjeux monétaires non-signés (Ep :  $\beta_{|E|} = 0.434 \pm 0.084$ ,  $p = 0.0001$ ; Em :  $\beta_{|E|} = 0.412 \pm 0.110$ ,  $p = 0.001$ ) et de même pour le dlPFC (Ep :  $\beta_{|E|} = 0.359 \pm 0.079$ ,  $p = 0.0002$  ; Em :  $\beta_{|E|} = 0.318 \pm 0.103$ ,  $p = 0.006$ ). Les enjeux monétaires signés étaient aussi corrélés avec l'activité du dmPFC (Ep :  $\beta_E = 0.074 \pm 0.021$ ,  $p = 0.002$  ; Em :  $\beta_E = 0.067 \pm 0.039$ ,  $p = 0.099$ ) et du dlPFC (Ep :  $\beta_E = 0.068 \pm 0.022$ ,  $p = 0.006$  ; Em :  $\beta_E = 0.022 \pm 0.026$ ,  $p = 0.414$ ) dans la tâche d'effort physique mais cette relation n'était pas significative dans la tâche d'effort mental. Enfin, aucune des deux aires liées au contrôle cognitif que nous avons étudié n'était impactée par le numéro d'essai, que ce soit le dmPFC (Ep :  $\beta_{NE} = 0.020 \pm 0.032$ ,  $p = 0.543$  ; Em :  $\beta_{NE} = 4 \cdot 10^{-5} \pm 0.025$ ,  $p = 0.999$ ) ou le dlPFC (Ep :  $\beta_{NE} = 0.011 \pm 0.029$ ,  $p = 0.710$  ; Em :  $\beta_{NE} = 0.003 \pm 0.017$ ,  $p = 0.846$ ).



**Figure 35 : Décomposition de l'activité du striatum (en haut), du dmPFC (au milieu) et du dIPFC (en bas) au moment de l'apparition des enjeux en fonction des variables impactant la performance.** Les régions d'intérêt ont été définies à partir de la plateforme Neurosynth avec le mot « *incentive* » pour le striatum et l'expression « *cognitive control* » pour le dmPFC et le dIPFC. Le sens des abréviations est le suivant : Ep pour la tâche d'effort

physique, Em pour la tâche d'effort mental, E pour les enjeux monétaires signés, |E| pour les enjeux monétaires non-signés et NE pour le numéro d'essai. N = 20 participants. Le sens des symboles est le suivant : \*p.value<0.05 ; \*\*p.value < 0.01 ; \*\*\*p.value < 0.005 ; \*\*\*\*p.value < 0.001. Les contrastes qui ne sont pas surplombés d'une étoile n'étaient pas significatives.

## II.4 Discussion

### II.4.a Différence entre gains et pertes

Le réseau des valeurs réagit différemment en fonction de si la situation implique un montant qu'on peut gagner ou un montant qu'on risque de perdre (Bartra et al., 2013; Tom et al., 2007). La théorie des perspectives suggère que, pour une valeur monétaire équivalente, le cerveau donne plus de poids à une valeur dans le cas de pertes potentielles que dans le cas de gains potentiels (voir **Figure 2**). Ce phénomène a été identifié comme une aversion à la perte (Kahneman and Tversky, 1984). Nos résultats semblent contradictoires avec cette proposition puisque nos participants avaient plutôt tendance à afficher une meilleure performance face à des gains équivalents (voir **Figure 33**). Plusieurs résultats expérimentaux remettent d'ailleurs aussi en question l'existence d'une aversion à la perte (Ert and Erev, 2007; Gal and Rucker, 2018; Walasek and Stewart, 2015, 2019; Yechiam and Hochman, 2013). Il a par exemple été proposé que cette aversion à la perte soit principalement le reflet d'une activité accrue du système autonome et de l'allocation de plus d'attention à la tâche en cours dans le cas des pertes (Yechiam and Hochman, 2013). Une autre étude a montré que le pattern s'inverse en fonction de la magnitude des enjeux et que les participants présentent une aversion à la perte uniquement lorsque les enjeux monétaires sont très élevés (Mukherjee et al., 2017). Il est de ce fait possible que les enjeux relativement modestes que nous avons employé dans nos tâches (0.01 à 20€) aient été dans l'ordre de grandeur où le pattern était inversé. Enfin, une autre théorie propose que l'aversion à la perte n'existe pas en tant que telle, mais qu'elle dépende des expériences précédentes du sujet et du rang des valeurs disponibles (Walasek and Stewart, 2015). Dans tous les cas, l'aversion à la perte semble dépendre des conditions de l'expérience et ne pas être ancrée comme un biais universel du comportement des Mammifères comme Daniel Kahneman l'avait initialement proposé.

Il a aussi été proposé que l'aversion à la perte, dans des tâches impliquant une performance, ne s'exprime pas par une meilleure performance, mais, au contraire, par une réduction de la performance lorsque les gains potentiels sont conséquents. Cette peur de ne pas réussir à obtenir une récompense élevée, c'est-à-dire en quelque sorte de la perdre, renvoie à un phénomène rendu par l'expression « *choking under pressure* » en anglais qui exprime l'idée



## Études expérimentales

plus générale que la performance baisse sous la pression (Chib et al., 2012). Selon cette logique, les participants, dans nos tâches, verraient leur performance baisser dans le cas des pertes en comparaison aux gains, non pas parce qu'ils donneraient une valeur plus élevée aux gains, mais justement parce qu'ils donnent plus d'importance aux pertes qu'aux gains. Dans la tâche de Chib *et al.*, la performance augmentait avec les enjeux jusqu'à ce que les gains potentiels soient tellement importants que la performance, et l'activité du striatum, baissent, dans la condition difficile. En d'autres termes, la transformation des enjeux en performance n'était pas linéaire mais dépendait de la magnitude des enjeux et de la difficulté de la tâche, mais un seul système motivationnel pourrait assurer cette transformation simple des enjeux en performance. Dans notre tâche cependant, ce raisonnement implique l'existence, à la fois d'un système motivationnel capable d'augmenter linéairement la performance en fonction des enjeux non-signés, indépendamment de leur valence, et d'un autre système, sensible à l'aversion à la perte, qui ferait baisser la performance dans le cas des pertes par rapport aux gains. Cela n'est pas impossible, mais cela paraît peu parcimonieux. Même si on acceptait cette hypothèse, elle impliquerait que l'aversion à la perte, dans une tâche de performance, s'exprime, paradoxalement, à la fois par une performance moindre dans le domaine des pertes en comparaison aux gains et aussi, par une activité moindre du striatum, qui devrait pourtant signaler la valeur motivationnelle de l'action. Or cette hypothèse part justement du principe que, c'est parce que les individus donnent une importance plus grande aux pertes qu'aux gains, que la performance ait moindre dans le cas des pertes. On s'attendrait donc à ce que le striatum soit moins actif, et non pas plus actif, en réponse aux enjeux monétaires signés, pour signaler que ces enjeux sont plus importants. Or, dans notre tâche, l'activité du striatum augmentait avec les enjeux monétaires signés. Il serait cependant intéressant de voir si, comme dans l'expérience de Chib *et al.*, il existe un lien entre la sensibilité à l'aversion à la perte dans une tâche de choix et la différence de performance entre les gains et les pertes qui puisse s'expliquer par d'autres facteurs que l'aversion à la perte telle qu'elle a initialement été définie par Daniel Kahneman et Amos Tversky.

On pourrait aussi nous rétorquer qu'il est possible que nos participants aient assimilé les deux conditions dans nos tâches. Ils auraient alors dû se comporter de manière strictement équivalente entre la condition où ils jouaient pour gagner une part du montant en jeu et la condition où ils jouaient pour ne pas perdre une part du montant en jeu. Or, ce que nous avons constaté est que les participants affichaient une performance légèrement supérieure lorsqu'ils jouaient pour gagner par rapport à lorsqu'ils jouaient pour ne pas perdre de l'argent (voir **Figure**

**33).** Ce résultat va aussi de pair avec le fait que l'activité du striatum variait linéairement avec les enjeux signés (voir **Figure 35**). L'activité des aires liées au contrôle cognitif et à la détection de situations saillantes et/ou périlleuses que sont le dmPFC et le dlPFC avait aussi tendance à augmenter avec les enjeux monétaires signés (voir **Figure 35**). En d'autres termes, que ce soit au niveau comportemental ou au niveau neural, les conditions étaient bien différenciées et les participants avaient plutôt tendance à donner un poids plus conséquent aux gains qu'aux pertes contrairement à ce que l'aversion à la perte suggère.

Plusieurs études montrent, par ailleurs, que le striatum est plus actif dans le cas du traitement de gains que dans le cas du traitement de pertes (Bartra et al., 2013; Fouragnan et al., 2018; Tom et al., 2007). Ce biais a été identifié à un biais pavlovien. Le cerveau réagirait de manière linéaire en fonction des enjeux afin de promouvoir une réaction d'approche plus rapide et plus intense en réaction à des stimuli appétitifs et au contraire une réaction d'évitement plus lente en réponse à des stimuli aversifs (Boureau and Dayan, 2011; Dayan, 2008; Oudiette et al., 2019; Shadmehr et al., 2019). Plus généralement, il a été proposé que les ganglions de la base permettent la mise en place d'un comportement plus rapide et d'un effort d'intensité plus importante en réponse à des enjeux plus importants (Dudman and Krakauer, 2016; Turner and Desmurget, 2010; Turner et al., 2003; Yttri and Dudman, 2016). Dans le cas de la tâche d'effort physique, ce biais pavlovien pourrait ainsi s'exprimer par un biais à répondre avec un effort plus intense dans le cas des gains par rapport aux pertes. Dans le cas de la tâche d'effort mental, où la performance dépend à la fois de la précision et de la vitesse des réponses, il est possible que cet effet passe par des réponses plus rapides, cependant, nous avons vu que pour la première paire de chaque essai du moins, la tendance à répondre plus vite pour les enjeux signés, bien que présente, était faible (voir **Figure 31**). Une possibilité serait que les ganglions de la base favorisent ainsi l'engagement de plus de ressources neurales afin de répondre à la fois plus rapidement et plus précisément en réponse aux enjeux signés dans le cadre de notre tâche d'effort mental. L'activité du striatum, variant linéairement avec les enjeux signés, pourrait ainsi favoriser la mise en place d'une réponse comportementale engageant plus de ressources, que ce soit au niveau physique ou mental.

### II.4.b Le striatum signale les enjeux

L'activité des ganglions de la base a depuis longtemps été associée à la capacité d'effectuer des mouvements (Neafsey et al., 1978; Turner and Desmurget, 2010; Turner et al., 2003). Plusieurs études montrent en outre que l'activité des ganglions de la base détermine l'intensité d'un effort physique en fonction des enjeux (Pessiglione et al., 2007; Schmidt et al.,

2008, 2012). Il a aussi été montré que leur activité permettait aussi de déterminer l'intensité d'un effort mental en fonction des enjeux (Locke and Braver, 2008; Schmidt et al., 2012). Nous avons vu que le striatum qui fait partie des ganglions de la base s'activait proportionnellement aux enjeux non-signés et aux enjeux signés (voir **Figure 35**). De même, la performance, tant dans la tâche d'effort physique que dans la tâche d'effort mental augmentait proportionnellement aux enjeux non-signés et aux enjeux signés (voir **Figure 33**). De plus, le réseau des valeurs tient son nom de par sa corrélation récurrente avec une mesure de l'estimation de la valeur subjective (Bartra et al., 2013). Il paraît ainsi plausible que le réseau des valeurs calcule, à chaque essai, un rapport entre les bénéfices en termes de gains ou de pertes monétaires et les coûts en termes d'effort physique ou mental à implémenter. Une fois ce calcul effectué au moment de l'affichage des enjeux à l'écran, ce réseau pourrait ainsi envoyer cette valeur nette à d'autres structures cérébrales en vue de les informer sur la quantité de ressources à allouer au moment où il sera possible d'effectuer un effort.

La seule nuance à cette proposition est que, bien que nous ayons constaté que le temps, via le numéro d'essai, impactait la performance (voir **Figure 33**), nous n'avons pas constaté de lien direct entre cette mesure et l'activité des structures du réseau des valeurs (voir **Figure 35**). Une explication possible est que le réseau des valeurs ne calcule que la partie de l'équation dévolue à l'anticipation des bénéfices  $B_i$  mais pas celle liée aux coûts de l'action  $C_i$  (voir (Équation 4)).

En résumé, premièrement, nos résultats sont en accord avec plusieurs études précédentes montrant que le striatum corrèle à la fois avec la valeur subjective non-signée, une variable souvent considérée comme reflétant la saillance des enjeux, et aussi avec la valeur subjective signée (Bartra et al., 2013; Cooper and Knutson, 2008; Cooper et al., 2009; Litt et al., 2011). Ce résultat va à l'encontre de l'idée que les individus donnent plus d'importance aux pertes, ce qui devrait au contraire se refléter par une activité accrue du striatum en réponse à des pertes équivalentes. Deuxièmement, l'activité du striatum dans nos tâches, qui impliquaient d'effectuer une performance de nature physique ou mentale, pourrait influencer la vigueur des efforts effectués de sorte à faciliter une meilleure performance en fonction des enjeux.

### II.4.c Fatigue ou baisse motivationnelle

Nous avons constaté une baisse de la performance au cours des essais dans chacune de nos tâches (voir **Figure 33**).

## Études expérimentales

Dans la tâche d'effort physique, une baisse de la performance au cours du temps aurait pu s'expliquer par la manière dont notre tâche fonctionnait. En effet, dans notre tâche d'effort physique, si, au cours d'un essai, les participants effectuaient une force maximale supérieure à la force pour laquelle ils ont été calibrés au début de chaque session, le seuil de force nécessaire pour atteindre le haut de l'échelle était ajusté pour les essais suivants (voir *Spécificités de la tâche d'effort physique*). De ce fait, la tâche pouvait nécessiter d'exercer une force supérieure pour atteindre un même niveau de performance à la fin par rapport au début de l'expérience. Ainsi, pour atteindre un même niveau de performance, ils devaient, le cas échéant, allouer un effort physique plus important pour compenser la modification de l'échelle. Nous avons en effet pu observer que 22 sur les 23 participants inclus dans l'étude avaient serré la poignée de force au moins une fois dans l'une des deux sessions plus fortement que la force maximale calibrée. La force maximale utilisée lors du dernier essai était, en moyenne, égale à  $129 \pm 5$  % (moyenne  $\pm$  erreur type de la moyenne à travers les participants) à la force maximale calibrée avant le début de la tâche. Il est ainsi difficile de savoir, pour la tâche d'effort physique, si la baisse de performance est due au fait que les participants étaient de plus en plus physiquement épuisés au cours de la tâche ou au fait qu'ils ne parvenaient pas à compenser suffisamment pour le fait que la tâche leur demandait d'effectuer un effort d'intensité croissante au cours du temps. Cependant, cela n'explique pas pourquoi le même effet a pu être observé dans la tâche d'effort mental où la performance des participants diminuait aussi au cours des essais.

Alternativement, si on cherche une explication commune à nos deux tâches, on peut interpréter cette réduction de la performance au cours des essais au moins de deux façons. Premièrement, il est possible que le coût lié à l'exécution d'un effort physique ou mental augmente au cours des essais. Le fait de serrer la poignée de force implique notamment d'effectuer un travail musculaire qui nécessite la consommation de ressources physiologiques comme l'Adénosine TriPhosphate (ATP) et l'accumulation de déchets physiologiques qu'il faut éliminer comme le lactate. Dans le cas de l'effort mental, de même plusieurs théories proposent que le fait d'utiliser du contrôle cognitif implique de consommer des ressources physiologiques et/ou de produire des déchets à nettoyer (Kool and Botvinick, 2018; Shenhav et al., 2017; Zénon et al., 2019). Deuxièmement, il est aussi possible que la perception subjective des montants en jeu varie au cours des essais. Le même montant objectif serait ainsi perçu comme étant de moins en moins intéressant au fur et à mesure des essais. À la fin de chaque essai, les participants pouvaient observer la mise à jour de leur cagnotte qui représentait les gains et les pertes accumulés au cours de tous les essais précédents d'une tâche donnée. Or, nous savons que les

## Études expérimentales

individus évaluent les gains et les pertes en fonction d'un point de référence subjectif (Kahneman, 2012). Il se pourrait ainsi que, au fur et à mesure que leurs gains cumulés augmentaient, la perception relative d'un gain ou d'une perte donnée diminue. Relativement au total déjà accumulé au cours de la tâche, le même montant paraîtrait en effet alors relativement moins important à la fin de la tâche, lorsqu'un certain montant d'argent a déjà été assuré, par rapport au début, lorsque la cagnotte est encore vide.

Pour ce qui est de cette deuxième hypothèse, c'est-à-dire qu'il existe une réduction de l'appétence des récompenses du fait d'une modification du point de référence, deux résultats dans nos données viennent la contredire. Premièrement, les participants ne présentaient pas de différence entre les deux sessions d'une même tâche alors que, lors de la deuxième session, ils avaient déjà en leur possession les gains cumulés lors de la première session (voir **Figure 32**). Deuxièmement, on s'attendrait à ce que le striatum, qui reflétait l'importance des enjeux dans nos tâches, soit impacté par une telle dévaluation. Si les enjeux étaient généralement moins importants au fur et à mesure des essais, l'activité du striatum aurait aussi dû baisser au cours des essais. Or nous n'avons pas constaté de corrélation significative entre le striatum et le numéro d'essai (voir **Figure 35**). Nous n'avons pas non plus observé de corrélation significative entre les aires liées au contrôle cognitif et le numéro d'essai. Une baisse d'activité du dlPFC a pourtant déjà été signalée en relation avec la notion de fatigue cognitive (Blain et al., 2016, 2019). En abaissant les seuils de nos cartes à des p.values très tolérantes ( $p < 0.05$  non-corrigée), les seules aires que nous avons pu voir apparaître sont le putamen ventral gauche dont l'activité diminuait au cours des essais. Nous n'avons ainsi pas vraiment réussi à complètement comprendre les facteurs expliquant la baisse de performance au cours du temps, ni à en trouver des corrélats cérébraux, mais la fatigue cognitive reste une possibilité. Une hypothèse alternative serait simplement que les participants sont moins impliqués dans la tâche au niveau global mais on se serait alors aussi attendu à ce que l'activité du striatum s'en trouve altérée ce qui n'était pas le cas.

### II.4.d Effort physique et effort mental : même combat ?

Notre modèle computationnel nous a permis d'estimer la quantité de ressources investies à chaque essai au sein de chacune de nos tâches. Nous n'avons pas précisé la nature de la ressource investie dans ce modèle. Cependant, il est intéressant de constater que la tâche d'effort physique et la tâche d'effort mental partagent des corrélats cérébraux en lien avec la quantité de ressources investies. En particulier, nous avons pu identifier l'activité du cortex préfrontal dorsolatéral (dlPFC), une zone classiquement considérée comme le siège de

## Études expérimentales

l'exécution du contrôle cognitif (van Belle et al., 2014; Blain et al., 2016; Hare et al., 2009; Wager et al., 2005), en lien avec la quantité de ressources investies à chaque essai (voir **Figure 34**). De plus, nous avons aussi pu observer que l'insula antérieure et le cortex préfrontal dorsomédian (dmPFC) étaient aussi associés à la quantité de ressources allouées à chaque essai. Or, le dmPFC a été associé à l'estimation de l'avantage de recruter du contrôle cognitif via l'activité du dlPFC (Shenhav et al., 2013). Ces résultats sont ainsi compatibles avec l'idée que le dmPFC, l'insula antérieure et le dlPFC sont liés à l'implémentation de contrôle cognitif et à la sensation aversive d'effort mental qui lui est associée.

De plus, la tâche de Stroop est classiquement utilisée pour stimuler l'utilisation de contrôle cognitif dans le cas des paires incongruentes où la réponse automatique n'est pas la réponse correcte. L'utilisation de contrôle cognitif dans cette tâche requiert ainsi l'activité du dlPFC et dmPFC afin de répondre correctement dans le cas des paires incongruentes (Glascher et al., 2012; MacDonald et al., 2000; Roberts and Hall, 2008). Le rôle du dmPFC dans la tâche de Stroop numérique est cependant à nuancer par le fait que des études lésionnelles ont montré que seul le dlPFC, mais pas le dmPFC, était essentiel afin de pouvoir effectuer la tâche de Stroop de manière optimale (Glascher et al., 2012). Quoiqu'il en soit, le fait que l'activité du dmPFC et celle du dlPFC soient corrélées avec la quantité de ressources inférées par notre modèle à chaque essai paraît cohérente avec les besoins de la tâche de Stroop numérique. D'un autre côté, le fait que ces mêmes aires soient aussi activées en réponse à la quantité de ressources inférée par notre modèle dans la tâche d'effort physique, comme le montre la conjonction, peut paraître plus surprenant. Une possibilité est que la tâche d'effort physique implique aussi du contrôle cognitif. Une hypothèse alternative est peut-être que le rôle de ces aires, dans le cas de la tâche d'effort physique au moins, ne soit pas liée à l'implémentation de contrôle cognitif, mais à d'autres formes de ressources physiologiques. L'activité de l'insula antérieure et du dmPFC a par exemple aussi été associée au recrutement du système nerveux autonome, ce qui pourrait globalement faciliter l'activité motrice et mentale, notamment en accélérant le pouls cardiaque (Menon and Uddin, 2010). Cependant, cela ne suffirait pas à expliquer l'augmentation de l'activité du dlPFC en réponse aux enjeux plus importants dans cette tâche. En résumé, nous avons pu voir que l'activité des aires liées au contrôle cognitif, le dmPFC et le dlPFC, était corrélée à l'allocation de plus de ressources, inférée par notre modèle, et donc, indirectement, au fait d'effectuer une meilleure performance, tant dans le cas d'une tâche impliquant un effort de nature physique que dans le cas d'une tâche impliquant un effort de

## Études expérimentales

nature mentale. Des approches complémentaires seraient nécessaires afin de vérifier le rôle causal de ces aires associées au contrôle cognitif dans le cas de la tâche d'effort physique.

En résumé, il semblerait que les efforts physiques et les efforts mentaux partagent des ressources neurales communes dans nos tâches. Il a d'ailleurs été montré qu'une fatigue au niveau cognitif pouvait avoir un impact sur une performance physique (Marcora et al., 2009; Van Cutsem et al., 2017) et qu'un surentraînement au niveau physique pouvait réduire les capacités de contrôle cognitif (Blain et al., 2019). Ceci confirme l'idée que des réseaux cérébraux similaires pourraient être impliqués dans ces deux types d'effort et, en particulier, que le dlPFC pourrait jouer un rôle dans ces deux types de tâches. L'interaction entre le striatum, dont l'activité augmente avec les enjeux non-signés et qui reflète aussi un possible biais pavlovien, reflété dans sa corrélation avec les enjeux signés, et les aires liées au contrôle cognitif reste encore à explorer en vue de mieux comprendre s'il existe un lien causal dans un sens ou dans l'autre entre ces différentes aires.

### **II.5 Conclusion**

Notre étude nous a permis de voir que, dans nos tâches, impliquant, non pas des choix, mais d'effectuer une performance physique ou mentale, nous n'avons pas observé d'aversion à la perte. Au contraire, les participants avaient plutôt tendance à afficher une performance légèrement supérieure pour des gains équivalents. Ce biais est en accord avec l'idée d'un système pavlovien automatique facilitant des réponses comportementales plus rapides et plus vigoureuses en réponse à des stimuli perçus comme étant plus appétitifs. Nous avons trouvé la trace d'un encodage des enjeux monétaires signés et aussi des enjeux monétaires non-signés dans le striatum. Enfin, nous avons utilisé un modèle computationnel en vue d'identifier des aires cérébrales associées avec la quantité de ressources investies dans chaque essai. Ce modèle nous a permis de voir qu'en plus du striatum, des aires classiquement associées à l'exécution de contrôle cognitif semblaient refléter la quantité de ressources investies à chaque essai et ce, aussi bien dans notre tâche d'effort physique que dans notre tâche d'effort mental.

## **III L'apprentissage par renforcement implique de l'effort mental**

### **III.1 Introduction**

L'apprentissage par renforcement est un mécanisme relativement universel. Il consiste dans l'idée simple qu'un évènement immédiatement suivi d'une récompense sera associé à l'obtention de cette récompense et qu'un évènement immédiatement suivi d'une punition sera

## Études expérimentales

associé au fait de subir cette punition. L'universalité de ce mode d'apprentissage dans le règne animal, des pigeons aux Primates, explique que ce mode d'apprentissage soit rattaché à des structures « primaires » qu'on retrouve chez les plupart des Vertébrés, comme les ganglions de la base et les neurones dopaminergiques (Averbeck and Costa, 2017; Grillner and Robertson, 2016; Mannella et al., 2013). De ce fait, on pourrait aussi s'attendre à ce que ce mode d'apprentissage, ayant prouvé son efficacité au fil de l'évolution, soit invariant à travers la plupart des espèces chez lesquelles on peut l'observer.

À l'inverse, la capacité d'exercer du contrôle cognitif est restreinte à quelques espèces. Cette capacité consiste plus ou moins à être capable d'établir une stratégie d'actions, notamment dans des situations où des distractions ou des automatismes doivent être inhibés en vue d'atteindre un but désiré. Le contrôle cognitif fait partie des capacités qui ont été massivement développées chez les Primates, et l'homme en particulier, grâce à l'expansion du cortex préfrontal (Carlén, 2017). L'apprentissage par renforcement, dans sa version la plus simple, sans changement dynamique des probabilités associées à chaque événement, peut être vu comme un processus automatique d'apprentissage par essai-erreur où les valeurs associées à une action donnée sont automatiquement mises à jour en fonction des feedbacks associés à chaque action. De ce fait, l'apprentissage par renforcement ne nécessiterait pas l'implémentation de contrôle cognitif. D'un autre côté, certaines versions de l'apprentissage par renforcement partent du principe que, dans le cadre d'une situation nouvelle ou lorsque l'incertitude est élevée, l'apprentissage par renforcement va de pair avec l'allocation de ressources mentales, jusqu'à ce que, à force de répétitions, le processus devienne éventuellement automatisé. John M. Pearce et Geoffrey Hall, dans le cadre du conditionnement classique, proposent par exemple que, tant qu'un stimulus n'est pas complètement prédictif, il implique l'utilisation d'un processeur cognitif qui permette de comparer les options à disposition et de délibérer (Pearce and Hall, 1980).

Plus généralement, la détection du fait d'avoir commis une erreur entraîne une augmentation de l'activité de l'insula antérieure et du cortex préfrontal dorsomédian (dmPFC) (Bastin et al., 2016; Iannaccone et al., 2015). Cette activité provoque ensuite un phénomène classiquement dénommé comme un ralentissement post-erreur (*post-error slowing*). Ce ralentissement pourrait notamment refléter le fait que les individus utilisent plus de contrôle cognitif après avoir identifié une situation où ils ont commis une erreur en vue d'adapter leur performance future. Dans des tâches d'apprentissage par renforcement, le dmPFC pourrait être capable de moduler, de manière dynamique, le taux d'apprentissage  $\alpha$  en fonction du degré de



## Études expérimentales

volatilité de la tâche (Behrens et al., 2007; Khamassi et al., 2013). L'activité du dmPFC a par ailleurs été corrélée à l'allocation de contrôle cognitif (Botvinick et al., 2001; Shenhav et al., 2013). De plus, des temps de réaction plus longs lors d'une décision sont généralement associés à l'idée que ces décisions sont plus difficiles et nécessitent plus d'effort mental. Dans notre première étude, nous avons d'ailleurs pu mettre en évidence que le temps de délibération était corrélé à l'activité du dmPFC dans des tâches nécessitant de prendre des décisions basées sur la valeur (voir *La valeur, la confiance et le temps de délibération au moment du compromis coûts/bénéfices ont des supports neuraux distincts dans le cortex préfrontal médian*). Nous avons ainsi dissocié, dans le cadre d'une prise de décision, les corrélats cérébraux de la valeur subjective, situés en partie dans le vmPFC, et les corrélats de l'effort lié à la délibération, situé dans le dmPFC et l'insula antérieure.

Dans cette étude, nous avons utilisé une tâche d'apprentissage par renforcement requérant des participants d'apprendre, par essai-erreur, à identifier les symboles leur permettant de maximiser leurs gains et de minimiser leurs pertes. Nous voulions voir si cette tâche impliquait, ou non, l'activité de structures liées à l'allocation de contrôle cognitif. L'apprentissage par renforcement, dans un cas simple qui n'incluait pas de changement dynamique des probabilités, pourrait avoir lieu de manière automatique sans nécessiter l'activité de structures liées à l'allocation de contrôle cognitif. Nous nous attendrions alors que les structures du réseau des valeurs, comme le vmPFC identifié dans notre première étude, permettent de représenter la valeur subjective associée aux différents symboles de la tâche et de la mettre automatiquement à jour. L'activité du dmPFC, que nous avons liée à l'effort implémenté au cours de la délibération dans une tâche impliquant une prise de décision (voir **Figure 27**), pourrait ne pas être impliqué dans cette tâche si elle se fait de manière automatique. A l'inverse, si cette tâche implique un effort au moment où les participants délibèrent, la tâche pourrait impliquer l'activité du dmPFC. Comme dans notre première étude, nous avons utilisé le temps de délibération comme un reflet des ressources éventuellement engagées au cours de la délibération. Nous avons ensuite regardé quels étaient les corrélats cérébraux de ce temps de délibération pour voir si le dmPFC était ou non impliqué au moment de la délibération suggérant que la tâche impliquait un certain effort mental.

## III.2 Méthodes

### III.2.a Participants

Les participants qui ont participé à cette étude étaient les mêmes que ceux recrutés pour la deuxième étude (voir *Participants*). Ils ont effectué, au cours de la même journée les tâches d'effort physique et d'effort mental détaillées dans la deuxième partie et la tâche d'apprentissage détaillée ici.

Nous avons aussi exclu le participant qui avait des résultats déplorables dans la tâche d'effort physique de l'analyse de la tâche d'apprentissage. Sa performance était en effet, là aussi, déplorable suggérant qu'il n'avait soit rien compris aux tâches que nous lui avions demandé d'exécuter soit qu'il n'avait pas la volonté de les effectuer correctement.

Pour les données d'IRMf, nous avons exclu 1 sujet en plus car il bougeait trop dans le scanner (mouvement supérieur à 5 mm). Ce sujet a cependant été inclus dans les études de comportement. Les données d'imagerie et du diamètre pupille sont donc basées sur 22 participants au total (12 femmes, 10 hommes) de  $26 \pm 4$  ans (moyenne  $\pm$  écart-type).

### III.2.b Tâche

La tâche a été programmée dans la toolbox Psychtoolbox (<http://psychtoolbox.org/>) au sein du logiciel Matlab 2012 (The MathWorks, Inc., USA). Les participants ont reçu des instructions à l'écrit et à l'oral. Le principe de l'expérience tel qu'il était indiqué aux participants était de maximiser leurs gains.

Nous avons utilisé une version similaire à celle déjà été utilisée dans notre équipe lors d'études précédentes (Palmineri et al., 2012; Pessiglione et al., 2006). Dans les instructions, nous disions aux participants que, pour une session donnée, ils allaient être confrontés à 6 symboles différents. Ils devaient chercher à identifier quels étaient les symboles qui leur permettraient de maximiser leurs gains. Nous leur disions que certains symboles étaient principalement associés à des gains de 10€, d'autres à rien (0€) et d'autres à des pertes de 10€. Ils devaient ainsi découvrir, par essai-erreur, quels étaient les symboles à sélectionner à chaque essai afin de maximiser leurs gains. Nous leur précisions que le sens associé à chaque symbole ne variait pas au cours de l'expérience, mais qu'un symbole donné ne permettait pas toujours d'obtenir le résultat escompté. Le but de ces instructions était d'essayer d'expliquer aux participants que les symboles étaient des prédicteurs probabilistes, mais que les probabilités qui y étaient associées étaient fixes pour éviter qu'ils cherchent à estimer la volatilité de la tâche.

## Études expérimentales

Nous voulions ainsi être sûrs que les participants avaient bien compris la structure de la tâche et qu'ils ne seraient pas en train d'apprendre la structure de la tâche, en plus d'apprendre les valeurs associées à chaque symbole.

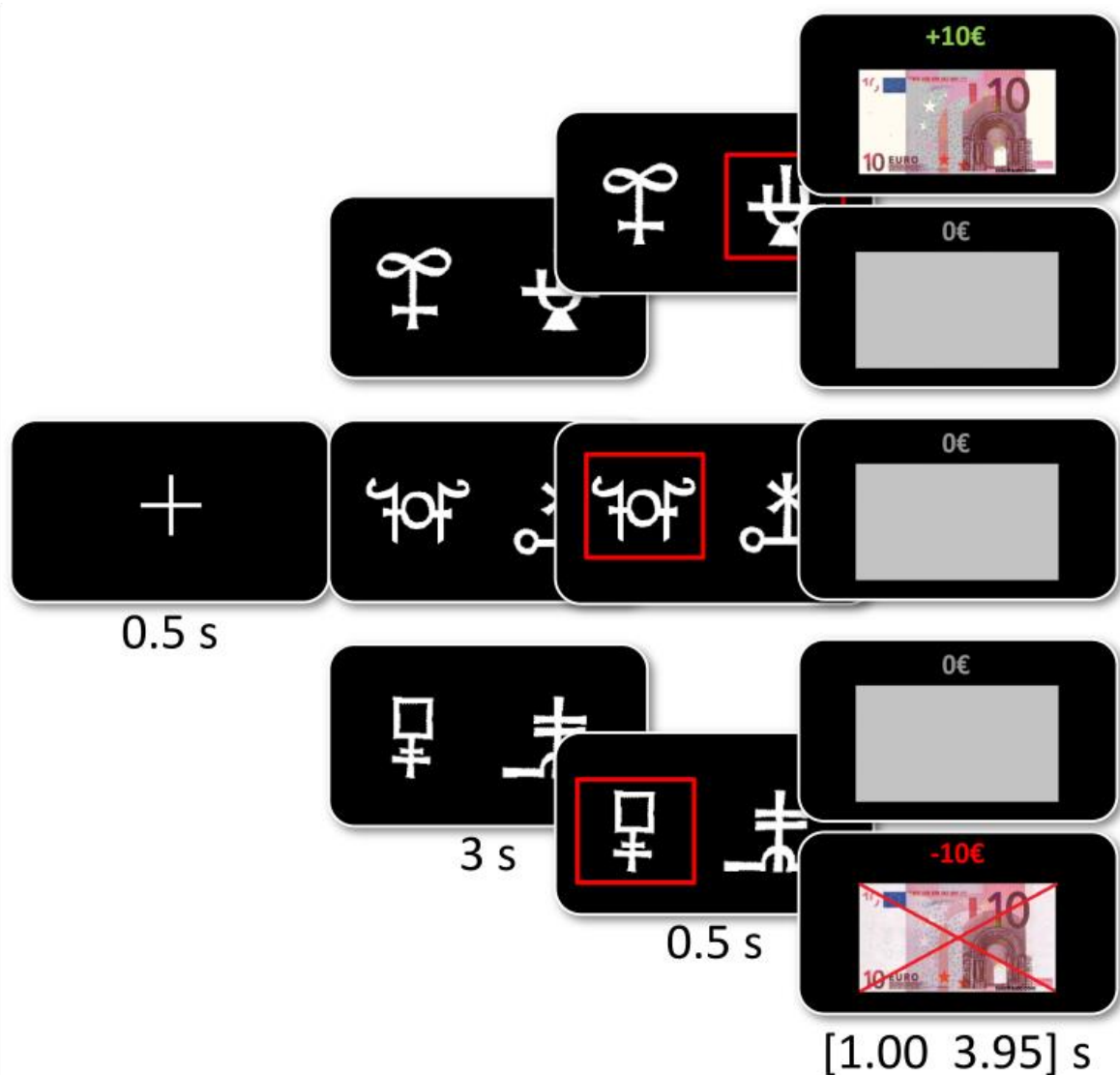
Cela n'était pas précisé dans les instructions, mais chaque symbole était systématiquement présenté au sein d'une même paire. Il y avait 3 types de paires : des paires associées à des gains (+10€ ou 0€), des paires associées à un résultat neutre (0€) et des paires associées à des pertes (-10€ ou 0€). Au sein de chaque paire, les probabilités des symboles étaient associées (75/25% et 25/75%). En d'autres termes, quand, dans une paire associée à des gains, le symbole A permettait d'obtenir un gain (+10€), le symbole B de la même paire était nécessairement associé à un résultat neutre (0€) pour l'essai concerné.

Chaque essai démarrait avec la présentation d'une croix de fixation pendant une durée de 500 millisecondes. Ensuite, les deux symboles d'une même paire apparaissaient à l'écran. Pour répondre, le sujet devait appuyer sur le bouton correspondant à la localisation spatiale du symbole qu'il voulait sélectionner (gauche/droite) et garder le doigt appuyé sur ce bouton jusqu'à la fin de la période de sélection. Pour répondre, les participants disposaient d'un boîtier de réponse (fORP 932, Current Designs Inc, Philadelphia, USA) placé sous leur main droite au niveau du torse au début de l'expérience. Leur index était posé sur le bouton permettant de sélectionner l'option de gauche et leur majeur sur le bouton permettant de sélectionner l'option de droite. Une fois cette période terminée, le symbole choisi apparaissait entouré d'un cadre rouge pendant 500 millisecondes. Enfin, le résultat obtenu suite à la sélection du symbole était affiché à l'écran (voir **Figure 36**). Si le participant n'avait appuyé sur aucun bouton pendant la période de sélection, s'il avait cessé d'appuyer avant la fin de la période de sélection ou s'il avait appuyé sur les deux boutons simultanément, nous considérons cela comme une omission de la part du sujet. Il recevait alors un message précisant l'erreur commise (« Réveillez-vous ! », « Maintenez votre appui » ou « Un seul bouton ! ») et, au moment du feedback, il obtenait le pire résultat possible pour la paire considérée. Pour les paires associées à des gains ou à un résultat neutre, il n'obtenait donc rien (0€) et pour les paires associées à des pertes, ils subissaient une perte de 10€. Ceci était effectué afin de motiver les participants à répondre à tous les essais, notamment dans le cas des paires associées à des pertes.

Chaque session de la tâche d'apprentissage était composée de 60 essais. Les symboles étaient des stimuli abstraits tirés de l'alphabet Agathodaimon. Les participants ont dû effectuer 4 sessions de cette tâche. La première session était effectuée sur un ordinateur portable en dehors du scanner. Elle servait à familiariser les participants avec la tâche avant l'expérience.

## Études expérimentales

Pour répondre, les participants utilisaient alors les flèches gauche et droite du clavier de l'ordinateur à la place du boîtier de réponse. Les trois sessions suivantes étaient effectuées dans le scanner. Entre chaque session, les participants effectuaient aussi une session de la tâche d'effort physique et une session de la tâche d'effort mental (voir *L'effort, tant physique que mental, est plus motivé par l'attrait de gagner que par la peur de perdre*). Les 6 symboles utilisés dans chaque session étaient différents pour chaque session. Pour effectuer ces 4 sessions, nous avons donc sélectionné 24 symboles identiques pour tous les participants. Pour l'entraînement, nous utilisons toujours les mêmes symboles pour tous les participants. Pour les sessions en IRMf cependant, le sens associé à chaque symbole était contrebalancé entre les participants. Nous voulions ainsi éviter que des propriétés visuelles liées aux symboles puissent éventuellement créer des biais dans l'apprentissage. Au sein d'une paire, la position de chaque symbole sur l'écran (gauche/droite) variait aléatoirement d'un essai à l'autre. Cela nous permettait de nous assurer que les participants apprenaient bien à associer des valeurs aux symboles et pas à leur localisation spatiale. Les symboles associés aux gains ainsi que ceux associés aux pertes apparaissaient dans 24 essais chacun. Les symboles toujours associés au résultat neutre n'apparaissaient, eux, que dans 12 essais. Nous avons réparti les essais en 12 mini-blocs de 5 essais. Chaque mini-bloc était constitué de 2 essais avec la paire associée à des gains, 2 essais avec la paire associée à des pertes et 1 essai avec la paire associée à un résultat neutre. L'ordre de présentation des essais au sein d'un mini-bloc était aléatoire. La structure en mini-blocs n'était pas connue des participants mais elle nous permettait de nous assurer que l'ordre de présentation des différents types de paires et de symboles était à peu près équivalent.



**Figure 36 : Tâche d'apprentissage.** Un essai démarrait toujours avec la croix de fixation affichée pendant 500 millisecondes. Ensuite, les deux symboles d'une paire apparaissaient à l'écran. Le participant devait sélectionner un des deux symboles à l'aide d'un boîtier de réponse où deux boutons correspondaient à la localisation spatiale (gauche/droite) de chaque option. Cette période de sélection durait 3 secondes. Le participant devait garder le doigt appuyé sur le bouton de l'option choisie, sans appuyer sur l'autre bouton, jusqu'à la fin de la période de sélection pour qu'une réponse soit considérée comme valide. Lorsque cette période était terminée, l'option sélectionnée, le cas échéant, était entourée d'un cadre rouge. Cette période durait 500 millisecondes. Si la réponse n'était pas valide, alors un message apparaissait à l'écran indiquant aux participants pourquoi la réponse n'était pas considérée comme valide et ils obtenaient le pire résultat possible pour la paire en question. Dans le cas de la paire de symboles associées à des gains (haut), l'un des deux symboles permettait d'obtenir une récompense de +10€ dans 75% des essais et rien (0€) dans les 25% d'essais restants, alors que l'autre symbole avait les probabilités inverses. Pour la paire de symboles associés à un résultat neutre (milieu), quelle que soit l'option choisie, le résultat était systématiquement 0€. Enfin, dans le cas de la paire associée à des pertes (bas), l'un des deux symboles faisait perdre 10€ dans 75% des essais et ne provoquait rien (0€) dans les 25% restants, alors que l'autre faisait de même mais avec les probabilités inverses. Le résultat final était affiché pour une durée variant aléatoirement entre

## Études expérimentales

1.00 et 3.95 secondes. Une fois cette période terminée, un nouvel essai démarrait avec à nouveau l’affichage de la croix de fixation.

La session d’entraînement avant de rentrer dans le scanner contenait 60 essais. A la fin de cette session, nous vérifions que le participant sélectionnait bien le meilleur symbole possible dans plus de 50% des essais pour les paires associées à des gains et les paires associées à des pertes. Si ce n’était pas le cas, nous demandions aux participants de refaire une session complète, avec les mêmes symboles associés aux mêmes résultats, jusqu’à ce que nous soyons assurés qu’ils avaient bien compris la structure de la tâche.

À la fin de chaque session d’apprentissage, les participants pouvaient voir le total des gains accumulés sur l’ensemble des sessions d’apprentissage effectuées dans le scanner. Ce total était indépendant des deux autres tâches effectuées.

### III.2.c Analyse comportementale

- **Modèle computationnel**

Nous avons modélisé le comportement de chaque individu séparément en utilisant un modèle de « Q-learning » standard (Watkins and Dayan, 1992) à l’aide de la VBA\_toolbox (Daunizeau et al., 2014). Le modèle a été inversé en utilisant l’approche variationnelle bayésienne avec une approximation de Laplace. Ce modèle dit de « Q-learning » est utilisé de longue date pour modéliser l’apprentissage par renforcement et rend bien compte des choix opérés par les participants dans des tâches de ce type (FitzGerald et al., 2012; Palminteri et al., 2009; Pessiglione et al., 2006). Le principe du modèle est que chaque option de chaque paire est associée à un état caché, une valeur  $Q$ , et que cet état évolue au fur et à mesure de la tâche en fonction de la séquence de choix et de feedbacks. Ces états cachés représentent la récompense (ou la punition) attendue du fait de sélectionner l’option en question. Toutes les valeurs  $Q$  initiales ont été définies avec une valeur de zéro. Nous partions ainsi du principe que comme les participants ne savaient pas quelle est la valeur associée à chaque option au début de chaque session, ils démarraient avec une valeur attendue nulle pour chaque option. Par la suite, la valeur associée à chaque option évolue, au sein d’une fonction d’évolution, selon la formule :

$$Q_{choisie}(t + 1) = Q_{choisie}(t) + \alpha * PE(t)$$

(Équation 10)

$Q_{choisie}(t+1)$  est la valeur de l’option choisie à l’essai  $t$  après l’avoir mise à jour au cours de l’essai,  $Q_{choisie}(t)$  est la valeur attendue de l’option choisie à l’essai  $t$ ,  $\alpha$  est le taux

## Études expérimentales

d'apprentissage et  $PE(t)$  est l'erreur de prédiction.  $PE(t)$  se décompose selon la formule  $PE(t) = feedback(t) - Q_{choisie}(t)$  où  $feedback(t)$  est une variable égale à 1 lorsque le feedback obtenu est un gain (+10€), à 0 lorsque le feedback obtenu est neutre (0€) et à -1 lorsque le feedback obtenu est une perte (-10€). Nous avons aussi défini le modèle selon le raisonnement contrefactuel, de sorte que les valeurs des deux options affichées soient mises à jour à la suite d'un essai, partant du principe que les participants ont compris, au cours de la session d'entraînement que les deux symboles d'une même paire étaient associés. La valeur de l'option non-choisie évoluait ainsi selon la formule :

$$Q_{non-choisie}(t+1) = Q_{non-choisie}(t) + \alpha * PE_{bis}(t)$$

(Équation 11)

$Q_{non-choisie}(t+1)$  est la valeur de l'option non-choisie à l'essai  $t$  après l'avoir mise à jour au cours de l'essai,  $Q_{non-choisie}(t)$  est la valeur attendue de l'option non-choisie à l'essai  $t$ ,  $\alpha$  est le taux d'apprentissage estimé par le modèle et  $PE_{bis}(t)$  est l'erreur de prédiction pour la valeur de l'option non-choisie.  $PE_{bis}(t)$  est basée sur l'idée que les participants mettaient à jour la valeur de l'option non-choisie comme s'ils avaient aussi pu observer le feedback associé à l'option non-choisie de sorte que  $PE_{bis}(t) = feedback_{bis}(t) - Q_{non-choisie}(t)$  où  $feedback_{bis}(t)$  vaut 1 lorsque l'option non-choisie était associée à un gain (+10€), à 0 lorsque elle était associée à un feedback neutre (0€) et à -1 lorsqu'elle était associée à une perte (-10€). Bien évidemment, dans le cas des paires neutres, aucun apprentissage ne pouvait être modélisé étant donné que les feedbacks obtenus étaient toujours de 0€.

Ce modèle nous a aussi permis de modéliser les choix effectués par les participants au sein d'une fonction d'observation qui permet de dériver la probabilité, au sein d'une paire, de choisir la meilleure des deux options selon la formule :

$$p(\text{choix} = \text{meilleure option}) = \frac{1}{1 + \exp\left(-\frac{DV}{\beta}\right)}$$

(Équation 12)

Où  $\beta$  est la température estimée par le modèle qui capte le fait que les participants répondent plus ou moins aléatoirement,  $DV$  est la différence entre l'état caché associé à la meilleure des deux options et l'état caché associé à la moins bonne des deux options.  $\beta$  était contraint à être positif dans le modèle.

## Études expérimentales

Nous avons utilisé les mêmes paramètres  $\alpha$  et  $\beta$  pour les symboles associés aux gains et les symboles associés aux pertes. Nous avons aussi utilisé les mêmes paramètres à travers les 3 sessions d'un individu donné. Le taux d'apprentissage  $\alpha$  permet de modéliser le fait que chaque individu pourrait accorder un poids différent aux feedbacks obtenus. La VBA\_toolbox permet de trouver, pour un participant donné, le set de paramètres  $\alpha$  et  $\beta$  qui maximise l'énergie libre du modèle selon l'approche variationnelle bayésienne.

- **Analyse des temps de réaction**

Nous avons voulu voir si les variables dérivées du modèle permettaient d'expliquer les temps de réaction. Pour la mesure du temps de réaction, nous avons pris en compte le temps mis entre l'apparition des stimuli à l'écran et le moment du premier appui sur une des touches de réponse. Pour l'analyse, nous avons d'abord regroupé les données à travers les trois sessions. Ensuite, nous avons regroupé les essais avec une paire associée à des gains avec les essais contenant une paire associée à des pertes, en laissant les essais avec une paire associée à un résultat toujours nul (paires neutres) à part. Nous avons ensuite effectué un GLM pour voir quelles variables expliquent le temps de réaction en incluant les régresseurs suivants :

- 1) une constante pour chaque session, au cas où il existe des différences d'une session à l'autre ;
- 2) Val : la somme des valeurs des deux options présentées à chaque essai ;
- 3) Conf : une mesure indirecte de la confiance dans la décision estimée selon la formule  $[p(\text{gauche}) - 0.5]^2$  où la probabilité  $p(\text{gauche})$  était la probabilité de choisir l'option située à gauche de l'écran qui a été dérivée du modèle d'apprentissage détaillé précédemment.

Ensuite, nous avons moyenné les betas issus de ce GLM à travers les participants. Puis, nous avons testé leur significativité à l'aide d'un t.test en les comparant à zéro.

### III.2.d IRM

- **Acquisition et pré-traitement**

Les sessions de cette tâche ayant été acquises en même temps que la tâche d'effort physique et la tâche d'effort mental, le lecteur est invité à se reporter à cette partie pour voir les paramètres d'acquisition des données (voir *Acquisition*). De plus, nous avons prétraité les données conjointement pour les 3 tâches et avons ainsi aussi utilisé les mêmes paramètres pour le pré-traitement que pour ces tâches (voir *Pré-traitement*).



- **Analyse des données d'IRMf : les modèles utilisés**

Après le prétraitement, nous avons analysé les données à l'aide de modèles linéaires généralisés (*general linear model*, GLM) dans SPM12 au premier niveau pour chaque sujet. Pour chaque sujet et chaque session, 6 paramètres de mouvement étaient directement estimés par SPM12 et rajoutés au sein du GLM. Ensuite, nous avons effectué une analyse au second niveau en regardant la significativité des paramètres issus du premier niveau au niveau du groupe.

Dans le GLM principal (**GLM1**) que nous avons effectué, nous avons modélisé le moment de l'affichage des symboles associés aux paires de gains et de pertes ensemble avec une fonction boxcar enveloppant la durée de l'affichage (3s). Cette fonction était modulée avec les régresseurs suivants, rentrés dans cet ordre : 1) Val, 2) Conf, 3) DT. Nous avons aussi modélisé séparément les paires neutres avec une fonction boxcar enveloppant la durée de l'affichage des symboles à l'écran (3s). Cet événement était modulé uniquement par le temps de délibération à chaque essai. Le modèle incluait aussi un événement pour le moment où le stimulus choisi apparaissait en rouge à l'écran. Cet événement était modélisé avec une fonction boxcar enveloppant la durée d'affichage de la réponse (0.5s). Enfin, nous avons modélisé le moment d'affichage du feedback pour les paires gains et les paires pertes ensemble avec une fonction delta de durée nulle. Cette fonction était modulée par 2 régresseurs rentrés dans cet ordre : 1) l'erreur de prédiction non-signée pour l'option choisie, 2) l'erreur de prédiction signée pour l'option choisie. Enfin nous avons aussi modélisé le moment où les participants voyaient le feedback des paires neutres avec une fonction delta de durée nulle mais sans modulateur paramétrique.

Les régresseurs étaient orthogonalisés par SPM de manière sérielle au sein de chaque bloc, c'est-à-dire que le temps de délibération DT était orthogonalisé par rapport aux régresseurs Val et Conf et l'erreur de prédiction signée était orthogonalisée par rapport à l'erreur de prédiction non-signée.

Nous avons effectué un deuxième GLM, le **GLM2**, dont le but était de voir dans quelle mesure les corrélations observées dans le GLM1 étaient, ou non, impactées par l'ordre des régresseurs dans le GLM1. Ce GLM était ainsi strictement identique au GLM1 à part pour l'ordre des modulateurs paramétriques modulant l'apparition des symboles à l'écran dont l'ordre était inversé de la manière suivante : 1) DT, 2) Conf et 3) Val. Dans ce GLM aussi, toutes les variables étaient orthogonalisées de manière sérielle.

Nous avons aussi effectué un autre GLM, le **GLM3**, dont le but était d'extraire l'activité de nos régions d'intérêt pour chaque essai. Dans cette optique, ce modèle n'incluait aucun modulateur paramétrique. Nous avons modélisé indépendamment chaque apparition de stimuli à l'écran, chaque apparition de l'option choisie avec un cadre rouge et chaque apparition du feedback à l'écran. Tous ces événements étaient modélisés avec une fonction delta, c'est-à-dire avec une durée nulle. Nous pouvions ainsi obtenir l'activité moyenne d'une région d'intérêt donnée pour chaque période de chaque essai.

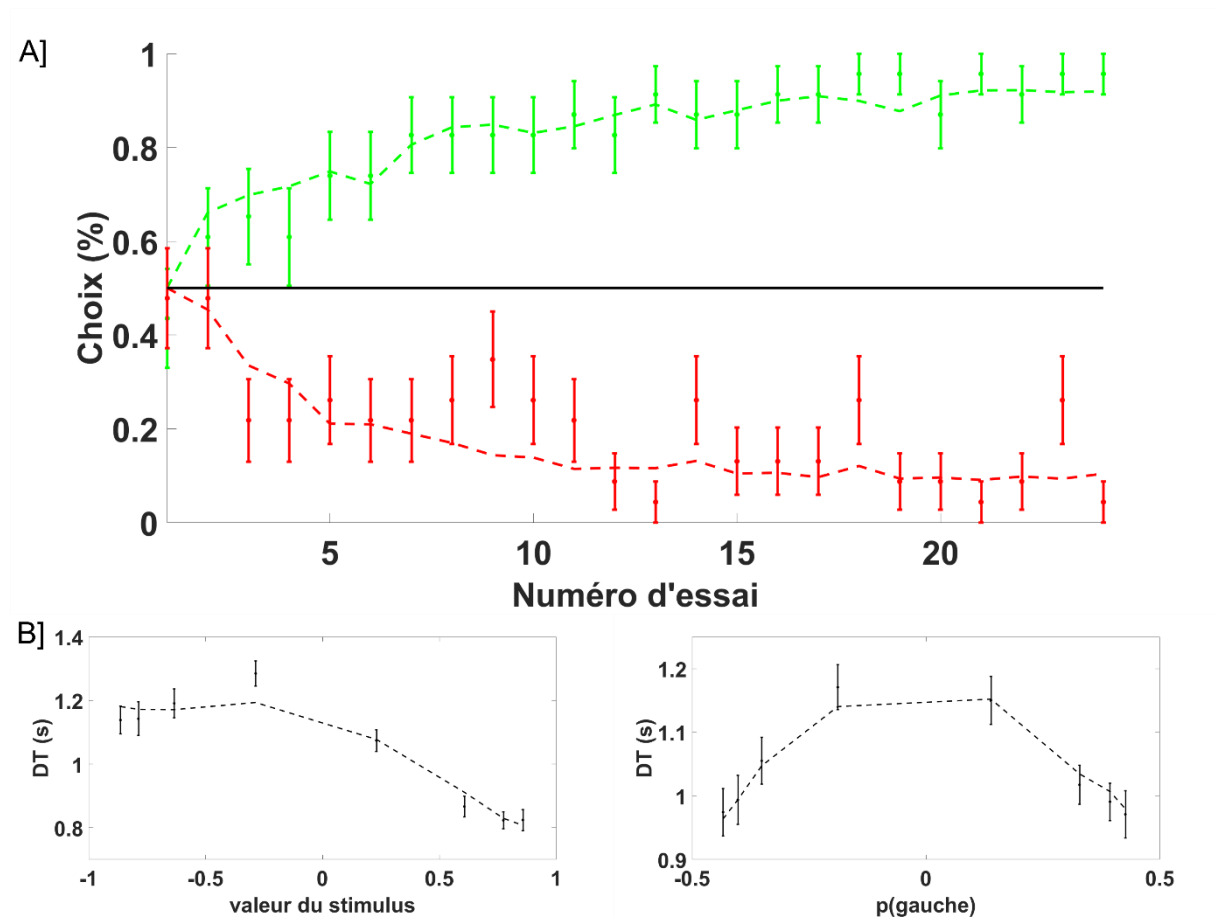
- **Régions d'intérêt**

Nous avons établi nos régions d'intérêt sur la base des corrélats cérébraux de Val et DT dans notre première étude (voir **Figure 27**). Le vmPFC a ainsi été défini comme une sphère avec un rayon de 8 millimètres autour des coordonnées MNI (-10 ; 48 ; -12) et le dmPFC comme une sphère avec un rayon de 8 millimètres autour des coordonnées MNI (10 ; 12 ; 48).

### III.3 Résultats

#### III.3.a Le comportement

Les participants ont effectué une tâche d'apprentissage par renforcement mêlant à la fois des symboles associés à des gains, des symboles neutres et des symboles associés à des pertes. Ils devaient essayer, à chaque essai, de choisir l'option qui leur permettrait de maximiser leurs gains et de minimiser leurs pertes. Comme on peut le voir sur la figure (voir **Figure 37**), les participants ont appris à sélectionner plus souvent le symbole qui leur permettait de maximiser leurs gains dans le cas des paires associées à des gains. Ils ont aussi appris à sélectionner plus souvent le symbole qui leur permettait de minimiser leurs pertes dans le cas des paires associées à des pertes (voir **Figure 37**). Le taux d'apprentissage  $\alpha$  estimé par le modèle était de  $0.12 \pm 0.07$  (moyenne  $\pm$  écart-type). La température moyenne  $\beta$  estimée par le modèle était de  $0.17 \pm 0.09$ . Le  $R^2$  moyen du modèle était de  $0.29 \pm 0.15$  (moyenne  $\pm$  écart-type) et son AIC de  $-73.49 \pm 20.66$  (moyenne  $\pm$  écart-type). En résumé, les participants, à l'exception de celui que nous avons exclu, ont bien compris la tâche. Ils ont réussi à apprendre comment maximiser leurs gains et minimiser leurs pertes en sélectionnant, dans chaque paire, le meilleur symbole possible. De plus, le modèle que nous avons utilisé semble relativement bien rendre compte du comportement des participants.



**Figure 37 : Le comportement dans la tâche d'apprentissage.**

A) **L'apprentissage au cours du temps.** Au sein de chaque session, chaque paire apparaît 24 fois. Les points de données en vert montrent le taux moyen, à travers les sessions et les participants, où l'option associée à des gains dans 75% des occurrences a été choisie pour chaque occurrence de la paire gain à l'écran. Les traits en pointillés vert montrent le résultat moyen à travers les sessions et les participants de l'estimation du modèle pour la probabilité de choisir la meilleure des deux options de la paire associée à des gains pour chaque occurrence. Les points de données en rouge montrent le taux moyen, à travers les sessions et les participants, où l'option associée à des pertes dans 75% des essais a été choisie pour chaque occurrence de la paire perte à l'écran. Les traits en pointillés rouge montrent le résultat moyen à travers les sessions et les participants de l'estimation du modèle pour la probabilité de choisir la pire des deux options de la paire associée à des pertes pour chaque occurrence. Les barres d'erreurs représentent l'erreur type de la moyenne à travers les participants dans les deux conditions. Cette figure a été établie sur la base des 23 participants inclus dans l'étude.

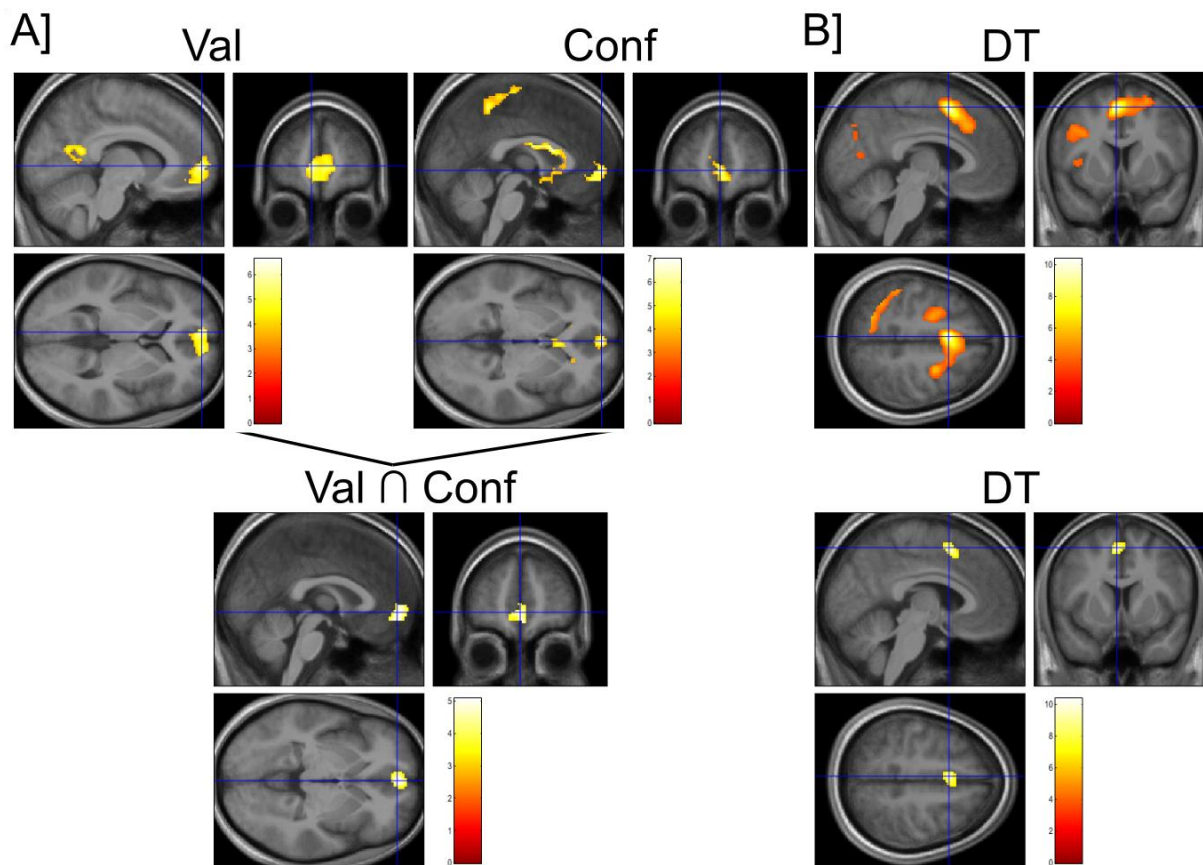
B) **Temps de délibération (DT) en fonction de la somme des valeurs dérivées du modèle computationnel, à gauche, et en fonction de la probabilité de choisir l'option de gauche centrée, à droite.** Les deux graphiques regroupent les données issues des 23 participants de l'étude pour les essais avec des paires associées à des gains ou à des pertes au cours des 3 sessions d'apprentissage. Les points avec les marges d'erreur représentent la moyenne ainsi que l'erreur type de la moyenne des temps de délibération divisés en 8 bins chez chaque participant en fonction de la somme des valeurs des stimuli à gauche et de la probabilité de choisir l'option de gauche centrée à droite. Les traits en pointillés représentent la moyenne à travers les participants, pour chaque bin, de la prédiction estimée par le modèle.

## Études expérimentales

Nous avons ensuite effectué une régression linéaire sur le temps de délibération (DT) pour voir s'il variait avec la valeur subjective (Val) et la confiance dans la réponse (Conf). Nous avons constaté que plus la valeur des options inférées par le modèle était élevée, plus le temps de délibération était rapide ( $\beta = -0,221 \pm 0.019$  ;  $p.value = 1 \cdot 10^{-10}$ ). Le DT était aussi d'autant plus rapide que Conf était élevée ( $\beta = -1.195 \pm 0.284$  ;  $p.value = 4 \cdot 10^{-4}$ ). Le DT variait donc avec la valeur du choix et avec la confiance (voir **Figure 37**).

### III.3.b Résultats neuraux

Dans un premier temps, nous avons voulu vérifier que nous arrivions à répliquer les résultats de notre première étude en reproduisant la triple dissociation, dans le cortex préfrontal médian, entre Val, Conf et le DT (voir **Figure 27**). Nous avons constaté que, dans cette tâche, Val et Conf partageaient les mêmes corrélats cérébraux. Nous avons en effet identifié un cluster dans le vmPFC avec un cluster de 747 voxels dont le pic était localisé aux coordonnées (-8 ; 56 ; 2) pour Val et un cluster de 299 voxels dont le pic était localisé aux coordonnées (0 ; 56 ; -2) pour Conf. Nous avons d'ailleurs confirmé ce résultat avec une conjonction entre ces deux contrastes (voir **Figure 38A**) qui ne révélait qu'un cluster unique de 289 voxels dont le pic était situé aux coordonnées (0 ; 52 ; -4). Au seuil utilisé, le seul autre cluster significatif qui corrélait avec Val était un cluster de 476 voxels localisé dans le cortex cingulaire postérieur avec un pic aux coordonnées MNI (-4 ; -54 ; 16). Cette aire est régulièrement associée au vmPFC en lien avec l'encodage de la valeur subjective. Pour Conf, au seuil utilisé, seuls deux autres clusters apparaissaient, dont un situé à la frontière entre les ventricules et la matière blanche, qui constituait probablement un artefact, et un autre de 440 voxels dans le précuneus, avec un pic aux coordonnées (2 ; -46 ; 54). Le DT était principalement corrélé à un cluster plus dorsal localisé dans le cortex préfrontal dorsomédian (dmPFC) (voir **Figure 38B**), malgré le fait que le DT était orthogonalisé par rapport à Val et Conf. La carte d'activation en lien avec le DT à  $p < 0.05$  corrigé au niveau des voxels pour les comparaisons multiples révélait aussi plusieurs autres clusters dont un cluster bilatéral dans l'insula antérieure avec un cluster de 37 voxels dans l'insula droite aux coordonnées (32 ; 22 ; 4) et un cluster de 3 voxels dans l'insula gauche aux coordonnées (-34 ; 14 ; 6). Contrairement aux clusters associés à Val et Conf, ce cluster survivait même pour un seuil avec  $p.value < 0.05$  corrigée au niveau des voxels (voir **Figure 38B** bas).



**Figure 38 : Les corrélats cérébraux de Val, Conf et DT dans la tâche d'apprentissage.** Les cartes d'activation sont superposées à la moyenne des scans anatomiques des 22 cerveaux inclus dans l'étude après le pré-traitement, en particulier après normalisation dans l'espace MNI. Les barres colorées montrent la correspondance entre les couleurs des cartes et les valeurs T pour chaque voxel.

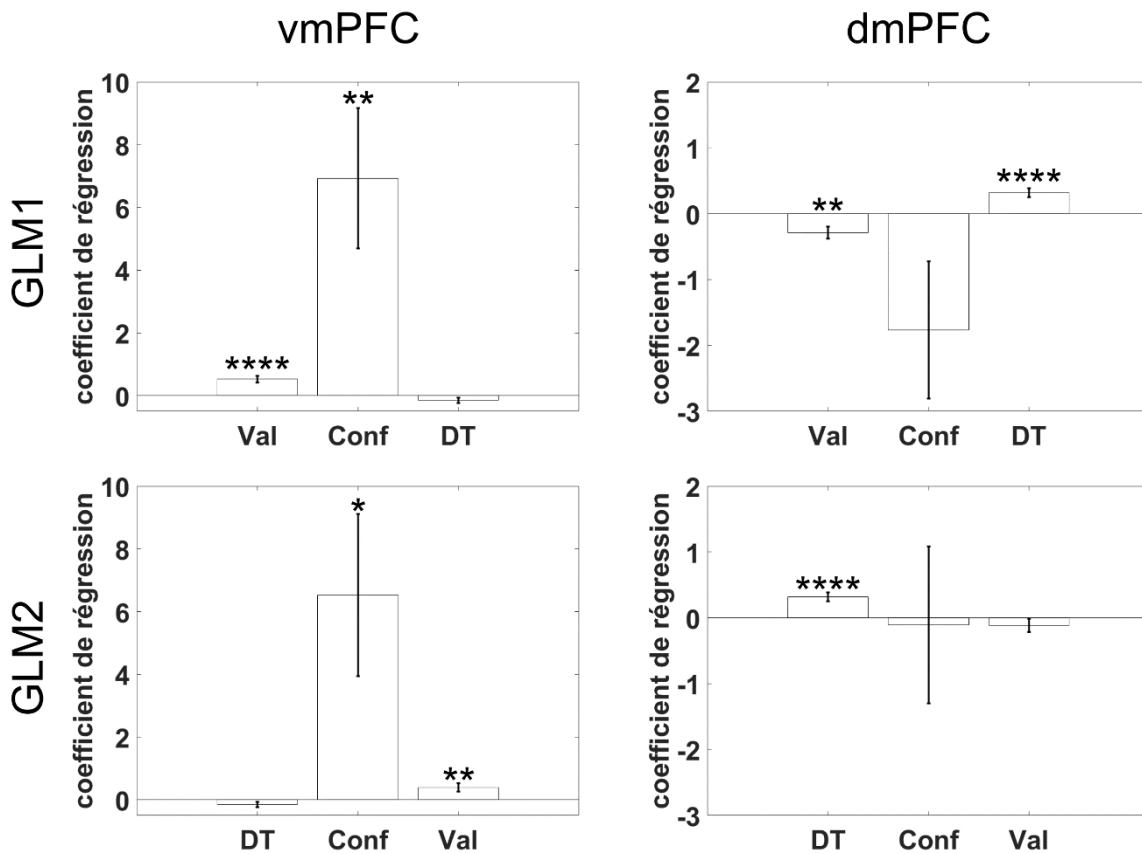
**A] Val et Conf corrént avec l'activité du vmPFC.** En haut, ces cartes représentent la carte d'activation, pour le GLM1, correspondant à Val à gauche et à Conf à droite à travers les paires associées à des gains et celles associés à des pertes. Toutes les cartes sont seuillées à  $p.value < 0.05$  corrigée au niveau des clusters pour les comparaisons multiples au niveau de la famille (*family wise error rate*). En bas, la figure montre le résultat d'une conjonction entre le contraste Val et le contraste Conf. La carte est aussi seuillée à  $p.value < 0.05$ , corrigée au niveau des clusters pour les comparaisons multiples au niveau de la famille. Le vmPFC est le seul cluster qui survit à ce seuil.

**B] DT corréle avec le dmPFC.** Corrélats cérébraux du temps de réaction pour les paires associées à des gains et les paires associées à des pertes dans le GLM1. Les deux images montrent la même carte, cependant la carte du haut est seuillée à  $p.value < 0.05$  corrigée au niveau des clusters pour les comparaisons multiples au niveau de la famille (*family wise error rate*), alors que la carte située en-dessous est seuillée à  $p.value < 0.05$  corrigée au niveau des voxels pour les comparaisons multiples au niveau de la famille.

Ensuite, nous avons défini deux régions d'intérêt sous la forme d'une sphère de 8 millimètres de rayon, centrée sur les coordonnées issues de la première étude pour Val et DT (voir **Table 1**). Nous avons ainsi confirmé que l'activité du vmPFC était corrélé à Val ( $\beta =$

## Études expérimentales

$0.523 \pm 0.112$ ,  $p = 0.0001$ ) et Conf ( $\beta = 6.931 \pm 2.232$ ,  $p = 0.005$ ), mais qu'il n'était pas significativement lié à DT ( $\beta = -0.148 \pm 0.088$ ,  $p = 0.108$ ) (voir **Figure 39**, haut). À l'inverse, l'activité du dmPFC corrélait négativement avec Val ( $\beta = -0.288 \pm 0.093$ ,  $p = 0.005$ ). L'activité du dmPFC avait aussi une tendance non-significative à décroître avec la confiance subjective ( $\beta = -1.767 \pm 1.044$ ,  $p = 0.105$ ) et à augmenter avec le DT ( $\beta = 0.319 \pm 0.066$ ,  $p = 0.0001$ ). De plus, l'activité du vmPFC était mieux corrélée à Val ( $p = 5 \cdot 10^{-7}$ ) et à Conf ( $p = 7 \cdot 10^{-8}$ ) que l'activité du dmPFC. À l'inverse, l'activité du dmPFC était mieux corrélée à DT que l'activité du vmPFC ( $2 \cdot 10^{-4}$ ). Pour voir si ces résultats étaient liés à la manière dont nos régresseurs avaient été orthogonaux, nous avons effectué un deuxième GLM où l'ordre des régresseurs était inversé, de sorte à ce que Val soit orthogonalisé à Conf et DT. En particulier, nous voulions voir si la corrélation entre le dmPFC et Val était liée au fait que la délibération était plus rapide à mesure que Val était plus élevée (voir **Figure 37**) ou si elle reflétait réellement un lien entre le dmPFC et la valeur subjective. Nous avons alors observé que les résultats étaient relativement inchangés dans le vmPFC (voir **Figure 39**, bas) dont l'activité était toujours liée à Val ( $\beta = 0.392 \pm 0.135$ ,  $p = 0.008$ ) et à Conf ( $\beta = 6.526 \pm 2.593$ ,  $p = 0.020$ ), mais pas à DT ( $\beta = -0.148 \pm 0.088$ ,  $p = 0.1075$ ), malgré la corrélation qui existait entre le DT et Val et Conf (voir **Figure 37**). Par contre, l'activité du dmPFC était toujours significativement corrélée à DT ( $\beta = 0.319 \pm 0.066$ ,  $p = 0.0001$ ) mais elle n'était plus significativement corrélée à Val ( $\beta = -0.114 \pm 0.104$ ,  $p = 0.284$ ). Elle n'était toujours pas corrélée à Conf non plus ( $\beta = -0.105 \pm 1.191$ ,  $p = 0.930$ ). Nous avons ainsi vu que le vmPFC était positivement corrélé à la fois à Val et à Conf, alors que le dmPFC était positivement corrélé avec le temps de délibération.

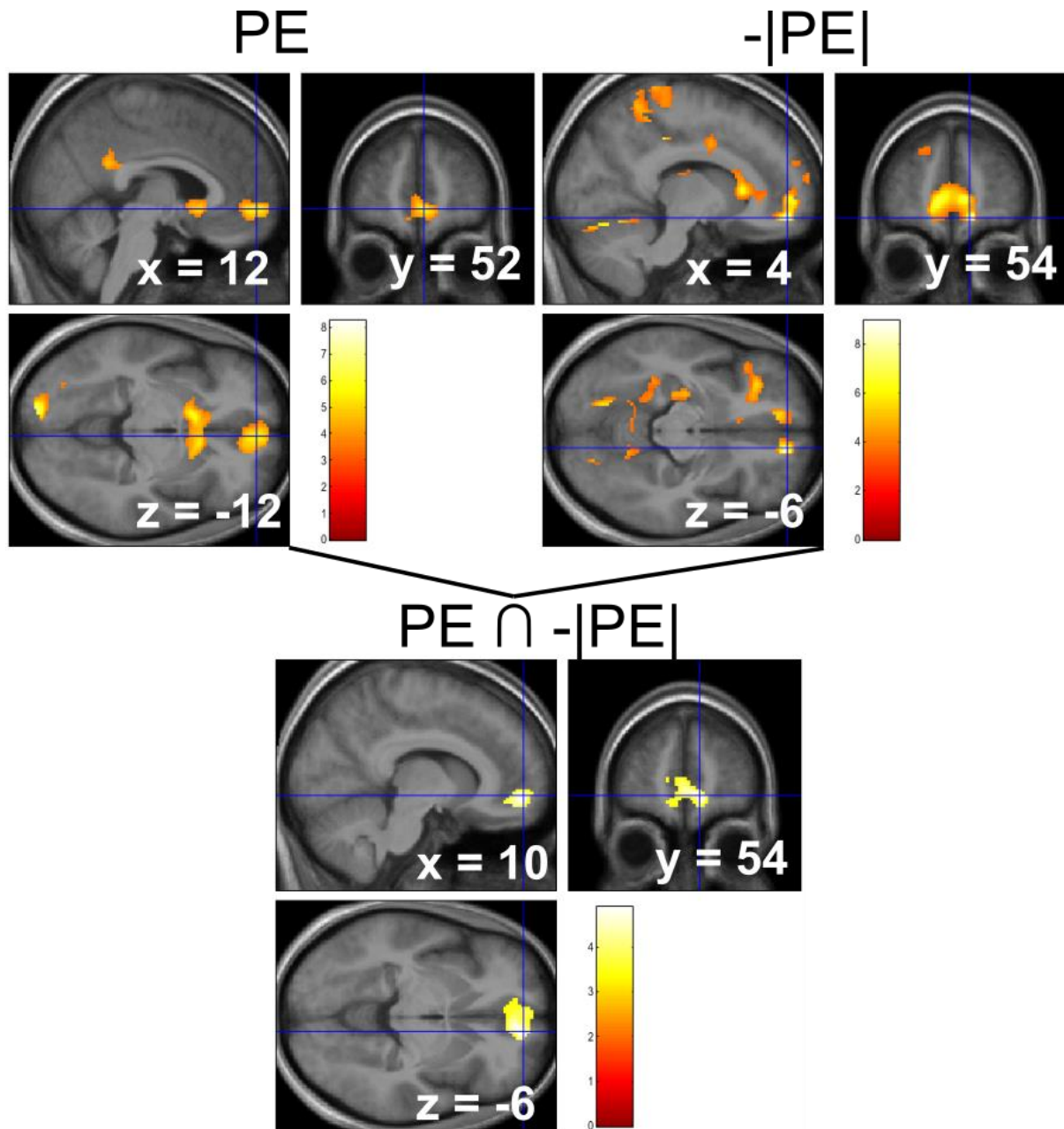


**Figure 39 : Lien entre Val, Conf, DT et l'activité du vmPFC et du dmPFC.** Les figures représentent la corrélation d'une sphère d'un rayon de 8 millimètres, centrée, à gauche, sur les coordonnées du vmPFC (-10 ; 48 ; -12) et, à droite, sur les coordonnées du dmPFC (10 ; 12 ; 48). Les coordonnées sont issues de notre première étude pour définir le vmPFC en lien avec Val et le dmPFC en lien avec le DT (voir **Table 1**). Les figures du haut sont issues du GLM1 où les régresseurs étaient orthogonalisés dans l'ordre sériel Val/Conf/DT. Les figures du bas sont issues du GLM2 où les régresseurs étaient orthogonalisés dans l'ordre sériel inverse DT/Conf/Val. Les barres et les barres d'erreurs montrent, respectivement, la moyenne et l'erreur type de la moyenne à travers les participants. Les étoiles indiquent la significativité à la suite d'un t-test des données contre zéro : \*p.value < 0.05 ; \*\*p.value < 0.01 ; \*\*\*p.value < 0.005 ; \*\*\*\*p.value < 0.001.

Nous avons aussi voulu observer les corrélats de l'erreur de prédiction signée et non-signée au moment de l'apparition du feedback à l'écran. De manière cohérente avec son rôle dans l'encodage de la valeur subjective, l'activité du vmPFC augmentait linéairement avec l'erreur de prédiction signée (voir **Figure 40**), c'est-à-dire que son activité était d'autant plus grande que le gain obtenu était plus important que ce qui était attendu et qu'elle était au contraire d'autant plus faible que la perte obtenue était plus conséquente que ce qui était attendu. On peut d'ailleurs aussi voir sur la carte deux autres composantes souvent associées au vmPFC dans ce qui a été appelé le réseau des valeurs, dans la présence du cortex postérieur cingulaire et du striatum ventral. L'activité du vmPFC était aussi négativement corrélée avec l'erreur de

## Études expérimentales

prédiction non-signée (voir **Figure 40**), c'est-à-dire que son activité était d'autant moins importante que la surprise liée au résultat donné dans le feedback était grande. A l'inverse, nous n'avons pas observé de corrélation significative entre le dmPFC et l'erreur de prédiction signée ( $\beta_{PE} = 0.190 \pm 0.261$ ,  $p = 0.474$ ) ou l'erreur de prédiction non-signée ( $\beta_{|PE|} = -0.152 \pm 0.219$ ,  $p = 0.494$ ).



**Figure 40 : Erreur de Prédiction au moment du feedback et activité du vmPFC.** En haut à gauche, zoom sur le pic d'un cluster dans le vmPFC (12 ; 52 ; -12) corrélé à l'erreur de prédiction signée. En haut à droite, zoom sur le pic d'un cluster dans le vmPFC (4 ; 54 ; -6) négativement corrélé à l'erreur de prédiction non-signée. En bas, la figure représente la conjonction entre le contraste correspondant à la corrélation positive avec l'erreur de prédiction signée et celui correspondant à la corrélation négative avec l'erreur de prédiction non-signée.



## Études expérimentales

Les cartes d'activation sont superposées à la moyenne des scans anatomiques des 22 cerveaux inclus dans l'étude après le pré-traitement, en particulier après normalisation dans l'espace MNI. Les barres colorées montrent la correspondance entre les couleurs des cartes et les valeurs T pour chaque voxel. Toutes les cartes sont montrées pour un seuil de  $p$ -value  $< 0.05$  corrigée au niveau des clusters pour les comparaisons multiples au niveau de la famille (*Family Wise Error Rate*).

### III.4 Discussion

#### III.4.a L'effort dans la délibération

L'apprentissage par renforcement est un mécanisme relativement universel. Il consiste dans l'idée simple qu'un évènement suivi d'une récompense sera associé à l'obtention de cette récompense et qu'un évènement suivi d'une punition sera associé au fait de subir cette punition, par association. Au contraire, il a été proposé que les Primates, et l'homme plus particulièrement, soit capable d'effectuer une forme plus perfectionnée de ce mode d'apprentissage, notamment en allouant plus d'effort mental (Pearce and Hall, 1980). De plus, il est connu que, dans des tâches impliquant d'explorer l'environnement par essai-erreur, le cortex préfrontal dorsomédian (dmPFC), ainsi que l'insula antérieure, qui est souvent co-activée avec lui, ont tendance à s'activer plus intensément pendant la période d'exploration que pendant la période d'exploitation de la réponse correcte (Amiez et al., 2012, 2013, 2016). Plus généralement, l'activité du dmPFC et de l'insula antérieure suite à des erreurs semblent entraîner une adaptation comportementale qui peut se résumer par un ralentissement de la vitesse de réponse et une baisse du nombre d'erreurs en vue d'améliorer la performance (Bastin et al., 2016; Iannaccone et al., 2015). Or, le dmPFC et l'insula antérieure constituent deux structures associées à l'exécution de contrôle cognitif (Menon and Uddin, 2010; Shenhav et al., 2013, 2016a), ou du moins au recrutement du dlPFC censé être la structure qui exerce le contrôle cognitif. Dans notre tâche d'apprentissage par renforcement, nous avons constaté une augmentation de l'activité du dmPFC en lien avec des temps de délibération plus longs (voir **Figure 38**). Cette corrélation était observée alors même que les temps de délibération étaient orthogonalisés par rapport à la valeur subjective et la confiance dans la décision, deux variables auxquelles ils étaient liés. Ce résultat suggère que même dans cette tâche d'apprentissage par renforcement relativement simple, les participants utilisaient du contrôle cognitif au moment de délibérer sur le symbole qu'ils allaient choisir en vue de maximiser leurs gains et de minimiser leurs pertes. Ce résultat est d'ailleurs cohérent avec les résultats de notre première étude où nous avons aussi observé une corrélation significative entre le temps de délibération dans chacune de nos trois tâches et l'activité du dmPFC et de l'insula antérieure bilatéralement

(voir **Figure 27** et **Table 1**). Enfin, nous avons vu que l'activité du dmPFC était aussi corrélée négativement avec Val dans le GLM1, mais que cette corrélation disparaissait dans notre GLM2, lorsque Val était orthogonalisé par rapport au DT. Cela suggère que, dans notre tâche du moins, si le dmPFC s'activait plus en réponse à l'anticipation de pertes plus importantes, c'était principalement lié au fait que cette condition induisait des temps de délibération plus longs. Dans notre tâche, il se pourrait ainsi que les symboles associés à des pertes impliquent une délibération plus importante. Une possibilité est que les participants présentent une aversion à la perte qui les amène à donner plus d'importance à des pertes potentielles par rapport à des gains potentiels (Kahneman and Tversky, 1984) d'où des temps de délibération plus importants. Une hypothèse alternative serait liée au fait que, dans le cas des pertes, les participants sont moins souvent exposés au feedback associé à la perte en comparaison à la situation associée à des gains. De ce fait, les participants pourraient être légèrement moins certains de la réponse à donner puisque le symbole associé à la perte n'a pas été négativement renforcé au même degré que le symbole associé au gain a été positivement renforcé. Enfin, il est possible qu'il existe un effet pavlovien qui pousse à prendre plus de temps à répondre dans l'anticipation de pertes par rapport à l'anticipation de gains (Dayan, 2008; Shadmehr et al., 2019) et que ce biais incite en quelque sorte le dmPFC à prendre plus de temps pour délibérer dans cette situation. Dans tous les cas, la corrélation entre le dmPFC et le DT au moment de la délibération suggère que notre tâche simple d'apprentissage par renforcement implique de l'effort mental. Les participants ne choisiraient pas un symbole de manière totalement automatisée, lorsqu'ils doivent choisir un symbole en vue d'optimiser leurs gains dans notre tâche, mais ils prendraient le temps de délibérer en exécutant un effort mental pour s'assurer de faire le meilleur choix possible.

### III.4.b La confiance subjective pendant la délibération

Le fait de prendre plus de temps avant de prendre une décision pourrait refléter la recherche, au niveau méta-décisionnel, d'augmenter la confiance subjective dans la décision prise (Lee and Daunizeau, 2019). Nous avons d'ailleurs pu voir que les corrélats neuraux de la confiance étaient localisés dans le cortex préfrontal ventromédian (vmPFC) au sein de la même zone que celle qui corrélait avec la valeur subjective (voir **Figure 38**). Notre analyse consiste d'ailleurs, à notre connaissance, de la première analyse montrant des corrélats cérébraux de la confiance dans une tâche d'apprentissage par renforcement. La corrélation entre Val et le vmPFC est cohérente avec des résultats précédents dans une tâche semblable (Palminteri et al., 2009) ainsi qu'avec le rôle plus général du vmPFC dans l'encodage de la valeur subjective de manière générique (Bartra et al., 2013; Levy and Glimcher, 2012). Par ailleurs, dans notre

## Études expérimentales

première étude incluant des tâches basées sur les préférences subjectives, le vmPFC corrélait aussi avec la valeur et avec la confiance subjective (voir **Figure 27**). Cependant, l'épicentre de la corrélation était dorsal au vmPFC. D'un autre côté, le lien entre le vmPFC et la confiance subjective est en accord avec une vaste littérature montrant un lien entre le vmPFC et la confiance subjective dans divers contextes (Chua et al., 2006; De Martino et al., 2013; Fleming et al., 2018; Gherman and Philiastides, 2018; Kuchinke et al., 2013; Lebreton et al., 2015; Moritz et al., 2006; Rolls et al., 2010; Shapiro and Grafton, 2020). Une limite de cette étude, dont ce n'était pas la question principale, est que les choix des participants avaient un rôle instrumental sur leur paiement final. Une confiance accrue correspondait ainsi à une espérance de gains plus élevée. De ce fait, dans cette tâche, contrairement à notre première étude où les décisions n'étaient pas instrumentales, le plaisir attendu à obtenir des gains, ou à éviter des pertes, et le plaisir lié au fait que la probabilité d'avoir choisi la meilleure des deux options était plus élevée, étaient confondus. Cela pourrait expliquer pourquoi l'épicentre de la corrélation avec la confiance, dans cette tâche, était centré sur le vmPFC. Il serait intéressant de voir si l'épicentre de la corrélation avec la confiance subjective varie selon

### III.4.c L'activité du vmPFC au moment du feedback

D'un autre côté, nous avons observé une corrélation significative dans le sens positif entre l'activité du vmPFC et l'erreur de prédiction signée et dans le sens négatif entre l'activité du vmPFC et l'erreur de prédiction non-signée (voir **Figure 40**). Ce résultat est cohérent avec l'idée que l'activité du vmPFC reflète la valeur subjective du feedback. En effet, premièrement, une erreur de prédiction signée plus élevée indique l'obtention de gains plus élevés que ce que le modèle interne prévoyait. À l'inverse, une erreur de prédiction signée plus faible signale que des pertes plus importantes que prévu ont été subies. Ce résultat est d'ailleurs en accord avec une vaste littérature liant les deux composantes principales du réseau des valeurs, qui inclue le vmPFC et le striatum ventral, avec l'erreur de prédiction signée dans des tâches d'apprentissage par renforcement (Fouragnan et al., 2018; Garrison et al., 2013). Deuxièmement, la corrélation négative entre l'activité du vmPFC et l'erreur de prédiction non-signée pourrait représenter la satisfaction subjective liée à la confirmation que le choix effectué était le bon, puisque le décalage entre les attentes et le résultat obtenu est minime.

## III.5 Conclusion

Premièrement, nous avons pu répliquer une partie des résultats de notre première étude et de la littérature dans cette tâche mélangeant des choix binaires liés à la valeur avec de

## Discussion générale

l'apprentissage par renforcement. Nous avons notamment pu confirmer une double dissociation entre le vmPFC qui encode la valeur subjective au moment du choix et l'erreur de prédiction signée au moment du feedback, alors que le dmPFC corrèle avec le DT. Deuxièmement, ce dernier résultat suggère que, même dans une tâche somme toute assez basique d'apprentissage par renforcement, les participants effectuent un effort mental, exprimé par l'activité du dmPFC, en vue de résoudre au mieux la tâche à chaque essai.

# Chapitre 3: Discussion générale

## I Le réseau des valeurs

### I.1 Le réseau des valeurs calcule un compromis coûts/bénéfices

Nos résultats confirment l'idée générale que le réseau des valeurs, un réseau principalement constitué du vmPFC et du striatum ventral, effectue un calcul de premier ordre pour identifier les bénéfices associés à la mise en place d'une série d'actions donnée.

Dans notre première étude, nous avons en effet pu observer que le vmPFC s'activait d'autant plus intensément que la valeur subjective était élevée, et ce, au sein de chacune de nos tâches (voir **Figure 27**). Nous avons aussi confirmé ce résultat dans notre troisième étude avec une tâche d'apprentissage par renforcement incluant à la fois des gains et des pertes, où le vmPFC s'activait linéairement avec la valeur subjective au moment de la décision et aussi avec l'erreur de prédiction signée au moment de l'apparition du feedback (voir **Figure 38** et **Figure 40**). Le vmPFC paraît ainsi capable d'estimer à la fois la valeur subjective au travers de plusieurs types de tâches, indépendamment du mode de présentation et de s'il estime la valeur de gains à obtenir ou de pertes à éviter. Dans notre seconde étude, nous avons aussi pu observer que le striatum était capable d'encoder les enjeux monétaires signés dans deux tâches impliquant un effort de nature physique ou mentale (voir **Figure 35**).

De plus, il a été maintes fois proposé que le vmPFC n'encode que la valeur subjective de stimuli dans « l'espace des biens » (*good space*). Il serait ainsi incapable d'encoder les coûts associés à des actions ou d'attribuer une valeur à une série d'actions directement (Padoa-Schioppa, 2011; Rangel and Hare, 2010; Schneider and Koenigs, 2017). D'un autre côté, il a été montré que le vmPFC était capable d'associer aussi bien un symbole qu'une action à l'obtention d'une récompense (Gläscher et al., 2009), c'est-à-dire qu'il est capable d'encoder la valeur subjective dans l'espace des actions (*action space*). Plusieurs études montrent que

## Discussion générale

l'encodage de la valeur subjective qui y est opéré prend aussi en compte le coût des actions nécessaires à l'obtention des stimuli (Arulpragasam et al., 2018; Seaman et al., 2018). Nos résultats vont aussi dans ce sens en démontrant que l'activité du vmPFC peut à la fois augmenter linéairement avec l'aspect appétitif de récompenses et aussi diminuer linéairement avec l'aspect aversif des coûts associés à une action de nature physique ou mentale (voir **Figure 27**). Le réseau des valeurs ne calcule donc pas que les bénéfices escomptés à la suite d'une action, mais aussi les coûts des actions nécessaires à l'obtention de ces bénéfices. Dans notre seconde étude, nous avons aussi pu observer que le striatum s'activait proportionnellement à la quantité de ressources à investir, aussi bien dans une tâche nécessitant d'effectuer un effort physique que dans une tâche nécessitant d'effectuer un effort mental (voir **Figure 34**), suggérant que ce réseau est non seulement capable d'encoder le coût des actions dans son estimation de la valeur subjective, mais aussi de déterminer la vigueur de l'action qui va être effectuée en vue d'un but.

Ces résultats sont en accord avec une littérature conséquente qui montre que l'activité de ce réseau des valeurs varie linéairement avec la valeur subjective d'un stimulus (voir *L'estimation des valeurs positives : le réseau des valeurs*) et qu'elle est dévaluée par les coûts de l'action, ainsi que par tout phénomène, comme l'incertitude ou le délai, qui réduit l'aspect appétitif d'une récompense anticipée (voir *Les dévaluateurs de la valeur subjective*). En résumé, l'activité du réseau des valeurs, un réseau principalement constitué du vmPFC et du striatum ventral, est responsable du calcul des bénéfices escomptés de la consommation d'un stimulus ou de l'évitement d'une punition, décomptés par le coût des actions visant à son obtention. Nos résultats, en plus de confirmer un rôle déjà très connu de ce réseau dans l'encodage de la valeur, ont permis de 1) confirmer ce rôle au travers de plusieurs tâches chez les mêmes participants, 2) confirmer ce rôle indépendamment du mode de présentation ou de la valence du stimulus évalué et 3) d'incorporer l'idée que ce réseau était capable de prendre en compte les coûts de l'action dans son calcul du compromis coûts/bénéfices.

### **I.2 Un réseau motivant l'action**

Nous avons vu que le réseau des valeurs calcule un compromis coûts/bénéfices. Qu'en est-il de l'impact de ce calcul sur le comportement ?

#### **I.2.a Une facilitation de l'approche ou de l'évitement en réaction à la valeur**

La théorie des perspectives développée par Daniel Kahneman et Amos Tversky établit que les individus ont une aversion pour la perte (voir *L'économie comportementale*). Ils auraient tendance à donner plus de poids aux événements aversifs qu'aux événements appétitifs

## Discussion générale

lorsqu'ils prennent une décision (Kahneman, 2012; Kahneman and Tversky, 1984). Ce biais a été identifié comme une aversion à la perte (*loss aversion*). Il s'exprimerait par le fait que la pente qui transforme les valeurs objective, au moins au niveau monétaire, en valeurs subjectives possède une pente plus importante dans le domaine des pertes que dans le domaine des gains (voir **Figure 2**). Cependant, dans nos études, nous avons observé au contraire un biais dans le sens inverse. Dans notre première étude, nos participants donnaient légèrement plus de poids à la valeur des récompenses qu'à la valeur des efforts ( $\beta_R = 1.50 > |\beta_E| = 1.12$  ;  $p.value = 0.002$ ) lorsqu'ils devaient choisir s'ils étaient prêts à exécuter un effort en vue d'une récompense. Cette différence dépend cependant bien évidemment des items que nous avons utilisé dans l'étude et ne peut pas nécessairement se généraliser à tout calcul impliquant un rapport coûts/bénéfices. Cependant, dans la tâche d'effort physique (voir *Ce qui détermine la performance*), la performance des participants était légèrement supérieure dans le cas des gains (54.31%) par rapport aux pertes (52.60%). Dans la tâche d'effort mental (voir *Ce qui détermine la performance*), cette faible différence était encore moins prononcée, mais elle allait aussi dans le même sens. Les participants avaient en effet tendance à avoir une performance équivalente voire légèrement meilleure dans le cas des gains (59.53%) que dans le cas des pertes (58.86%). Dans ces deux tâches, la performance des participants augmentait linéairement avec les enjeux signés.

De plus, dans presque toutes les tâches de la première étude (voir **Figure 26**) et aussi dans la troisième étude (voir **Figure 37**) le temps de délibération était d'autant plus rapide que la valeur subjective signée était élevée. On retrouve d'ailleurs la même tendance au niveau des temps de réaction dans les tâches de la deuxième étude, même si, dans ce cas, les corrélations n'étaient pas significatives (voir **Figure 31**).

En résumé, dans quasiment toutes nos tâches, on peut observer une tendance linéaire à répondre plus rapidement et à effectuer des efforts, tant physiques que mentaux, plus importants en fonction de la valeur subjective signée. Il a déjà été proposé que le système pavlovien serve de support à ce mécanisme. En particulier, ce système promouvrait des réactions d'approche, qui pourraient ainsi se refléter par des temps de réaction plus rapides, et d'engagement de ressources, en particulier en termes d'intensité d'un effort, comme dans nos tâches, en réponse à des stimuli appétitifs. Au contraire, ce système favoriserait des réactions d'évitement et de « désengagement » en réaction à des stimuli aversifs (Boureau and Dayan, 2011; Dayan, 2008). Il a d'ailleurs été montré qu'il existe une tendance à exercer une force proportionnelle aux enjeux même lorsque cela n'était pas instrumental. Dans une étude où la rémunération était

## Discussion générale

proportionnelle à la précision et non à l'intensité d'un effort physique, Oudiette *et al.*, ont en effet montré que les participants avaient tendance à effectuer un effort largement excédentaire par rapport à ce qu'ils auraient dû faire pour augmenter leurs gains. Ils ont en outre montré que cet effet était lié aux enjeux. Les participants avaient en effet tendance à exécuter un effort d'intensité plus élevée lorsque les enjeux étaient plus élevés, indépendamment du niveau de force qu'ils auraient dû faire pour être précis dans la tâche (Oudiette *et al.*, 2019). Plus généralement, il a été observé que les mouvements étaient d'autant plus rapides et d'intensité d'autant plus importante que la valeur signée était importante (Shadmehr *et al.*, 2019).

Or le système pavlovien est lié à au moins l'une des deux composantes du réseau des valeurs : le striatum ventral (Daw *et al.*, 2005; Rangel *et al.*, 2008). De plus, l'activité des ganglions de la base, un réseau dont le striatum ventral fait partie, a directement été associée à la fois à la vitesse et à l'intensité d'un mouvement corporel (Dudman and Krakauer, 2016; Turner and Desmurget, 2010; Turner *et al.*, 2003; Yttri and Dudman, 2016). En résumé, les ganglions de la base et le réseau des valeurs réagissent linéairement avec la valeur et la valence d'un stimulus. Ce biais engendre ensuite une réponse plus ou moins rapide et aussi plus ou moins intense en fonction de la valeur associée au stimulus en question. Il s'agit de plus là d'un biais automatique qui ne prend pas en compte les contingences d'une situation donnée. En effet, dans nos données, de même que dans l'étude de Oudiette *et al* où ce biais entre les enjeux et la force exercée se faisait au détriment de la performance dans la tâche, un comportement optimal n'aurait, *a priori*, pas dû prendre en compte de biais lié à la valeur signée. Or, dans notre seconde étude, les participants avaient tendance à effectuer un effort d'intensité plus importante selon la valeur signée. De même, dans les tâches de décision de la première étude, on voit mal l'intérêt au niveau instrumental, de répondre plus rapidement lorsque la valeur signée est plus élevée. Dans la troisième étude, comme nous l'avons déjà évoqué, ce biais pourrait refléter un temps de délibération plus long dû à une incertitude plus grande dans le cas des pertes qui ont été moins souvent renforcée que les gains, mais il n'est pas impossible non plus qu'il soit inhérent à la valence des stimuli, indépendamment de la confiance dans la réponse qui est d'ailleurs aussi incluse dans le modèle. Cependant, il est possible que ce biais ne s'exprime que dans un certain rang de valeurs. Nous avons utilisé des récompenses et des punitions de valeurs modérées dans nos tâches. Or, il a été proposé que l'aversion à la perte ne se manifeste que pour des enjeux très élevés (Mukherjee *et al.*, 2017). Il serait ainsi intéressant de constater si ce biais pavlovien sur la vitesse et la vigueur d'une réaction face à un item dépend du rang des valeurs considérées ou s'il se généralise à toutes les situations.

## Discussion générale

L'activité du réseau des valeurs, au moins au niveau du striatum ventral, prend en compte un biais pavlovien. Elle est d'autant plus importante que la valeur subjective signée est élevée. Ceci se reflète ensuite dans le comportement par une réaction d'approche plus rapide et/ou un effort, physique ou mental, d'intensité plus élevée.

### I.2.b Un réseau qui motive l'action en fonction d'un but

En résumé, le réseau des valeurs calcule la valeur nette d'un compromis bénéfices/coûts en vue de maximiser le taux de récompenses obtenues et de minimiser le taux de punition subies, tout en minimisant l'effort fourni dans ce but. De plus, considérons que le réseau des valeurs permet de faciliter l'exécution d'un effort, physique ou mental, en vue d'un but déterminé après avoir calculé le compromis coûts/bénéfices associé à un but donné et avoir considéré que le but en question en valait la peine. Il paraît alors logique que, dans notre seconde étude incluant une tâche d'effort physique et une tâche d'effort mental et où effectuer un effort plus important était instrumental aussi bien dans le cas des gains que dans le cas des pertes, l'activité du striatum (voir **Figure 34** et **Figure 35**), de même que la performance (voir **Figure 33**), variait linéairement avec les enjeux monétaires non-signés. Nos études sont simplement de nature corrélacionnelle de sorte qu'il est impossible d'en tirer réellement un lien de causalité direct. Il est en effet aussi possible d'interpréter ce lien comme un simple reflet de l'anticipation de la récompense obtenue, ou de la punition évitée, avant même que la performance ait eu lieu. Cependant, il a déjà été montré que le striatum, et les ganglions de la base plus généralement, jouent un rôle dans la vigueur d'un effort effectué en vue d'un but (Schmidt et al., 2008). Il serait cependant fort intéressant de voir dans quelle mesure l'activité du vmPFC détermine aussi la quantité d'effort qui va être fourni ou si, à l'inverse des ganglions de la base, il ne jouerait qu'un rôle passif en reflétant l'anticipation du plaisir associé à l'obtention d'une récompense ou à l'évitement d'une punition sans avoir de rôle direct dans la rapidité ou l'intensité de l'effort effectué en vue d'un but.

## I.3 Qu'est-ce que le vmPFC ?

S'il fallait choisir les aires anatomiques à identifier avec les clusters fonctionnels du réseau des valeurs, avec le terme « cortex préfrontal ventromédian » (*ventromedial prefrontal cortex* ou *vmPFC* en anglais), je choisirais l'aire 14m et, avec le striatum ventral, je choisirais le noyau accumbens (voir **Figure 8**). Cependant, il faut noter que, dans le domaine, ces équivalences sont loin d'être consensuelles, en particulier en ce qui concerne le vmPFC. Les chercheurs utilisant le terme de vmPFC l'emploient en général en référence à des activations



## Discussion générale

fonctionnelles plus qu'à une position et une aire anatomique bien définie au niveau du cortex préfrontal. Sous ce terme, on peut ainsi trouver des études qui parlent de la partie périgénuale du cortex cingulaire, c'est-à-dire de l'aire 24 (Levy and Glimcher, 2011; Lopez-Persem et al., 2016) ou encore de la partie subgénuale du cortex cingulaire (Smith et al., 2010a), c'est-à-dire de l'aire 25. Ce manque de cohérence dans la nomenclature pourrait créer des malentendus dans l'interprétation des résultats puisque le même terme, le vmPFC, englobe des populations neurales avec une connectivité et des rôles fonctionnels potentiellement très différents. D'un autre côté, il faut noter que nos résultats ne permettent pas *a priori* d'identifier le cluster fonctionnel constitué par le vmPFC à une aire anatomique bien définie. En effet, dans notre première étude par exemple, lorsque nous avons observé les corrélats cérébraux de la valeur subjective à travers nos tâches, nous avons aussi bien pu observer un cluster que nous avons nommé vmPFC avec un pic aux coordonnées MNI (-10 ; 48 ; -12) et qui correspond plus ou moins à l'aire 14m, mais aussi un cluster dans la partie périgénuale du cortex cingulaire avec un pic aux coordonnées (-8 ; 38 ; 6) (voir **Table 1**). Pour ce qui est du striatum, dans notre deuxième étude (voir **Figure 34**), tout comme dans l'article original de Schmidt *et al.* (Schmidt et al., 2012), l'aire correspondant aux enjeux paraît légèrement plus dorsale que la localisation où le noyau accumbens est généralement supposé se trouver. Il se pourrait ainsi qu'il s'agisse d'une autre composante du striatum qui soit impliquée dans la transformation des enjeux monétaires en performance, physique ou mentale.

De plus, il a été montré qu'il existe au moins trois patterns anatomiques différents à travers les participants au niveau de cette zone (Chiavaras and Petrides, 2000; Mackey and Petrides, 2014). Plusieurs études montrent d'ailleurs que les clusters d'activation fonctionnels sont plus ou moins décalés entre les participants en fonction du pattern anatomique auxquels ils appartiennent (Li et al., 2015; Lopez-Persem et al., 2018). Mieux prendre en compte la diversité des participants en termes d'architecture cérébrale nécessite certes des pools de participants plus importants que ceux que nous avons réunis dans nos tâches. Cependant, cela permettrait d'améliorer notre compréhension du fonctionnement du cerveau. En particulier, nous pourrions voir quelles sous-parties du cortex préfrontal ventromédian sont réellement impliquées et nécessaires dans l'évaluation de la valeur subjective. Une telle approche permettra à l'avenir aussi de mieux identifier les structures correspondantes chez les modèles animaux, en particulier chez le primate non-humain.

Un avantage de notre approche, que ce soit dans la première ou dans la seconde étude, est que nous avons cherché à identifier les corrélats cérébraux des concepts cognitifs qui nous

## Discussion générale

intéressaient, la valeur subjective, la confiance et l'effort lié à la décision, au sein des mêmes participants au travers de plusieurs tâches. Nos résultats ne pouvaient ainsi pas être biaisés par des différences anatomiques entre les participants associés à chaque tâche puisque les mêmes participants effectuaient toutes les tâches.

### **I.4 Conclusion sur le réseau des valeurs et la valeur nette**

*« When you are in a state of cognitive ease, you are probably in a good mood, like what you see, believe what you hear, trust your intuitions, and feel that the current situation is comfortably familiar. You are also likely to be relatively casual and superficial in your thinking. »*

*Chapitre 5, Thinking, Fast and Slow, Daniel Kahneman (Kahneman, 2012)*

En résumé, le réseau des valeurs calcule le compromis bénéfices/coûts. Il s'agit d'une valeur nette qui prend en compte d'un côté les bénéfices éventuels associés à une action, comme le fait de profiter d'un bon repas ou d'éviter de subir une perte monétaire, et de l'autre les coûts associés à cette action, comme l'activité physique ou mentale nécessaire en vue d'obtenir cette récompense.

Dans ce calcul, le réseau des valeurs est cependant soumis à plusieurs biais qui sont liés à son fonctionnement. Nous avons vu premièrement qu'il pourrait exister un biais pavlovien dans la manière dont ce réseau considère tout stimulus. Le réseau des valeurs va ainsi systématiquement faciliter l'exécution plus rapide d'un effort d'intensité plus importante en réponse à des stimuli appétitifs, même si cela n'est pas instrumental. De plus, tout phénomène faisant varier l'activité basale de ce réseau qui varie avec l'humeur subjective d'un individu, peut impacter l'évaluation qu'il va effectuer d'un stimulus donné. Il a ainsi été montré qu'une activité plus élevée dans le vmPFC, liée à une humeur plus joyeuse, augmenterait la sensibilité à des gains monétaires (Vinckier et al., 2018). De même, il a été montré que des participants à qui on demandait d'évaluer des tableaux étaient impactés, dans leur évaluation, par le fait d'écouter une musique qu'ils jugeaient agréable. Cette modulation passait, là encore, par l'activité d'une des deux composantes principales du réseau des valeurs : le vmPFC (Abitbol et al., 2015). Plus généralement, l'économie comportementale a identifié toute une série de biais comportementaux qui affectent la prise de décision. Nous en avons déjà évoqué certains et nous ne chercherons pas ici à en faire la liste (voir (Kahneman, 2012) pour une liste exhaustive). Ces biais concerneraient principalement le fonctionnement de ce que Daniel Kahneman, reprenant des termes introduits par Keith Stanovich et Richard West en psychologie (Stanovich and West,

2000), appellera le Système 1 (Kahneman, 2012). Daniel Kahneman affirme que cette notion est purement opérationnelle et qu'elle n'est pas ancrée dans l'idée qu'il en existerait des bases neurales<sup>25</sup>. Il serait cependant intéressant de vérifier dans quelle mesure les biais identifiés par l'économie comportementale sont supportés par le fonctionnement du réseau des valeurs, afin de confronter à nouveau l'étude des bases neurales de la prise de décision et l'économie comportementale. Il a par exemple été montré que le striatum ventral déformait les probabilités de manière similaire aux prédictions de la théorie des perspectives (voir **Figure 11**) (Hsu et al., 2009). Nos résultats sont cependant en contradiction avec l'une des prédictions de la théorie des perspectives qui suggère que l'activité du réseau des valeurs devrait être biaisée par une aversion à la perte en s'activant plus fortement face à des pertes qu'à des gains équivalents. Nous avons en effet constaté un biais dans le sens inverse puisque l'activité du réseau des valeurs était positivement corrélée avec les enjeux monétaires signés. Quoiqu'il en soit, il paraît clair que non seulement le réseau des valeurs calcule un compromis bénéfices/coûts, mais aussi que ce calcul est sujet à plusieurs biais innés.

Enfin, une fois que ce calcul a été effectué en intégrant la valeur subjective des coûts et des bénéfices associés à l'action ainsi que les biais qui influent sur l'activité de ce réseau, alors ce signal pourrait servir à motiver la performance. Ce que l'on entend par là consiste à allouer plus ou moins de ressources en vue d'atteindre le but qui aura obtenu la valeur nette la plus élevée.

## II La confiance

### II.1 Valeur et confiance

L'encodage de la confiance au cours d'une tâche pourrait avoir un intérêt double en termes de survie. Premièrement, encoder la **confiance** de manière **prospective**, c'est-à-dire avant de prendre une décision, permet de ne prendre une décision que lorsqu'un certain seuil de confiance est atteint. La confiance, vue comme l'estimation de la probabilité qu'une série d'actions donnée permette d'obtenir une récompense, permettrait de moduler la mise en action, notamment en vue d'éviter d'engager inutilement des ressources, qu'il s'agisse simplement de temps, comme proposé par la théorie du coût d'opportunité, ou de ressources physiologiques si la probabilité de succès de la série d'actions envisagée est très faible. Deuxièmement, encoder

---

<sup>25</sup> « *there is no one part of the brain that either of the systems would call home* », c'est-à-dire « *il n'y a aucune partie du cerveau où ces deux systèmes logeraient* », Chapitre I, *Thinking, Fast and Slow*, Daniel Kahneman (Kahneman, 2012)

## Discussion générale

la **confiance** de manière **rétrospective**, c'est-à-dire par rapport à une série d'actions qui a déjà eu lieu, pourrait permettre de moduler le comportement futur de l'individu. Plusieurs études montrent que la confiance subjective à chaque essai pourrait avoir un impact sur les essais suivants, notamment en termes d'allocations de ressources (voir *La confiance*). Il est d'ailleurs connu depuis plusieurs années que la détection de la commission d'une erreur entraîne un ralentissement de la réponse lors des essais suivants, ainsi qu'une réduction de la proportion d'erreurs lors des essais suivants. Cet effet est connu sous le nom de ralentissement post-erreur (*post-error slowing*). Dans nos études, nous nous sommes principalement intéressés à la confiance rétrospective, c'est-à-dire la probabilité d'avoir pris la bonne décision, mais sans nécessairement chercher à voir l'impact de celle-ci sur les essais ultérieurs ce qui n'était pas une question centrale de nos études. Il serait intéressant cependant d'étudier dans quelle mesure le niveau de confiance à un essai donné pourrait influencer le traitement des essais suivants dans nos tâches de décision.

Nous venons d'évoquer les biais qui peuvent concerner le fonctionnement du réseau des valeurs (voir *Conclusion sur le réseau des valeurs et la valeur nette*). Or, il a été montré que la confiance prospective dans la performance augmente linéairement avec la valeur des enjeux signés, indépendamment de la performance (Bioud, 2019; Lebreton et al., 2018). Ce résultat suggère que les mêmes aires cérébrales qui encodent la valeur subjective sont aussi responsables d'encoder la confiance. Lorsque l'une de ces deux variables est impactée, cela pourrait ainsi aussi influencer l'encodage de l'autre variable. Il a d'ailleurs déjà été montré que l'activité du réseau des valeurs était dévaluée par l'incertitude, au moins dans le cas de décisions probabilistes et avant que la décision n'ait lieu (voir *Le réseau des valeurs est impacté par l'incertitude*). Dans nos résultats, nous avons confirmé, comme beaucoup d'études précédentes dans plusieurs contextes allant de la mémoire, la prise de décision basée sur la valeur ou des informations perceptuelles (Chua et al., 2006; De Martino et al., 2013; Fleming et al., 2018; Gherman and Philiastides, 2018; Lebreton et al., 2015; Lopez-Persem et al., 2020)), que le vmPFC était positivement corrélé à notre mesure de confiance dans notre première étude, comme dans notre troisième étude (voir **Figure 27**, **Figure 38** et **Figure 39**). Ce résultat pointe vers l'idée que le réseau des valeurs encode à la fois le résultat du compromis coûts/bénéfices et la confiance dans ce calcul. Cela expliquerait pourquoi des enjeux signés plus importants pourraient aussi biaiser l'estimation de la confiance prospective si c'est la même zone qui est en charge de ces deux calculs.

## II.2 La confiance peut être dissociée de la valeur

Dans notre première étude, il apparaît pourtant clairement que le cluster fonctionnel principalement lié à la confiance peut être dissocié du vmPFC (voir **Figure 27**). Ce résultat est d'ailleurs en accord avec l'idée d'un gradient dorso-ventral au sein du cortex préfrontal médian (voir **Figure 22**) qui a été développée dans une tâche d'évaluation montrant que la confiance était mieux encodée dans la partie dorsale du cortex préfrontal médian alors que la valeur était mieux encodée dans la partie ventrale (De Martino et al., 2017). Les résultats de notre première étude offrent l'avantage qu'ils montrent une représentation de la confiance dans le mPFC de manière stable au travers de trois tâches incluant à la fois une tâche d'évaluation et des tâches de choix. Nous n'avons cependant pas réussi à répliquer cette dissociation dans notre troisième étude où, d'ailleurs, la corrélation entre une sphère d'un rayon de 8 millimètres centrée sur les coordonnées du cluster lié à la confiance de notre première étude et la confiance n'était qu'au bord de la significativité ( $\beta = 7.644 \pm 3.833$ ,  $p = 0.059$ ), contrairement au vmPFC où la corrélation était très claire. Cependant, le résultat de notre première étude montre que, au moins en principe, il existe des aires cérébrales capables de corrélérer avec la confiance, c'est-à-dire la probabilité d'avoir pris la bonne décision, qui sont distinctes des aires liées au calcul direct du compromis coûts/bénéfices. L'idée que la confiance soit codée de manière générique dans une échelle commune (de Gardelle and Mamassian, 2014; de Gardelle et al., 2016) suggère que, comme pour la valeur subjective et le réseau des valeurs, il devrait aussi être possible d'identifier des corrélats cérébraux stables de celle-ci au travers des paradigmes, comme notre première étude nous a permis de le voir. Notre première étude pointe vers l'idée que le mPFC constitue une aire cérébrale capable d'encoder la confiance de manière générique, c'est-à-dire indépendamment de la tâche effectuée, mais des études complémentaires sont nécessaires pour voir si ce résultat est généralisable à d'autres tâches. Nous n'avons en effet pas réussi à le répliquer dans notre troisième étude où la valeur et la confiance avaient les mêmes corrélats cérébraux.

Une possibilité pour expliquer le fait qu'il n'existe pas de consensus autour de la localisation spatiale dans le cerveau de l'encodage de la confiance (voir *Y a-t-il une zone cérébrale de la confiance ?*) serait que la dynamique temporelle ait un impact sur les activations cérébrales observables en IRMf. En effet, si l'encodage de la confiance transite au cours du temps d'une zone cérébrale à une autre, il est possible que selon les paramètres d'acquisition et d'analyse utilisés dans chaque expérience, on n'arrive qu'à prendre une sorte de photo instantanée de la position de l'encodage de la confiance à un moment donné, sans avoir accès

## Discussion générale

aux différentes zones par lesquelles ce signal a transité. Des études complémentaires en magnétoencéphalographie, en électroencéphalographie, ou en électroencéphalographie intracrâniennes permettront peut-être de mieux comprendre la dynamique de l'encodage de la confiance dans les prochaines années. Il a par exemple été montré, dans une tâche impliquant un compromis entre des bénéfices monétaires et le fait de subir un choc électrique (Shapiro and Grafton, 2020), que le vmPFC encodait d'abord la valeur subjective, puis la confiance dans la décision (voir **Figure 24**). Il serait ainsi particulièrement intéressant de voir si le signal encodant la confiance transite, au cours du temps, le long de l'axe dorso-ventral du cortex préfrontal médian.

## III Les ressources neurales allouées au compromis coûts/bénéfices

### III.1 Le dmPFC et la prise de décision

Considérons l'équation :  $SV = R \cdot p - P \cdot (1 - p) - C$  où  $SV$  est la valeur subjective,  $R$  est une récompense à obtenir,  $P$  une punition à éviter,  $p$  la probabilité de réussir, c'est-à-dire la confiance prospective, et  $C$  les coûts des actions nécessaires à atteindre le but fixé (obtenir  $R$  ou éviter  $P$ ). Le dmPFC a été associé à une quantité impressionnante de concepts cognitifs très variés dont plusieurs sont en lien avec différentes parties de cette équation. Plus spécifiquement, si l'on reprend cette formule, le dmPFC a notamment été associé à :

- 1) l'encodage des valeurs négatives ( $P$  ou  $-(R-P)$ ) (Bartra et al., 2013; Hayes and Northoff, 2011), c'est-à-dire un encodage de la valeur associée à un stimulus qui irait dans le sens opposé au codage du réseau des valeurs (voir *Un réseau appétitif et un réseau aversif*) ;
- 2) le calcul de l'incertitude ( $p$ ) (voir *Les bases neurales de l'incertitude*) (Christopoulos et al., 2009; Volz et al., 2005) ;
- 3) le calcul des coûts associés à une action ( $C$ ) (Kurniawan et al., 2013; Skvortsova et al., 2014) ;
- 4) l'intégration dynamique du compromis coûts/bénéfices en vue d'estimer une valeur nette ( $SV$ ) (Pessiglione et al., 2018; Pisauro et al., 2017) ;
- 5) le calcul de la valeur d'implémenter du contrôle cognitif (Shenhav et al., 2013), c'est-à-dire une méta-variable qui module le calcul de  $SV$ .

En résumé, le dmPFC a été connecté à pratiquement toutes les variables de l'équation régulant la  $SV$ , et même au calcul de la  $SV$  elle-même. Nous n'allons bien évidemment pas prétendre avoir tranché cette question dans cette thèse. Cependant, nos résultats nous ont permis d'obtenir

## Discussion générale

quelques éléments en faveur ou en défaveur de certaines de ces hypothèses. Nous allons ici chercher à reprendre les différents résultats au travers de nos études pour voir s'il est possible d'y voir plus clair sur la base des résultats présentés ici.

- **Hypothèse 1 : la valeur négative des stimuli**

Nos résultats paraissent partiellement compatibles avec cette hypothèse en ce que, dans la première étude, nous avons pu voir que l'activité du dmPFC diminuait avec l'aspect appétitif des récompenses présentées dans la tâche d'évaluation ( $\beta = -0.229 \pm 0.107$ ,  $p = 0.038$ ). De même, dans la troisième étude, l'activité du dmPFC diminuait avec la valeur subjective des items présentés ( $\beta_{\text{Val}} = -0.288 \pm 0.093$ ,  $p = 0.005$ ). Cependant, on ne peut pas non plus voir le dmPFC comme une zone dont l'activité serait simplement l'opposée du vmPFC puisque, comme nous l'avons vu, alors que l'activité du vmPFC décroissait avec la pénibilité des efforts ( $\beta = -0.670 \pm 0.205$ ,  $p = 0.002$ ), le dmPFC affichait une tendance similaire ( $\beta = -0.225 \pm 0.116$ ,  $p = 0.060$ ). Si cette hypothèse paraît donc possible, la vision selon laquelle le vmPFC et le dmPFC formeraient deux systèmes neuraux tirant chacun dans une direction opposée, paraît peu plausible, en tout cas comme un processus général qu'on pourrait retrouver dans toutes les situations.

- **Hypothèse 2 : l'incertitude**

Nous n'avons pas pu observer de corrélation claire entre le dmPFC et notre mesure indirecte de la confiance, c'est-à-dire l'équivalent de la variable  $p$  de l'équation ci-dessus, dans notre première étude même si l'activité du dmPFC avait en effet tendance à diminuer avec  $p$  ( $\beta_{\text{Conf}} = -0.098 \pm 0.089$ ,  $p = 0.276$ ). Dans notre troisième étude, l'activité du dmPFC avait aussi tendance à diminuer avec le degré de confiance ( $\beta_{\text{Conf}} = -1.767 \pm 1.044$ ,  $p = 0.105$ ). Même si nos résultats suggèrent que, si effet il y a, il est relativement faible, ils ne permettent donc pas de complètement infirmer cette hypothèse.

- **Hypothèse 3 : calculer le coût des actions pour retirer ce coût à la SV**

Nous avons vu que l'activité du dmPFC ne variait pas avec les coûts associés à une activité physique ou mentale envisagée (voir *Le cortex préfrontal médian et les différents aspects du compromis coûts/bénéfices*) comme le prédirait l'hypothèse (3) évoquée ci-dessus. L'activité du dmPFC paraît en effet augmenter avec le fait d'exécuter un effort plus qu'avec le coût fictif d'envisager un effort. Ce résultat est partiellement contradictoire avec l'étude de Kurniawan *et al.*, où l'activité du dmPFC augmentait linéairement avec le coût des efforts même dans le cas où un effort n'avait pas à être implémenté par la suite (Kurniawan *et al.*, 2013).

## Discussion générale

Cependant, le design de la tâche en question, qui alternait des essais où un effort devait effectivement être exécuté et d'autres essais où l'effort était purement fictif, pourrait pousser le dmPFC à investir des ressources pour répondre au besoin d'exécuter un effort même dans le cas fictif, d'autant plus qu'au moment où la corrélation a été observée, les participants ne savaient pas encore s'ils devaient, ou non, effectuer l'effort requis. Dans notre cas où tous les efforts étaient purement fictifs, nous avons pu voir clairement que l'activité du dmPFC non seulement n'augmentait pas, mais elle avait même plutôt tendance à diminuer avec la pénibilité des efforts. Au contraire, nous avons vu que ce calcul pouvait avoir lieu dans les structures du réseau des valeurs comme le vmPFC ou le striatum ventral (voir **Figure 27**). Ces résultats sont d'ailleurs en accord avec d'autres études ayant montré que l'évaluation des coûts d'un effort est effectuée au niveau du vmPFC, alors que la quantité d'effort qui doit être allouée au cours de la prise de décision est liée à l'activité du dmPFC (Aridan et al., 2019; Hogan et al., 2019; Westbrook et al., 2019).

- **Hypothèse 4 : la valeur nette**

L'activité du dmPFC n'était pas significativement corrélée à la valeur nette issue du compromis coûts/bénéfices dans notre première étude ( $\beta = 0.214 \pm 0.122$ ,  $p = 0.087$ ), contrairement à ce que l'hypothèse (1) propose, bien que la tendance aille néanmoins dans cette direction. Sans éliminer complètement cette possibilité, nos résultats suggèrent néanmoins que cette corrélation, si elle est valable, est trop faible pour être visible dans nos données et que l'activité du dmPFC est mieux associée à d'autres variables lors du compromis coûts/bénéfices.

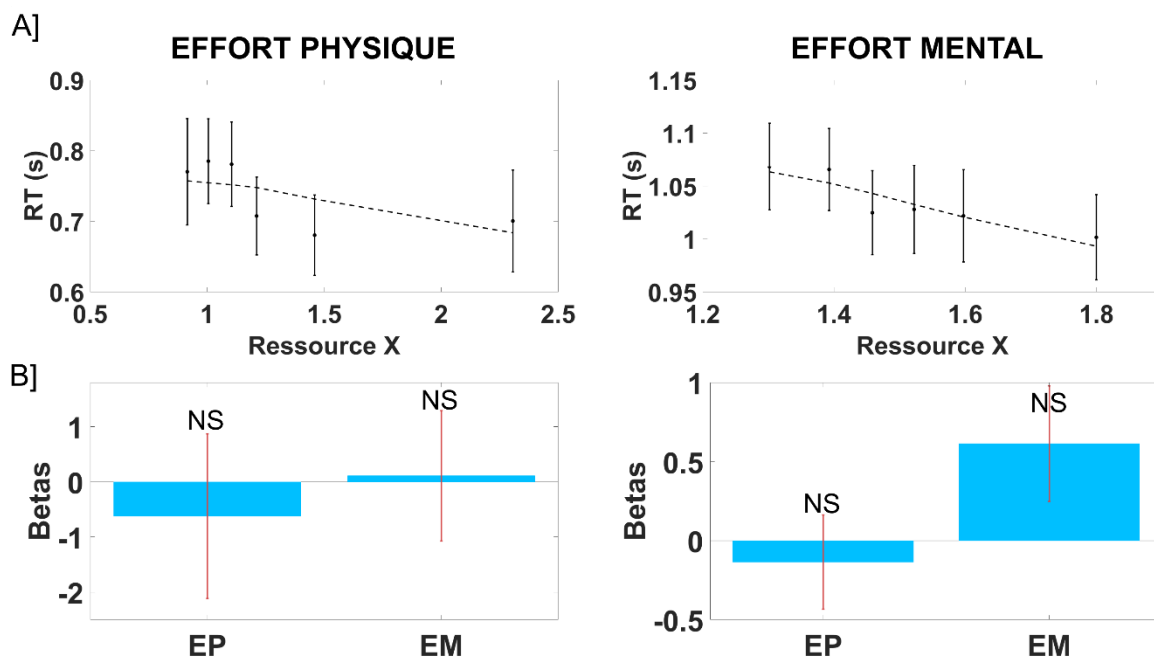
- **Hypothèse 5 : une variable méta-décisionnelle**

Comme nous venons de le voir, les éléments issus de nos résultats ne permettent pas de définitivement éliminer toutes les hypothèses sur le rôle du dmPFC. Si ce n'est l'hypothèse que le dmPFC s'occupe systématiquement d'encoder la pénibilité des efforts associés à l'obtention d'un stimulus que nous avons invalidé au vu de nos résultats, les autres hypothèses paraissent toujours plausibles, en tout cas il est impossible de les invalider sur la base de nos études. Cependant, s'il y a un résultat qui paraît robuste dans nos études, c'est la corrélation positive entre l'activité du dmPFC et le temps de délibération dans des tâches impliquant une prise de décision. Nous avons répliqué ce résultat dans pas moins de quatre tâches différentes (voir **Figure 27** et **Figure 38**), dont les trois premières ont été effectuées par les mêmes participants. Or, le temps de réaction est régulièrement considéré comme un reflet des ressources allouées à une décision, en particulier en réponse à des situations où la décision est difficile et nécessite



## Discussion générale

de l'effort mental (Alnaes et al., 2014; Ferguson et al., 2013; Haji et al., 2015; Kamiński et al., 2017; Pals et al., 2015; Robinson et al., 1997; Takeuchi et al., 2011; Vassena et al., 2014). Pour vérifier que cette corrélation entre le dmPFC et le temps de réaction était spécifique aux tâches où un temps plus long est associé à l'investissement de plus de ressources neurales, nous avons voulu vérifier que ce lien n'était valable que dans les tâches de notre première et de notre troisième étude qui impliquaient un temps de délibération plus long en lien avec une prise de décision plus difficile. Pour ce faire, nous avons vérifié que, dans les tâches de notre deuxième étude, impliquant un effort physique et un effort mental, le temps de réaction n'augmentait pas avec la quantité de ressources investies d'après notre modèle. En effet, dans ces tâches, le temps de réaction était *a priori* indépendant de la performance dans la tâche d'effort physique et, dans la tâche d'effort mental, investir plus de ressources devrait au contraire amener à des temps de réaction plus rapides pour résoudre la tâche plus rapidement. Nous avons en effet pu voir que le temps de réaction avait plutôt tendance à accélérer avec la quantité de ressources investies (voir **Figure 41A**) que ce soit dans la tâche d'effort physique ( $\beta_X = -0.042 \pm 0.056$  ; p.value = 0.460), où cette corrélation n'était pas significative, ou dans la tâche d'effort mental ( $\beta_X = -0.115 \pm 0.029$  ; p.value =  $6 \cdot 10^{-4}$ ). Nous avons déjà vu que, dans ces tâches, le dmPFC s'activait proportionnellement à la quantité de ressources à investir (voir **Figure 34**). Pour vérifier qu'il n'était pas lié au temps de réaction, nous avons défini deux GLM supplémentaires où le temps de réaction était inclus, soit en modulateur paramétrique de la période d'affichage des enjeux, soit en modulateur paramétrique de la période d'affichage de l'échelle d'effort à l'écran, en plus des autres variables impactant la performance (les enjeux monétaires signés, les enjeux monétaires non-signés et le numéro d'essai). Nous avons alors pu confirmer que, alors que son activité augmente avec les enjeux monétaires non-signés et la quantité de ressources à investir, son activité, dans nos tâches d'effort physique (modulation pendant les enjeux :  $\beta = -0.621 \pm 1.492$  ; p.value = 0.682 ; modulation pendant l'effort :  $\beta = -0.136 \pm 0.298$  ; p.value = 0.654) et d'effort mental (modulation pendant les enjeux :  $\beta = 0.110 \pm 1.18$  ; p.value = 0.927 ; modulation pendant l'effort :  $\beta = 0.613 \pm 0.366$  ; p.value = 0.110), n'était pas significativement corrélée aux temps de réaction (voir **Figure 41B**). En résumé, l'activité du dmPFC était positivement corrélée à la quantité de ressources investies dans toutes les tâches de toutes nos études. Dans le cas des tâches impliquant une prise de décision, ce lien se reflétait dans la corrélation entre l'activité du dmPFC et la durée du temps de délibération. Dans les tâches impliquant un effort physique ou mental, ce lien se reflétait dans la corrélation entre l'activité du dmPFC et la quantité optimale de ressources à investir à chaque essai d'après notre modèle computationnel.



**Figure 41 : Le dmPFC reflète les ressources allouées et pas le temps de réaction.**

A] **Temps de réaction en fonction de la quantité de ressources X investies.** Les points ainsi que les barres d'erreur représentent la moyenne ainsi que l'erreur type de la moyenne à travers les participants et les sessions, du temps de réaction en fonction de la quantité de ressources X investies dans la tâche d'effort physique, à gauche, et dans la tâche d'effort mental, à droite. La quantité de ressources X a été estimée à l'aide de notre second modèle computationnel. Les traits en pointillés représentent la moyenne correspondante à travers les participants de l'estimation par notre GLM du temps de réaction en fonction de X. Les données des figures sont basées sur les 23 participants inclus dans l'étude.

B] **Le dmPFC n'est pas significativement corrélé aux temps de réaction dans la tâche d'effort physique et dans la tâche d'effort mental.** Les deux figures représentent la corrélation entre le dmPFC et les temps de réaction dans la tâche d'effort physique (EP) et dans la tâche d'effort mental (EM). Pour la tâche d'effort physique, nous avons considéré le temps de réaction comme la différence entre le moment d'apparition de l'échelle d'effort à l'écran et le moment où le participant commençait à serrer la poignée de force. Pour la tâche d'effort mental, nous nous sommes concentrés sur le temps mis pour répondre à la première paire de chaque essai à partir de l'apparition des paires à l'écran. La figure de gauche représente la corrélation avec le temps de réaction dans un GLM où le temps de réaction modulait le moment d'apparition des enjeux à l'écran. La figure de droite représente la corrélation avec le temps de réaction dans un GLM où le temps de réaction modulait le moment où les participants pouvaient commencer à effectuer un effort. N = 20 participants. Le sens des symboles est le suivant : NS : non-significatif ; \*p.value<0.05 ; \*\*p.value < 0.01 ; \*\*\*p.value < 0.005 ; \*\*\*\*p.value < 0.001.

Ces résultats sont compatibles avec l'idée que le dmPFC encode la valeur attendue d'exercer du contrôle cognitif et qu'il régulerait ainsi l'exécution du contrôle cognitif par d'autres structures cérébrales comme le dlPFC (Shenhav et al., 2013, 2014, 2016a). Ceci expliquerait d'ailleurs pourquoi l'activité du dmPFC est corrélée à certaines des variables mentionnées précédemment. En effet, le calcul de la valeur attendue d'exercer du contrôle

## Discussion générale

cognitif doit aussi bien prendre en compte l'importance des enjeux, c'est-à-dire les bénéfices à obtenir ( $R$ ) et/ou les pertes à éviter ( $P$ ) en exerçant du contrôle cognitif, que l'incertitude ( $p$ ) et la difficulté de la situation qui requiert, de ce fait, d'exécuter du contrôle cognitif. Plus généralement, il est aussi possible que l'activité du dmPFC reflète l'engagement de ressources physiologiques au niveau de l'organisme entier, en stimulant par exemple une accélération du pouls cardiaque via une régulation de l'activité du système nerveux autonome dans des situations saillantes nécessitant d'allouer plus d'effort physique ou mental (Beissner et al., 2013; Critchley et al., 2003; Valenza et al., 2019). Certains auteurs ont même proposé que la corrélation entre les aires cérébrales du réseau exécutif et le contrôle cognitif soient purement liées à la régulation du système nerveux autonome (Beissner et al., 2013), ce qui permettrait d'expliquer éventuellement le lien entre le dmPFC, l'insula antérieure, le contrôle cognitif et l'éveil.

Dans tous les cas, nos résultats confirment l'idée que le dmPFC pourrait jouer un rôle au niveau méta-décisionnel en déterminant la quantité de ressources allouée à la décision en cours et, plus généralement, la quantité de ressources à allouer à la tâche en cours d'exécution. Nos résultats ne permettent cependant pas de déterminer la nature de la ressource que le dmPFC serait éventuellement capable de réguler. Il pourrait donc aussi bien s'agir du système nerveux autonome, du système de contrôle cognitif ou même de l'activité musculaire puisqu'il est aussi corrélé à l'intensité d'une activité motrice.

### **III.2 Le dmPFC : la vigie du cerveau ?**

*« tous les bœufs et les chevaux et toutes les bêtes à l'envi témoignent du contraire par leur chasse à la jouissance ; le vulgaire s'y fie, comme les devins aux oiseaux, pour juger que les plaisirs sont les facteurs les plus puissants de la vie bonne, et regarde les amours des bêtes comme des témoins plus autorisés que ne sont les amours nourris aux intuitions rationnelles de la muse philosophique. »*

*67b, Philèbe, Platon (Platon, 2018)*

Depuis l'Antiquité, les philosophes considèrent qu'il existe une dualité entre les appétits animaux, autrement dit les désirs, et de l'autre un système rationnel orienté vers des buts. Les appétits animaux seraient mauvais parce qu'ils ramèneraient l'homme à la part qu'il a en commun avec la plupart des autres Mammifères, comme l'exprime Platon dans la citation ci-dessus via un discours de Socrate. Ces appétits amèneraient l'homme, comme dans le test du marshmallow (voir *L'aversion pour le délai*), à prendre des décisions sous-optimales sur le long

## Discussion générale

terme. À l'opposé, le système rationnel serait guidé par la philosophie, la prise en compte du futur et pas seulement du besoin présent. Cette vision duelle du cerveau humain parcourt depuis longtemps la philosophie occidentale et la philosophie orientale. On en retrouve des traces au moins depuis le V<sup>ème</sup> siècle avant Jésus-Christ. On retrouve en effet aussi bien chez Platon (V<sup>ème</sup> siècle av. JC), comme dans la citation ci-dessus, ou chez Aristote (IV<sup>ème</sup> siècle av. JC) (voir *Bases neurales du délai*) en Occident, que chez Confucius (VI<sup>ème</sup> siècle av. JC) en Orient à peu près à la même période, un ensemble de réflexions promouvant le contrôle de soi plutôt que d'obéir à ses appétits. Confucius aurait dit par exemple : « *Aimer à donner libre cours à ses convoitises, aimer à perdre son temps et à courir çà et là, aimer les festins et les plaisirs déshonnêtes, ces trois passions sont nuisibles.* » (Confucius and Lao-Tseu, 1973). Descartes, en France, au XVII<sup>ème</sup> siècle, dans son *Discours de la méthode*, prônera de même la maîtrise des désirs par la rationalité. Avec une vision dualiste du cerveau très similaire à celle de l'Antiquité, Daniel Kahneman, reprenant des termes introduits par Keith Stanovich et Richard West en psychologie (Stanovich and West, 2000), développera amplement l'idée qu'il existe deux systèmes dans le cerveau dénommés simplement le Système 1 et le Système 2 (Kahneman, 2012). Comme le titre de son livre l'indique, le Système 1 est censé être un système intuitif qui nous permet de répondre vite (*Thinking fast*) à la plupart des situations, alors que le Système 2 s'activerait de manière plus lente et avec une notion de rationalité (*Thinking slow*). Ce Système 2 irait de pair avec la sensation d'effort subjectif, une pupille plus dilatée ainsi que des temps de réponse plus longs et il permettrait de voir s'il est nécessaire d'allouer plus d'effort pour effectuer la tâche en cours. Daniel Kahneman a d'ailleurs participé à initier les travaux montrant que le diamètre pupillaire pouvait refléter la quantité d'effort mental en cours d'exécution (Kahneman and Beatty, 1966). Cette idée dualiste, profondément ancrée dans nos esprits, a mené à l'intuition qu'il pourrait exister différents systèmes d'évaluation qui agiraient en parallèle dans le cerveau, l'un qui agirait automatiquement en vue d'un plaisir immédiat et un autre avec une vision à plus long terme (McClure, 2004). Depuis, on considère plutôt qu'il existe un réseau capable de contrôle cognitif capable de réguler l'activité du réseau des valeurs, par exemple en considérant dans quelle mesure des aliments sont sains, en plus de considérer leur aspect gustatif (Hare et al., 2009, 2011a). Le contrôle cognitif est une capacité permettant justement de mettre en place un plan d'action en vue d'un but, par exemple le fait d'avoir une meilleure santé, et de réguler les automatismes du comportement. Dans nos études, nous avons vu que l'activité du dmPFC, ainsi que celle de l'insula antérieure qui lui est régulièrement associée, allait de pair avec une pupille plus dilatée (voir **Figure 28**) et des temps de délibération

## Discussion générale

plus importants (voir **Figure 27** et **Figure 38**), suggérant, en accord avec une littérature conséquente sur le sujet, qu'il était associé à la capacité de contrôle cognitif.

L'activité du dmPFC et de l'insula antérieure a de plus été associée à l'identification de situations liées à un évènement aversif, comme l'exécution d'un effort physique (Kurniawan et al., 2013) ou d'un effort mental (Alnaes et al., 2014; Botvinick et al., 1999, 2004; Collette and Salmon, 2014; Duverne and Koechlin, 2017; Koechlin et al., 2003; MacDonald et al., 2000; Niendam et al., 2012), ou le fait de subir une punition d'ordre physique (Craig et al., 1996; Hayes and Northoff, 2012; Krishnan et al., 2016; Wager et al., 2013) ou psychologique (Bartra et al., 2013; Eisenberger and Lieberman, 2004; Palminteri et al., 2012). Plus généralement, l'activité du dmPFC et de l'insula antérieure a été assimilée à un réseau central pour la détection des évènements saillants (Craig, 2009; Menon and Uddin, 2010; Seeley et al., 2007), qu'il s'agisse d'une saillance associée à l'éventualité d'un gain ou d'une perte élevée (Litt et al., 2011), ou à une surprise élevée (Amiez et al., 2005; Fouragnan et al., 2018; Sridharan et al., 2008). Les évènements saillants impliquent potentiellement de reconsidérer le plan d'action en cours car ils permettent justement de signaler des situations cruciales pour la survie. Dans le cas de la potentialité d'obtenir une récompense élevée ou de subir une punition élevée, cela paraît évident. Pour ce qui est de la surprise, un tel signal invite, *a minima*, à reconsidérer le modèle interne du monde et, selon la cause de cette surprise, à adapter, ou non, son comportement. Il est intéressant, par ailleurs, de constater que dans notre première étude où les items étaient fictifs, l'activité du dmPFC n'augmentait pas avec l'aspect appétitif des items de récompense ou avec l'aspect aversif des items d'efforts, mais avait au contraire même tendance à baisser. Une explication un peu spéculative, mais possible, pourrait être justement que, dans cette étude, les participants sachant que tous les items étaient purement fictifs, savaient qu'il n'y avait pas besoin d'investir du contrôle cognitif en réponse à la saillance des items puisqu'ils n'allaient pas les obtenir. Au contraire, dans notre seconde étude où la rémunération finale était liée à la performance des participants (voir **Figure 35**), l'activité du dmPFC et de l'insula antérieure avait tendance à augmenter avec les enjeux monétaires non-signés, peut-être justement en vue d'adapter le comportement en fonction des enjeux. En résumé, l'activité du dmPFC et de l'insula antérieure semble augmenter en réponse à des situations saillantes, en particulier lorsque celles-ci requièrent une adaptation comportementale. Il a d'ailleurs déjà été proposé que l'insula antérieure joue un rôle d'interrupteur en détectant d'un côté les situations saillantes, notamment grâce au fait qu'elle est liée à l'intéroception, et en facilitant de l'autre côté le recrutement des structures liées au contrôle cognitif (Menon and Uddin, 2010; Sridharan

et al., 2008). Le réseau constitué de l'insula antérieure et du dmPFC permettrait ainsi à la fois de détecter les situations nécessitant du contrôle cognitif, de calculer dans quelle mesure il est pertinent d'allouer du contrôle cognitif au vu de la situation et enfin d'allouer du contrôle cognitif en vue de mettre en place une réponse comportementale adaptée.

### III.3 Nature des ressources engagées

Nous avons déjà évoqué les différentes hypothèses sur la nature des ressources engagées par des activités nécessitant du contrôle cognitif et allant de pair avec une sensation d'effort mental (voir *De quoi l'aversion pour l'activité mentale nous protège-t-elle ?*). Nous n'avons pas spécialement cherché à en identifier la nature dans nos expériences, aussi nous est-il difficile d'en dire beaucoup plus à ce stade. Cependant, il est intéressant de remarquer que, dans nos tâches d'effort physique et d'effort mental, les mêmes aires cérébrales étaient activées en lien avec la variable dérivée de notre modèle qui était censée représenter la quantité de ressources allouée à chaque essai (voir **Figure 34**). De plus, lorsque nous utilisons un seuil plus conservatif, le seul cluster qui survivait était localisé dans le dlPFC droit. Ce résultat suggère, si l'on accepte que l'activité du dlPFC reflète l'exécution de contrôle cognitif, que du contrôle cognitif pourrait aussi être requis même dans des tâches principalement basées sur un effort physique. De plus, la question de savoir si le rôle du dmPFC est principalement de recruter l'activité du dlPFC, ou s'il peut aussi influencer d'autres systèmes comme le système nerveux autonome, permettra aussi de savoir dans quelle mesure les ressources que le dmPFC alloue sont uniques (l'activité du réseau exécutif) ou multiples et dépendantes de la situation envisagée. Dans tous les cas, nos résultats convergent vers l'idée que l'activité du dmPFC induit l'allocation de ressources en réponse à des situations saillantes, qu'il s'agisse de l'activité des aires liées au contrôle cognitif, d'un temps de délibération plus conséquent et/ou de l'activité du système nerveux autonome.

### III.4 Le dmPFC *versus* l'insula antérieure

#### III.4.a Délimitations neuroanatomiques du dmPFC et de l'insula antérieure

Le dmPFC a été impliqué dans une proportion considérable de concepts cognitifs. Cependant, les résultats ne sont pas toujours consistants entre différentes équipes (Kolling et al., 2016b; Shenhav et al., 2014). Il a été suggéré que les divisions sur le rôle fonctionnel du dmPFC proviennent de divergences anatomiques. Différentes études utiliseraient ainsi un nom similaire pour parler, en fait, de populations neurales différentes. Le contrôle cognitif et la difficulté seraient encodés dans une zone plus dorsale au cortex cingulaire alors que la valeur

## Discussion générale

d'explorer l'environnement serait située directement au sein du cortex cingulaire par exemple (Kolling et al., 2016b, 2016a). Cependant, plusieurs explications alternatives pourraient aussi expliquer ces divergences.

En effet, les études en IRMf sont toujours basées sur une étape de lissage des données afin de donner du sens aux tests statistiques qui sont opérés. La manière dont les données ont été lissées et les paramètres d'acquisition de l'IRMf peuvent ainsi déjà entraîner de légers décalages dans les coordonnées extraites entre deux études. De plus, il a été montré que les humains diffèrent, parfois même d'un hémisphère à l'autre de la même personne, au niveau de cette zone par la présence ou l'absence d'un sulcus : le sulcus para-cingulaire. En effet, si 100% des sujets possèdent généralement un sulcus cingulaire dans leurs deux hémisphères, la proportion de sujets possédant un sulcus para-cingulaire est de l'ordre de 60% dans l'hémisphère gauche et de 40% dans l'hémisphère droit (Amiez et al., 2018). Or, il a été montré que le fait de posséder ou non ce sulcus pouvait décaler la localisation des corrélations fonctionnelles au niveau de cette zone dans une tâche requérant du contrôle cognitif et de l'exploration (Amiez et al., 2013). En moyenne, les participants possédant ce sulcus avaient un pic d'activation décalé sur l'axe vertical de 11mm par rapport aux autres participants. La proportion de participants possédant ou non ce sulcus dans chaque hémisphère pourrait ainsi entraîner des différences dans les coordonnées moyennes liées à cette zone dans chaque étude. Cela pourrait expliquer en partie le fait que certaines études estiment que ce cluster est juste au-dessus ou juste en-dessous du cortex cingulaire. Nous n'avons d'ailleurs pas vérifié, au sein de chacune de nos études, la proportion des hémisphères de nos participants possédant ou non ce sulcus. Il serait intéressant de voir dans quelle mesure le fait de prendre ces différences anatomiques en compte à l'avenir permettra, ou non, d'améliorer la cartographie cérébrale et la correspondance entre les activations fonctionnelles et leur correspondance au niveau des structures anatomiques sous-jacentes. Il semble en tout cas raisonnable de se demander si les études sur cette zone qui ont permis de révéler des corrélations avec une diversité impressionnante de concepts peut s'expliquer par le fait que ces études ne parlent pas toutes de la même zone ou ne prennent pas assez compte des différences anatomiques des participants. Alternativement, comme nous venons de l'évoquer, il est aussi envisageable que des études pensant parler de régions distinctes parlent en fait de la même zone. Dans tous les cas, il reste intéressant de voir si un seul principe unique, comme la valeur attendue du contrôle cognitif, peut rendre compte de toute cette diversité indépendamment des biais qui peuvent exister dans la localisation spatiale de ce cluster fonctionnel. Dans nos trois études, nous avons pu observer

## Discussion générale

que le cluster fonctionnel que nous avons appelé le dmPFC était systématiquement corrélé à la notion d'allocation de ressources, qu'il s'agisse du temps de délibération, dans nos tâches de décision dans notre première et dans notre troisième étude, ou de la quantité de ressources inférées par notre modèle, dans notre seconde étude impliquant un effort physique et/ou mental. Ce résultat est d'ailleurs compatible aussi bien avec le fait que le dmPFC est plus actif pendant l'exploration par rapport à l'exploitation, avec la nécessité de mettre des modèles internes du monde à jour ou en lien avec le contrôle cognitif.

Par ailleurs, nous avons à plusieurs reprises évoqué l'insula antérieure, qui semble faire partie du même réseau que le dmPFC pour ce qui est du traitement des situations requérant l'allocation de ressources mentales. Il a cependant été montré que les analyses de groupe peuvent biaiser les résultats et donner l'impression qu'un cluster dont le pic est localisé dans l'operculum frontal peut parfois apparaître comme étant situé dans l'insula antérieure qui lui est adjacente (Amiez et al., 2016). Une meilleure prise en compte des distinctions entre l'operculum frontal et l'insula antérieure à l'avenir permettra de voir si ces deux aires adjacentes sont réellement dissociables au niveau fonctionnel.

### III.4.b Rôles respectifs de chacune des zones

Nous avons vu que le dmPFC était systématiquement co-activé avec l'insula antérieure dans nos études, que ce soit en lien avec le temps de délibération dans notre première et notre troisième étude, avec l'erreur de prédiction non-signée dans notre troisième étude, ou encore avec la quantité de ressources à allouer dans la deuxième étude (voir **Table 1** et **Table 2**). Il paraît de ce fait légitime de se demander si le dmPFC et l'insula antérieure ont des rôles redondants, ou si ces deux aires ont des rôles différents dans l'évaluation de situations nécessitant d'allouer des ressources physiques ou mentales.

Dans la littérature sur la douleur, nous avons vu que ces deux structures semblaient avoir des rôles distincts (voir *Bases neurales des punitions*). Stimuler l'activité de l'insula antérieure semble en effet directement éliciter une sensation de douleur (Afif et al., 2010; Frot and Mauguière, 2003; Isnard et al., 2004; Mazzola et al., 2012, 2017, 2019; Stephani et al., 2011). Au contraire, stimuler le dmPFC ne semble pas provoquer de sensation de douleur (Caruana et al., 2018; Hutchison et al., 1999). De plus, l'insula antérieure est régulièrement considérée comme une zone réagissant à la saillance d'événements externes, indépendamment de leur valence (Litt et al., 2011; Wiech et al., 2010). Elle est aussi présentée comme une zone ayant un rôle prépondérant dans l'intéroception, c'est-à-dire le suivi permanent de l'état des



## Conclusion

différentes parties du corps (Craig, 2002, 2009), elle permettrait ainsi de signaler des situations requérant d'agir. Le dmPFC lui servirait alors à initier une chaîne d'actions en fonction de la situation (Craig, 2009). On constate d'ailleurs que des stimulations autour du dmPFC provoquent des schémas moteurs comme en réponse à une menace dans l'environnement (Caruana et al., 2018).

L'insula antérieure pourrait ainsi identifier les situations importantes pour la survie, en particulier dans le domaine aversif. Elle semble d'ailleurs idéale pour ce rôle puisque, comme nous venons de le voir, elle encode aussi bien les événements externes au corps qui pourraient avoir un impact sur la survie, via sa réaction aux événements saillants, que les événements internes en surveillant, via l'intéroception, l'état des différentes parties du corps. L'activité de l'insula antérieure pourrait ensuite entraîner l'activité du dmPFC en vue de mettre en place du contrôle cognitif et une réponse comportementale adaptée en vue de réagir à la situation identifiée par l'insula antérieure.

## Chapitre 4: Conclusion

Dans cette thèse, nous avons étudié les bases cérébrales du compromis coûts/bénéfices et des structures responsables de la régulation de ce compromis. En ce qui concerne l'estimation du compromis coûts/bénéfices, nous avons pu confirmer le rôle prépondérant du réseau des valeurs, un réseau constitué du cortex préfrontal ventromédian (vmPFC) et du striatum ventral, dans le calcul du compromis coûts/bénéfices. Plus précisément, nous avons vu que :

- Le vmPFC représentait la valeur subjective issue du compromis coûts/bénéfices de manière **stable** au travers de pas moins de quatre paradigmes différents. Ces paradigmes, dont trois ont été effectués par les mêmes participants, impliquaient une tâche d'évaluation, deux tâches de choix basés sur la valeur et une tâche d'apprentissage par renforcement. Cette représentation était aussi stable indépendamment du mode de présentation, qu'il s'agisse d'un affichage d'une récompense sous forme de texte ou d'un affichage d'une récompense sous forme de texte accompagnée d'une image comme dans notre première étude, ou encore d'un symbole associé à la valeur comme dans notre troisième étude. Le striatum était aussi impliqué dans l'estimation d'un compromis coûts/bénéfices en vue d'estimer la quantité optimale de ressources à investir dans des tâches impliquant une performance ;

## Conclusion

- Nous avons aussi découvert que **le vmPFC était capable d'estimer le coût associé à des efforts** alors que plusieurs chercheurs pensaient qu'il n'est capable que d'encoder la valeur de biens, mais pas d'actions ;
- Nous avons aussi constaté que lorsque ce réseau était plus actif, la **vigueur**, c'est-à-dire l'intensité et la rapidité, des efforts physiques et mentaux qui sont fournis en réponse à des enjeux croissants augmentait. Dans la plupart de nos études, même lorsque cela n'était pas instrumental, les participants avaient tendance à répondre plus rapidement lorsque ce réseau était plus actif. Dans notre seconde étude, nous avons constaté que l'activité du striatum était plus élevée en réponse à des gains par rapport à des pertes équivalentes et que cette différence était aussi reflétée dans la performance. Bien que nos études soient purement corrélationnelles et ne permettent pas de faire d'inférence sur le rôle causal des aires du réseau des valeurs, nos résultats suggèrent qu'elles permettent d'influencer le comportement en vue de faciliter l'obtention de récompenses et l'évitement de punitions.

Prendre une décision implique à la fois d'effectuer un compromis entre les bénéfices et les coûts associés à une action, mais aussi d'effectuer une méta-décision sur les coûts et les bénéfices que l'on est prêt à investir dans ce compromis. Dépenser plus de temps et d'effort mental au cours d'une prise de décision doit aider à augmenter le bénéfice lié à la décision : **la confiance**. En ce qui concerne la régulation du compromis coûts/bénéfices, nous avons ainsi vu que :

- Il était possible de **dissocier l'encodage de la confiance**, c'est-à-dire les bénéfices associés au calcul du compromis coûts/bénéfices, **de l'encodage de la valeur** issue du compromis coûts/bénéfices. Dans notre première comme dans notre troisième étude, le **vmPFC**, une zone liée à l'encodage du compromis coûts/bénéfices comme nous venons de le répéter, était corrélé avec la confiance. Ce résultat est en accord avec une littérature conséquente montrant qu'il est impliqué à la fois dans le calcul du compromis coûts/bénéfices et dans le calcul de la confiance de ce calcul. Cependant, notre première étude nous a permis de mettre en évidence qu'il était possible d'aussi identifier un autre cluster, le **mPFC**, principalement associé à l'encodage de la **confiance** dans la décision. Ce résultat était d'ailleurs **stable** au travers des trois tâches de notre première étude ;

## Conclusion

- Les **coûts liés à la prise de décision**, en particulier le temps passé à délibérer, étaient systématiquement corrélés à l'activité d'une autre aire cérébrale, le **dmPFC**, au sein de notre première et de notre troisième étude ;
- Nous avons pu voir que, même une tâche relativement simple d'**apprentissage par renforcement** qui est effectuée par la plupart des Mammifères de manière quotidienne, impliquait l'activité du dmPFC. Ce résultat suggérant que, chez l'homme au moins, ce type d'activité n'est pas nécessairement automatique et peut entraîner une délibération et la dépense d'un effort mental prenant la forme de **contrôle cognitif** ;
- L'activité du dmPFC pourrait refléter le fait d'effectuer un effort mental, ou plus généralement le fait **d'allouer des ressources à la tâche en cours**, qu'il s'agisse de ressources sous la forme de l'activité du système nerveux autonome, d'activité musculaire, de temps, ou du fait d'effectuer du contrôle cognitif. Dans notre deuxième étude, nous avons d'ailleurs pu voir que l'activité du dmPFC était à la fois corrélée au montant de ressources, inféré par notre modèle computationnel, investi dans le cadre d'une tâche d'effort mental, qu'à celui investi dans le cadre d'une tâche d'effort physique suggérant soit que notre tâche d'effort physique impliquait aussi du contrôle cognitif, soit que les ressources que le dmPFC est capable de recruter sont en effet plus larges.

Il faut noter que toutes nos études sont purement basées sur des corrélations entre l'activité cérébrale et les concepts cognitifs qui nous intéressaient. De ce fait, des approches complémentaires sont nécessaires afin de tester si les aires cérébrales que nous avons identifiées jouent bien un rôle causal dans la représentation mentale des concepts auxquels nous nous sommes intéressés ici. Des projets sont d'ailleurs en cours actuellement dans notre équipe afin de voir dans quelle mesure des patients atteints de lésions en différentes zones du cerveau se comportent dans des tâches similaires aux nôtres. Les résultats que nous venons de présenter suggèrent par exemple qu'une lésion dans le dmPFC induirait une incapacité dans le fait d'allouer des ressources neurales à la prise de décision. Cela amènerait la personne affectée d'une lésion au dmPFC à prendre des décisions hâtivement, sans prendre le temps de délibérer, avec une confiance résultante potentiellement aussi plus faible dans la décision effectuée. Alternativement, si cette zone reflète uniquement la pénibilité de l'effort effectué sans avoir de rôle causal dans la quantité de ressources allouées pour la décision, on s'attendrait plutôt à ce

## Bibliographie

que l'effort lié à la délibération ne paraisse pas pénible, tout en étant équivalent à celui effectué par des sujets sains.

Nos résultats permettent de mieux comprendre comment s'effectue le compromis coûts/bénéfices, autant au niveau comportemental qu'au niveau cérébral. Au niveau comportemental, nous avons par exemple montré des différences dans le traitement des gains par rapport aux pertes dans des tâches de performance avec des enjeux modérés. Au niveau cérébral, nous avons réussi à mieux caractériser le fonctionnement du réseau des valeurs qui est associée au calcul du compromis coûts/bénéfices, ainsi que le fonctionnement des aires cérébrales associées à la régulation de ce calcul. Nous avons en effet démontré que le mPFC pouvait être lié aux bénéfices attendus de ce calcul, un gain de confiance, et que le dmPFC pouvait être associé aux coûts de ce calcul, comme le temps passé à délibérer. Avec ces études, nous espérons avoir ainsi à la fois ouvert une nouvelle approche visant à regarder les corrélats neuraux stables d'un concept au travers de plusieurs tâches effectuées par les mêmes participants, ainsi qu'avoir permis d'augmenter notre connaissance du fonctionnement du compromis coûts/bénéfices.

## Bibliographie

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., and Pessiglione, M. (2015). Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. *Journal of Neuroscience* 35, 2308–2320.

Afif, A., Minotti, L., Kahane, P., and Hoffmann, D. (2010). Anatomofunctional organization of the insular cortex: A study using intracerebral electrical stimulation in epileptic patients. *Epilepsia* 51, 2305–2315.

Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica* 21, 503–546.

Alnaes, D., Sneve, M.H., Espeseth, T., Endestad, T., van de Pavert, S.H.P., and Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision* 14, 1–1.

Amiez, C., Joseph, J.-P., and Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *European Journal of Neuroscience* 21, 3447–3452.

Amiez, C., Sallet, J., Procyk, E., and Petrides, M. (2012). Modulation of feedback related activity in the rostral anterior cingulate cortex during trial and error exploration. *NeuroImage* 63, 1078–1090.

## Bibliographie

- Amiez, C., Neveu, R., Warrot, D., Petrides, M., Knoblauch, K., and Procyk, E. (2013). The Location of Feedback-Related Activity in the Midcingulate Cortex Is Predicted by Local Morphology. *Journal of Neuroscience* *33*, 2217–2228.
- Amiez, C., Wutte, M.G., Faillenot, I., Petrides, M., Burle, B., and Procyk, E. (2016). Single subject analyses reveal consistent recruitment of frontal operculum in performance monitoring. *NeuroImage* *133*, 266–278.
- Amiez, C., Wilson, C.R.E., and Procyk, E. (2018). Variations of cingulate sulcal organization and link with cognitive performance. *Sci Rep* *8*, 13988.
- Anderson, B.A. (2016). Value-driven attentional capture in the auditory domain. *Atten Percept Psychophys* *78*, 242–250.
- Anderson, B.A. (2019). Neurobiology of value-driven attention. *Current Opinion in Psychology* *29*, 27–33.
- Anderson, M.L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behav Brain Sci* *33*, 245–266.
- Anderson, B.A., Laurent, P.A., and Yantis, S. (2011). Value-driven attentional capture. *Proc. Natl. Acad. Sci. U.S.A.* *108*, 10367–10371.
- Andrade, L.F., and Petry, N.M. (2012). Delay and probability discounting in pathological gamblers with and without a history of substance use problems. *Psychopharmacology (Berl.)* *219*, 491–499.
- Apps, M.A.J., Grima, L.L., Manohar, S., and Husain, M. (2015). The role of cognitive effort in subjective reward devaluation and risky decision-making. *Sci Rep* *5*, 16880.
- Aridan, N., Pelletier, G., Fellows, L.K., and Schonberg, T. (2019). Is ventromedial prefrontal cortex critical for behavior change without external reinforcement? *Neuropsychologia* *124*, 208–215.
- Aristote (1969). *De l'âme*.
- Aristote (2004). *Ethique à Nicomaque* (GF-Flammarion).
- Arnsten, A.F., and Goldman-Rakic, P.S. (1984). Selective prefrontal cortical projections to the region of the locus coeruleus and raphe nuclei in the rhesus monkey. *Brain Res.* *306*, 9–18.
- Arulpragasam, A.R., Cooper, J.A., Nuutinen, M.R., and Treadway, M.T. (2018). Corticoinsular circuits encode subjective value expectation and violation for effortful goal-directed behavior. *Proceedings of the National Academy of Sciences* *115*, E5233–E5242.
- Aston-Jones, G., and Waterhouse, B. (2016). Locus coeruleus: From global projection system to adaptive regulation of behavior. *Brain Res.* *1645*, 75–78.
- Averbeck, B.B., and Costa, V.D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience* *20*, 505–512.

## Bibliographie

- Baliki, M.N., Geha, P.Y., and Apkarian, A.V. (2009). Parsing Pain Perception Between Nociceptive Representation and Magnitude Estimation. *J Neurophysiol* *101*, 875–887.
- Balleine, B.W., Daw, N.D., and O’Doherty, J.P. (2009). Multiple forms of value learning and the function of dopamine. *Neuroeconomics* 367–387.
- Ballesta, S., and Padoa-Schioppa, C. (2019). Economic Decisions through Circuit Inhibition. *Current Biology* *29*, 3814–3824.e5.
- Bandettini, P.A., Wong, E.C., Hinks, R.S., Tikofsky, R.S., and Hyde, J.S. (1992). Time course EPI of human brain function during task activation. *Magn. Reson. Med.* *25*, 390–397.
- Bang, D., and Fleming, S.M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. *Proceedings of the National Academy of Sciences* *115*, 6082–6087.
- Barch, D.M., Braver, T.S., Nystrom, L.E., Forman, S.D., Noll, D.C., and Cohen, J.D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia* *35*, 1373–1380.
- Barlow, H.B. (2012). Possible Principles Underlying the Transformations of Sensory Messages. In *Sensory Communication*, W.A. Rosenblith, ed. (The MIT Press), pp. 216–234.
- Barron, G., and Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making* *16*, 215–233.
- Bartra, O., McGuire, J.T., and 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.
- Basten, U., Biele, G., Heekeren, H.R., and Fiebach, C.J. (2010). How the brain integrates costs and benefits during decision making. *PNAS* *107*, 21767–21772.
- Bastin, J., Deman, P., David, O., Gueguen, M., Benis, D., Minotti, L., Hoffman, D., Combrisson, E., Kujala, J., Perrone-Bertolotti, M., et al. (2016). Direct Recordings from Human Anterior Insula Reveal its Leading Role within the Error-Monitoring Network. *Cerebral Cortex* bhv352.
- Baumeister, R.F., Bratslavsky, E., Muraven, M., and Tice, D.M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology* *74*, 1252–1265.
- Bechara, A., Damasio, A.R., Damasio, H., and Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* *50*, 7–15.
- Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., and Damasio, A.R. (1995). Double Dissociation of Conditioning and Declarative Knowledge Relative to the Amygdala and Hippocampus in Humans. *269*, 5.
- Bégonne-Bret, C. (2011). *Le désir. De Platon à Sartre* (Eyrolles).
- Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience* *10*, 1214–1221.

## Bibliographie

Beissner, F., Meissner, K., Bär, K.-J., and Napadow, V. (2013). The Autonomic Brain: An Activation Likelihood Estimation Meta-Analysis for Central Processing of Autonomic Function. *J. Neurosci.* *33*, 10503–10511.

Békés, V., Perry, J.C., and Robertson, B.M. (2018). Psychological masochism: A systematic review of the literature on conflicts, defenses, and motives. *Psychother Res* *28*, 470–483.

van Belle, J., Vink, M., Durston, S., and Zandbelt, B.B. (2014). Common and unique neural networks for proactive and reactive response inhibition revealed by independent component analysis of functional MRI data. *NeuroImage* *103*, 65–74.

Benarroch, E.E. (2018). Locus coeruleus. *Cell Tissue Res.* *373*, 221–232.

Benoit, C.-E., Solopchuk, O., Borragán, G., Carbonnelle, A., Van Durme, S., and Zénon, A. (2019). Cognitive task avoidance correlates with fatigue-induced performance decrement but not with subjective fatigue. *Neuropsychologia* *123*, 30–40.

Bergeron, V. (2007). Anatomical and Functional Modularity in Cognitive Science: Shifting the Focus. *Philosophical Psychology* *20*, 175–195.

Bernoulli, D. (1738). *Specimen Theoriae Novae de Mensura Sortis*.

Berridge, C.W., and Spencer, R.C. (2017). Noradrenergic Control of Arousal and Stress. In *Stress: Neuroendocrinology and Neurobiology*, (Elsevier), pp. 39–47.

Berridge, K.C., and Kringelbach, M.L. (2015). Pleasure Systems in the Brain. *Neuron* *86*, 646–664.

Berridge, K.C., and Valenstein, E.S. (1991). What Psychological Process Mediates Feeding Evoked by Electrical Stimulation of the Lateral Hypothalamus? *Behavioral Neuroscience* *105*, 12.

Berridge, K.C., Robinson, T.E., and Aldridge, J.W. (2009). Dissecting components of reward: ‘liking’, ‘wanting’, and learning. *Current Opinion in Pharmacology* *9*, 65–73.

Białaszek, W., Marcowski, P., and Ostaszewski, P. (2017). Physical and cognitive effort discounting across different reward magnitudes: Tests of discounting models. *PLOS ONE* *12*, e0182353.

Bijleveld, E. (2018). The feeling of effort during mental activity. *Consciousness and Cognition* *63*, 218–227.

Bioud, E. (2019). Interferences between decisional variables. Sorbonne Université.

Blain, B., Hollard, G., and Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. *Proceedings of the National Academy of Sciences* *113*, 6967–6972.

Blain, B., Schmit, C., Aubry, A., Hausswirth, C., Le Meur, Y., and Pessiglione, M. (2019). Neuro-computational Impact of Physical Training Overload on Economic Decision-Making. *Current Biology* *29*, 3289-3297.e4.

## Bibliographie

- Bludau, S., Eickhoff, S.B., Mohlberg, H., Caspers, S., Laird, A.R., Fox, P.T., Schleicher, A., Zilles, K., and Amunts, K. (2014). Cytoarchitecture, probability maps and functions of the human frontal pole. *NeuroImage* 93, 260–275.
- Boccard, S.G.J., Fitzgerald, J.J., Pereira, E.A.C., Moir, L., Van Hartevelt, T.J., Kringelbach, M.L., Green, A.L., and Aziz, T.Z. (2014). Targeting the Affective Component of Chronic Pain. *Neurosurgery* 74, 628–637.
- Boldt, A., Blundell, C., and De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neurosci Conscious* 2019, niz004.
- Boly, M., Baeteau, E., Schnakers, C., Degueldre, C., Moonen, G., Luxen, A., Phillips, C., Peigneux, P., Maquet, P., and Laureys, S. (2007). Baseline brain activity fluctuations predict somatosensory perception in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12187–12192.
- Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., and Rushworth, M.F.S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. *Neuron* 62, 733–743.
- Borg, G. (1970). Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med* 2, 92–98.
- Borg, G. (1978). Subjective Aspects of Physical and Mental Load. *Ergonomics* 21, 215–220.
- Borg, G. a. V. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise* 14, 377–381.
- Borg, G., and Linderholm, H. (1970). Exercise performance and perceived exertion in patients with coronary insufficiency, arterial hypertension and vasoregulatory asthenia. *Acta Med Scand* 187, 17–26.
- Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., and Cohen, J.D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. 402, 3.
- Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., and Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychol Rev* 108, 624–652.
- Botvinick, M.M., Cohen, J.D., and Carter, C.S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences* 8, 539–546.
- Botvinick, M.M., Huffstetler, S., and McGuire, J.T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective, & Behavioral Neuroscience* 9, 16–27.
- Boureau, Y.-L., and Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36, 74–97.
- Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., and Noll, D.C. (1997). A Parametric Study of Prefrontal Cortex Involvement in Human Working Memory. *NeuroImage* 5, 49–62.
- Bréal, M. (1897). *Essai de Sémantique : (science des significations)* (Paris : Hachette).



## Bibliographie

- Breiter, H.C., Berke, J.D., Kennedy, W.A., Rosen, B.R., and Hyman, S.E. (1996). Activation of striatum and amygdala during reward conditioning: An fMRI study. *NeuroImage* 3, S220.
- Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses. *Neuron* 30, 619–639.
- Briggs, R.G., Conner, A.K., Sali, G., Rahimi, M., Baker, C.M., Burks, J.D., Glenn, C.A., Battiste, J.D., and Sughrue, M.E. (2018a). A Connectomic Atlas of the Human Cerebrum—Chapter 17: Tractographic Description of the Cingulum. *Operative Neurosurgery* 15, S462–S469.
- Briggs, R.G., Conner, A.K., Baker, C.M., Burks, J.D., Glenn, C.A., Sali, G., Battiste, J.D., O’Donoghue, D.L., and Sughrue, M.E. (2018b). A Connectomic Atlas of the Human Cerebrum—Chapter 18: The Connectional Anatomy of Human Brain Networks. *Operative Neurosurgery* 15, S470–S480.
- Bromiley, P., and Curley, S.P. (1992). Individual differences in risk-taking. In *Risk Taking Behavior*, (John Wiley & Sons Ltd), pp. 87–132.
- Brown, T.E., and Landgraf, J.M. (2010). Improvements in executive function correlate with enhanced performance and functioning and health-related quality of life: evidence from 2 large, double-blind, randomized, placebo-controlled trials in ADHD. *Postgrad Med* 122, 42–51.
- Buckner, R.L., and DiNicola, L.M. (2019). The brain’s default network: updated anatomy, physiology and evolving insights. *Nat Rev Neurosci* 20, 593–608.
- Burke, C.J., Brünger, C., Kahnt, T., Park, S.Q., and Tobler, P.N. (2013). Neural Integration of Risk and Effort Costs by the Frontal Pole: Only upon Request. *J. Neurosci.* 33, 1706–1713.
- Bush, R.R., and Mosteller, F. (1951). A mathematical model for simple learning. *The Psychological Review* 58, 11.
- Cacciaglia, R., Pohlack, S.T., Flor, H., and Nees, F. (2015). Dissociable roles for hippocampal and amygdalar volume in human fear conditioning. *Brain Struct Funct* 220, 2575–2586.
- Cacioppo, J.T., and Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology* 42, 116–131.
- Cai, X., Kim, S., and Lee, D. (2011). Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* 69, 170–182.
- Camille, N., Griffiths, C.A., Vo, K., Fellows, L.K., and Kable, J.W. (2011). Ventromedial Frontal Lobe Damage Disrupts Value Maximization in Humans. *Journal of Neuroscience* 31, 7527–7532.
- Camilleri, A.R., and Newell, B.R. (2013). Mind the gap? Description, experience, and the continuum of uncertainty in risky choice. In *Progress in Brain Research*, (Elsevier), pp. 55–71.
- Caplin, A., and Glimcher, P.W. (2014). Chapter 1 - Basic Methods from Neoclassical Economics. In *Neuroeconomics (Second Edition)*, P.W. Glimcher, and E. Fehr, eds. (San Diego: Academic Press), pp. 3–17.

## Bibliographie

- Carlén, M. (2017). What constitutes the prefrontal cortex? *Science* 358, 478–482.
- Carlezon, W.A., and Wise, R.A. (1996). Rewarding Actions of Phencyclidine and Related Drugs in Nucleus Accumbens Shell and Frontal Cortex. *J. Neurosci.* 16, 3112–3122.
- Carroll, L. (2009). *Through the Looking-Glass* (Oxford University Press: New York).
- Carroll, L. (2017). *Alice de l'autre côté du miroir* (Soleil).
- Carter, C.S., and van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience* 7, 367–379.
- Carter, E.C., and McCullough, M.E. (2014). Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated? *Front. Psychol.* 5.
- Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., and Cohen, J.D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280, 747–749.
- Carter, C.S., Macdonald, A.M., Botvinick, M., Ross, L.L., Stenger, V.A., Noll, D., and Cohen, J.D. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *PNAS* 97, 1944–1948.
- Caruana, F., Gerbella, M., Avanzini, P., Gozzo, F., Pelliccia, V., Mai, R., Abdollahi, R.O., Cardinale, F., Sartori, I., Lo Russo, G., et al. (2018). Motor and emotional behaviours elicited by electrical stimulation of the human cingulate cortex. *Brain* 141, 3035–3051.
- Casey, B.J., Thomas, K.M., Welsh, T.F., Badgaiyan, R.D., Eccard, C.H., Jennings, J.R., and Crone, E.A. (2000). Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. *PNAS* 97, 8728–8733.
- Cechetto, D.F. (2014). Cortical control of the autonomic nervous system: Cortical autonomic control. *Experimental Physiology* 99, 326–331.
- Chatham, C.H., and Badre, D. (2015). Multiple gates on working memory. *Current Opinion in Behavioral Sciences* 1, 23–31.
- Chiavaras, M.M., and Petrides, M. (2000). Orbitofrontal sulci of the human and macaque monkey brain. *Journal of Comparative Neurology* 422, 35–54.
- Chiavaras, M.M., LeGoualher, G., Evans, A., and Petrides, M. (2001). Three-Dimensional Probabilistic Atlas of the Human Orbitofrontal Sulci in Standardized Stereotaxic Space. *NeuroImage* 13, 479–496.
- Chib, V.S., Rangel, A., Shimojo, S., and O'Doherty, J.P. (2009). Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *Journal of Neuroscience* 29, 12315–12320.
- Chib, V.S., De Martino, B., Shimojo, S., and O'Doherty, J.P. (2012). Neural Mechanisms Underlying Paradoxical Performance for Monetary Incentives Are Driven by Loss Aversion. *Neuron* 74, 582–594.

## Bibliographie

- Choe, K.W., Jenifer, J.B., Rozek, C.S., Berman, M.G., and Beilock, S.L. (2019). Calculated avoidance: Math anxiety predicts math avoidance in effort-based decision-making. *Sci Adv* 5, eaay1062.
- Choi, E.Y., Yeo, B.T.T., and Buckner, R.L. (2012). The organization of the human striatum estimated by intrinsic functional connectivity. *Journal of Neurophysiology* 108, 2242–2263.
- Chong, T.T.-J., Apps, M., Giehl, K., Sillence, A., Grima, L.L., and Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biol* 15, e1002598.
- Christie, S.T., and Schrater, P. (2015). Cognitive cost as dynamic allocation of energetic resources. *Front. Neurosci.* 9.
- Christopoulos, G.I., Tobler, P.N., Bossaerts, P., Dolan, R.J., and Schultz, W. (2009). Neural Correlates of Value, Risk, and Risk Aversion Contributing to Decision Making under Risk. *Journal of Neuroscience* 29, 12574–12583.
- Chua, E.F., Schacter, D.L., Rand-Giovannetti, E., and Sperling, R.A. (2006). Understanding metamemory: neural correlates of the cognitive process and subjective level of confidence in recognition memory. *Neuroimage* 29, 1150–1160.
- Chua, E.F., Schacter, D.L., and Sperling, R.A. (2009). Neural correlates of metamemory: a comparison of feeling-of-knowing and retrospective confidence judgments. *J Cogn Neurosci* 21, 1751–1765.
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 362, 1585–1599.
- Cisek, P. (2012). Making decisions through a distributed consensus. *Current Opinion in Neurobiology* 22, 927–936.
- Cisek, P. (2019). Resynthesizing behavior through phylogenetic refinement. *Atten Percept Psychophys* 81, 2265–2287.
- Cisek, P., Puskas, G.A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* 29, 11560–11571.
- Clithero, J.A., and Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience* 9, 1289–1302.
- Cohen, A.R.S., Ezra Wolfe, Donald M. (1955). An experimental investigation of need for cognition. *Journal of Abnormal and Social Psychology* 51, 291–294.
- Cohen, J.D., Perlstein, W.M., Braver, T.S., Nystrom, L.E., Noll, D.C., Jonides, J., and Smith, E.E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608.
- Colizoli, O., de Gee, J.W., Urai, A.E., and Donner, T.H. (2018). Task-evoked pupil responses reflect internal belief states. *Sci Rep* 8, 13702.

## Bibliographie

Collette, F., and Salmon, É. (2014). Fonctionnement exécutif et réseaux cérébraux. *Revue de neuropsychologie* *Volume 6*, 256–266.

Conen, K.E., and Padoa-Schioppa, C. (2018). Neuronal Adaptation to the Value Range in the Macaque Orbitofrontal Cortex. *BioRxiv*.

Confucius, and Lao-Tseu (1973). *Confucius lao tseu* (Paris-Bievres).

Cooper, J.C., and Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *NeuroImage* *39*, 538–547.

Cooper, J.C., Hollon, N.G., Wimmer, G.E., and Knutson, B. (2009). Available alternative incentives modulate anticipatory nucleus accumbens activation. *Soc Cogn Affect Neurosci* *4*, 409–416.

Córcoles-Parada, M., Müller, N.C.J., Ubero, M., Serrano-del-Pueblo, V.M., Mansilla, F., Marcos-Rabal, P., Artacho-Pérula, E., Dresler, M., Insausti, R., Fernández, G., et al. (2017). Anatomical segmentation of the human medial prefrontal cortex: CÓRCOLES-PARADA et al. *J Comp Neurol* *525*, 2376–2393.

Coull, J.T., Cheng, R.-K., and Meck, W.H. (2011). Neuroanatomical and Neurochemical Substrates of Timing. *Neuropsychopharmacol* *36*, 3–25.

Courtney, S.M., Ungerleider, L.G., Keil, K., and Haxby, J.V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature* *386*, 608–611.

Craig, A.D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* *3*, 655–666.

Craig, A.D. (2009). How do you feel—now? the anterior insula and human awareness. *Nature Reviews Neuroscience* *10*.

Craig, A.D., Reiman, E.M., Evans, A., and Bushnell, M.C. (1996). Functional imaging of an illusion of pain. *Nature* *384*, 258–260.

Cramer, S.C., Weisskoff, R.M., Schaechter, J.D., Nelles, G., Foley, M., Finklestein, S.P., and Rosen, B.R. (2002). Motor cortex activation is related to force of squeezing. *Human Brain Mapping* *16*, 197–205.

Critchley, H.D., Mathias, C.J., Josephs, O., O’Doherty, J., Zanini, S., Dewar, B.-K., Cipolotti, L., Shallice, T., and Dolan, R.J. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain* *126*, 2139–2152.

Damasio, A.R. (2010). *L’erreur de Descartes: La raison des émotions* (Odile Jacob).

Danziger, S., Levav, J., and Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proc. Natl. Acad. Sci. U.S.A.* *108*, 6889–6892.

Danziger, S., Levav, J., and Avnaim-Pesso, L. (2015). « Qu’a mangé le juge à son petit-déjeuner ? » De l’impact des conditions de travail sur la décision de justice. *Les Cahiers de la Justice* *N° 4*, 579–587.

## Bibliographie

- Daunizeau, J., Adam, V., and Rigoux, L. (2014). VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. *PLoS Comput Biol* *10*, e1003441.
- Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* *8*, 1704–1711.
- Dayan, P. (2008). The role of value systems in decision making. In *Better than Conscious? Decision Making, the Human Mind, and Implications for Institutions*, (Cambridge, MA, US: MIT Press), pp. 51–70.
- De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Confidence in value-based choice. *Nature Neuroscience* *16*, 105–110.
- De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., and Love, B.C. (2017). Social Information Is Integrated into Value and Confidence Judgments According to Its Reliability. *The Journal of Neuroscience* *37*, 6066–6074.
- Dehaene, S., and Cohen, L. (2007). Cultural Recycling of Cortical Maps. *Neuron* *56*, 384–398.
- Deichmann, R., Gottfried, J.A., Hutton, C., and Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage* *19*, 430–441.
- Desender, K., Boldt, A., Verguts, T., and Donner, T.H. (2019). Confidence predicts speed-accuracy tradeoff for subsequent decisions. *Elife* *8*.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology* *64*, 135–168.
- Ditterich, J. (2006). Evidence for time-variant decision making. *European Journal of Neuroscience* *24*, 3628–3641.
- Domsch, S., Linke, J., Heiler, P.M., Kroll, A., Flor, H., Wessa, M., and Schad, L.R. (2013). Increased BOLD sensitivity in the orbitofrontal cortex using slice-dependent echo times at 3 T. *Magnetic Resonance Imaging* *31*, 201–211.
- Drugowitsch, J., Moreno-Bote, R., Churchland, A.K., Shadlen, M.N., and Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *J. Neurosci.* *32*, 3612–3628.
- Dudman, J.T., and Krakauer, J.W. (2016). The basal ganglia: from motor commands to the control of vigor. *Curr. Opin. Neurobiol.* *37*, 158–166.
- Durston, S. (2003). Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *NeuroImage* *20*, 2135–2141.
- Dutilh, G., and Rieskamp, J. (2016). Comparing perceptual and preferential decision making. *Psychon Bull Rev* *23*, 723–737.
- Duverne, S., and Koechlin, E. (2017). Rewards and Cognitive Control in the Human Prefrontal Cortex. *Cerebral Cortex* *27*, 5024–5039.
- Eisenberger, R. (1992). Learned industriousness. *Psychol Rev* *99*, 248–267.

## Bibliographie

Eisenberger, N.I., and Lieberman, M.D. (2004). Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in Cognitive Sciences* 8, 294–300.

Eldar, E., Rutledge, R.B., Dolan, R.J., and Niv, Y. (2016). Mood as Representation of Momentum. *Trends in Cognitive Sciences* 20, 15–24.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal Of Economics* 75, 643–669.

Engström, M., Karlsson, T., Landtblom, A.-M., and Craig, A.D. (Bud) (2015). Evidence of Conjoint Activation of the Anterior Insular and Cingulate Cortices during Effortful Tasks. *Frontiers in Human Neuroscience* 8.

Epictète (2007). *Manuel d'Epictète* (GF-Flammarion).

Ert, E., and Erev, I. (2007). *On the Descriptive Value of Loss Aversion in Decisions under Risk: Six Clarifications* (Rochester, NY: Social Science Research Network).

Fawcett, T.W., McNamara, J.M., and Houston, A.I. (2012). When is it adaptive to be patient? A general framework for evaluating delayed rewards. *Behavioural Processes* 89, 128–136.

Feinstein, J.S. (2013). Lesion studies of human emotion and feeling. *Current Opinion in Neurobiology* 23, 304–309.

Ferenczi, E.A., Zalocusky, K.A., Liston, C., Grosenick, L., Warden, M.R., Amatya, D., Katovich, K., Mehta, H., Patenaude, B., Ramakrishnan, C., et al. (2016). Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. *Science* 351, aac9698–aac9698.

Ferguson, H.J., Tresh, M., and Leblond, J. (2013). Examining mental simulations of uncertain events. *Psychon Bull Rev* 20, 391–399.

Ferrero, G. (1894). L'INERTIE MENTALE ET LA LOI DU MOINDRE EFFORT. *Revue Philosophique de La France et de l'Étranger* 37, 169–182.

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., and Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nat. Neurosci.* 22, 2066–2077.

FitzGerald, T.H.B., Friston, K.J., and Dolan, R.J. (2012). Action-specific value signals in reward-related regions of the human brain. *J. Neurosci.* 32, 16417–16423a.

Fleming, S.M., van der Putten, E.J., and Daw, N.D. (2018). Neural mediators of changes of mind about perceptual decisions. *Nature Neuroscience* 21, 617–624.

Folke, T., Jacobsen, C., Fleming, S.M., and De Martino, B. (2016). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour* 1, 0002.

Fouragnan, E., Retzler, C., Mullinger, K., and Philiastides, M.G. (2015). Two spatiotemporally distinct value systems shape reward-based learning in the human brain. *Nat Commun* 6, 8107.

## Bibliographie

- Fouragnan, E., Retzler, C., and Philiastides, M.G. (2018). Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Hum. Brain. Mapp.* 39, 2887–2906.
- Fox, C.R., and Poldrack, R.A. (2014). Prospect Theory and the Brain. In *Neuroeconomics: Decision-Making and the Brain*, p.
- Franklin, B. (1748). *Advice to a Young Tradesman*. 2.
- Frederick, S., Loewenstein, G., and O'donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* 40, 351–401.
- Freud, S. (1988). *Sur le rêve* (Gallimard).
- Friedman, D., Hakerem, G., Sutton, S., and Fleiss, J.L. (1973). Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalography and Clinical Neurophysiology* 34, 475–484.
- Frot, M., and Mauguière, F. (2003). Dual representation of pain in the operculo-insular cortex in humans. *Brain* 126, 438–450.
- Fullana, M.A., Albajes-Eizagirre, A., Soriano-Mas, C., Vervliet, B., Cardoner, N., Benet, O., Radua, J., and Harrison, B.J. (2018). Fear extinction in the human brain: A meta-analysis of fMRI studies in healthy participants. *Neuroscience & Biobehavioral Reviews* 88, 16–25.
- Gagliano, M., Vyazovskiy, V.V., Borbély, A.A., Grimmonprez, M., and Depczynski, M. (2016). Learning by Association in Plants. *Sci Rep* 6, 38427.
- Gailliot, M.T., and Baumeister, R.F. (2007). The physiology of willpower: linking blood glucose to self-control. *Pers Soc Psychol Rev* 11, 303–327.
- Gailliot, M.T., Baumeister, R.F., DeWall, C.N., Maner, J.K., Plant, E.A., Tice, D.M., Brewer, L.E., and Schmeichel, B.J. (2007). Self-control relies on glucose as a limited energy source: willpower is more than a metaphor. *J Pers Soc Psychol* 92, 325–336.
- Gal, D., and Rucker, D.D. (2018). The Loss of Loss Aversion: Will It Loom Larger Than Its Gain? *J Consum Psychol* 28, 497–516.
- Gallagher, K.M., Fadel, P.J., Strømstad, M., Ide, K., Smith, S.A., Querry, R.G., Raven, P.B., and Secher, N.H. (2001). Effects of partial neuromuscular blockade on carotid baroreflex function during exercise in humans. *The Journal of Physiology* 533, 861–870.
- Garcia-Larrea, L., Perchet, C., Creac'h, C., Convers, P., Peyron, R., Laurent, B., Mauguière, F., and Magnin, M. (2010). Operculo-insular pain (parasylyvian pain): a distinct central pain syndrome. *Brain* 133, 2528–2539.
- de Gardelle, V., and Mamassian, P. (2014). Does confidence use a common currency across two visual tasks? *Psychol Sci* 25, 1286–1288.
- de Gardelle, V., Le Corre, F., and Mamassian, P. (2016). Confidence as a Common Currency between Vision and Audition. *PLOS ONE* 11, e0147901.

## Bibliographie

- Garrison, J., Erdeniz, B., and Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews* 37, 1297–1310.
- Geng, J.J., Blumenfeld, Z., Tyson, T.L., and Minzenberg, M.J. (2015). Pupil diameter reflects uncertainty in attentional selection during visual search. *Front Hum Neurosci* 9, 435.
- Gengerelli, J.A. (1930). The principle of maxima and minima in animal learning. *Journal of Comparative Psychology* 11, 193–236.
- Georgiadis, J.R., and Kringelbach, M.L. (2012). The human sexual response cycle: Brain imaging evidence linking sex to other pleasures. *Progress in Neurobiology* 98, 49–81.
- Gergelyfi, M., Jacob, B., Olivier, E., and Zénon, A. (2015). Dissociation between mental fatigue and motivational state during prolonged mental activity. *Front. Behav. Neurosci.* 9.
- Gervain, J., and Geffen, M.N. (2019). Efficient Neural Coding in Auditory and Speech Perception. *Trends in Neurosciences* 42, 56–65.
- Gherman, S., and Philiastides, M.G. (2018). Human VMPFC encodes early signatures of confidence in perceptual decisions. *ELife* 7, e38293.
- Gläscher, J., Hampton, A.N., and O’Doherty, J.P. (2009). Determining a Role for Ventromedial Prefrontal Cortex in Encoding Action-Based Value Signals During Reward-Related Decision Making. *Cerebral Cortex* 19, 483–495.
- Glascher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., Paul, L.K., and Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences* 109, 14681–14686.
- Glimcher, P.W. (2014). Chapter 20 - Value-Based Decision Making. In *Neuroeconomics (Second Edition)*, P.W. Glimcher, and E. Fehr, eds. (San Diego: Academic Press), pp. 373–391.
- Glimcher, P.W., and Fehr, E. (2014). Introduction: A Brief History of Neuroeconomics. In *Neuroeconomics (Second Edition)*, P.W. Glimcher, and E. Fehr, eds. (San Diego: Academic Press), pp. xvii–xxviii.
- Gluth, S., Sommer, T., Rieskamp, J., and Büchel, C. (2015). Effective Connectivity between Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices from Memory. *Neuron* 86, 1078–1090.
- Gluth, S., Kern, N., Kortmann, M., and Vitali, C.L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nat Hum Behav* 1–12.
- Gold, J.I., and Shadlen, M.N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience* 30, 535–574.
- Green, D.I. (1894). Pain-Cost and Opportunity-Cost. *Q J Econ* 8, 218–229.
- Greicius, M.D., Supekar, K., Menon, V., and Dougherty, R.F. (2009). Resting-State Functional Connectivity Reflects Structural Connectivity in the Default Mode Network. *Cerebral Cortex* 19, 72–78.



## Bibliographie

Grillner, S., and Robertson, B. (2016). The Basal Ganglia Over 500 Million Years. *Current Biology* 26, R1088–R1100.

Grueschow, M., Polania, R., Hare, T.A., and Ruff, C.C. (2015). Automatic versus Choice-Dependent Value Representations in the Human Brain. *Neuron* 85, 874–885.

Haji, F.A., Rojas, D., Childs, R., de Ribaupierre, S., and Dubrowski, A. (2015). Measuring cognitive load: performance, mental effort and simulation task complexity. *Med Educ* 49, 815–827.

Halai, A.D., Welbourne, S.R., Embleton, K., and Parkes, L.M. (2014). A comparison of dual gradient-echo and spin-echo fMRI of the inferior temporal lobe: Comparison of Dual GE and Spin-Echo fMRI. *Human Brain Mapping* 35, 4118–4128.

Harber, V.J., and Sutton, J.R. (1984). Endorphins and Exercise. *Sports Medicine* 1, 154–171.

Hare, T.A., Camerer, C.F., and Rangel, A. (2009). Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. *324*, 4.

Hare, T.A., Malmaud, J., and Rangel, A. (2011a). Focusing Attention on the Health Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary Choice. *Journal of Neuroscience* 31, 11077–11087.

Hare, T.A., Schultz, W., Camerer, C.F., O’Doherty, J.P., and Rangel, A. (2011b). Transformation of stimulus value signals into motor commands during simple choice. *PNAS* 108, 18120–18125.

Hare, T.A., Hakimi, S., and Rangel, A. (2014). Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. *Front. Neurosci.* 8.

Hartmann, M.N., Hager, O.M., Tobler, P.N., and Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes* 100, 192–196.

Hausenblas, H.A., Schreiber, K., and Smoliga, J.M. (2017). Addiction to exercise. *BMJ* 357.

Hawkins, G.E., Forstmann, B.U., Wagenmakers, E.-J., Ratcliff, R., and Brown, S.D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *J. Neurosci.* 35, 2476–2484.

Hayes, D.J., and Northoff, G. (2011). Identifying a Network of Brain Regions Involved in Aversion-Related Processing: A Cross-Species Translational Investigation. *Frontiers in Integrative Neuroscience* 5.

Hayes, D.J., and Northoff, G. (2012). Common brain activations for painful and non-painful aversive stimuli. *BMC Neuroscience* 13, 60.

Hayes, S.M., Buchler, N., Stokes, J., Kragel, J., and Cabeza, R. (2011). Neural Correlates of Confidence during Item Recognition and Source Memory Retrieval: Evidence for Both Dual-process and Strength Memory Theories. *Journal of Cognitive Neuroscience* 23, 3959–3971.

## Bibliographie

- Heath, R.G. (1972). PLEASURE AND BRAIN ACTIVITY IN MAN: Deep and Surface Electroencephalograms During Orgasm. *The Journal of Nervous and Mental Disease* 154, 3–18.
- Hebart, M.N., Schriever, Y., Donner, T.H., and Haynes, J.-D. (2016). The Relationship between Perceptual Decision Variables and Confidence in the Human Brain. *Cerebral Cortex* 26, 118–130.
- Hebb, D.O. (1955). Drives and the C. N. S. (conceptual nervous system). *Psychological Review* 62, 243–254.
- Henri-Bhargava, A., Simioni, A., and Fellows, L.K. (2012). Ventromedial frontal lobe damage disrupts the accuracy, but not the speed, of value-based preference judgments. *Neuropsychologia* 50, 1536–1542.
- Herman, A.M., Critchley, H.D., and Duka, T. (2018). The role of emotions and physiological arousal in modulating impulsive behaviour. *Biol Psychol* 133, 30–43.
- Hertwig, R., and Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences* 13, 517–523.
- Hertwig, R., Wulff, D.U., and Mata, R. Three gaps and what they may mean for risk preference. 10.
- Hidalgo-Muñoz, A.R., Mouratille, D., Matton, N., Causse, M., Rouillard, Y., and El-Yagoubi, R. (2018). Cardiovascular correlates of emotional state, cognitive workload and time-on-task effect during a realistic flight simulation. *International Journal of Psychophysiology* 128, 62–69.
- Hockey, G.R.J. (2011). A motivational control theory of cognitive fatigue. In *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications.*, P.L. Ackerman, ed. (Washington: American Psychological Association), pp. 167–187.
- Hoehl, S., Hellmer, K., Johansson, M., and Gredebäck, G. (2017). Itsy Bitsy Spider...: Infants React with Increased Arousal to Spiders and Snakes. *Front. Psychol.* 8.
- Hogan, P.S., Galaro, J.K., and Chib, V.S. (2019). Roles of Ventromedial Prefrontal Cortex and Anterior Cingulate in Subjective Valuation of Prospective Effort. *Cerebral Cortex* 29, 4277–4290.
- Hogeveen, J., Hauner, K.K., Chau, A., Krueger, F., and Grafman, J. (2016). Impaired Valuation Leads to Increased Apathy Following Ventromedial Prefrontal Cortex Damage. *Cerebral Cortex* bhv317.
- Holroyd, C.B. (2015). The Waste Disposal Problem of Effortful Control.
- Houthakker, H.S. (1950). Revealed Preference and the Utility Function. *Economica* 17, 159–174.
- Hsu, M., Krajbich, I., Zhao, C., and Camerer, C.F. (2009). Neural Response to Reward Anticipation under Risk Is Nonlinear in Probabilities. *Journal of Neuroscience* 29, 2231–2237.

## Bibliographie

- Huettel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., and Platt, M.L. (2006). Neural Signatures of Economic Preferences for Risk and Ambiguity. *Neuron* 49, 765–775.
- Huk, A.C., and Shadlen, M.N. (2005). Neural Activity in Macaque Parietal Cortex Reflects Temporal Integration of Visual Motion Signals during Perceptual Decision Making. *J. Neurosci.* 25, 10420–10436.
- Hull, C.L. (1943). *Principles of Behavior: An Introduction to Behavior Theory* (D. Appleton-Century Company, Incorporated).
- Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F.S., and Behrens, T.E.J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience* 15, 470–476.
- Hutchison, W.D., Davis, K.D., Lozano, A.M., Tasker, R.R., and Dostrovsky, J.O. (1999). Pain-related neurons in the human cingulate cortex. *Nat. Neurosci.* 2, 403–405.
- Iannaccone, R., Hauser, T.U., Staempfli, P., Walitza, S., Brandeis, D., and Brem, S. (2015). Conflict monitoring and error processing: New insights from simultaneous EEG–fMRI. *NeuroImage* 105, 395–407.
- Ikeda, M., Brown, J., Holland, A.J., Fukuhara, R., and Hodges, J.R. (2002). Changes in appetite, food preference, and eating habits in frontotemporal dementia and Alzheimer’s disease. *Journal of Neurology, Neurosurgery & Psychiatry* 73, 371–376.
- Ikemoto, S. (2010). Brain reward circuitry beyond the mesolimbic dopamine system: A neurobiological theory. *Neuroscience & Biobehavioral Reviews* 35, 129–150.
- Ikemoto, S., Glazier, B.S., Murphy, J.M., and McBride, W.J. (1997). Role of Dopamine D1 and D2 Receptors in the Nucleus Accumbens in Mediating Reward. *J. Neurosci.* 17, 8580–8587.
- Inzlicht, M., Schmeichel, B.J., and Macrae, C.N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences* 18, 127–133.
- Inzlicht, M., Shenhav, A., and Olivola, C.Y. (2018). The Effort Paradox: Effort Is Both Costly and Valued. *Trends Cogn. Sci. (Regul. Ed.)* 22, 337–349.
- Isnard, J., Guénot, M., Sindou, M., and Mauguière, F. (2004). Clinical Manifestations of Insular Lobe Seizures: A Stereo-electroencephalographic Study. *Epilepsia* 45, 1079–1090.
- Izquierdo, I., Furini, C.R.G., and Myskiw, J.C. (2016). Fear Memory. *Physiological Reviews* 96, 695–750.
- Jin, F., Zheng, P., Liu, H., Guo, H., and Sun, Z. (2018). Functional and anatomical connectivity-based parcellation of human cingulate cortex. *Brain Behav* 8, e01070.
- Job, V., Walton, G.M., Bernecker, K., and Dweck, C.S. (2013). Beliefs about willpower determine the impact of glucose on self-control. *Proceedings of the National Academy of Sciences* 110, 14837–14842.

## Bibliographie

- Jocham, G., Hunt, L.T., Near, J., and Behrens, T.E.J. (2012). A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat Neurosci* *15*, 960–961.
- Jocham, G., Furlong, P.M., Kröger, I.L., Kahn, M.C., Hunt, L.T., and Behrens, T.E.J. (2014). Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. *Neuroimage* *100*, 498–506.
- Kable, J.W. (2014). Chapter 10 - Valuation, Intertemporal Choice, and Self-Control. In *Neuroeconomics (Second Edition)*, P.W. Glimcher, and E. Fehr, eds. (San Diego: Academic Press), pp. 173–192.
- Kable, J.W., and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* *10*, 1625–1633.
- Kagel, J.H., Green, L., and Caraco, T. (1986). When foragers discount the future: constraint or adaptation? *Animal Behaviour* *34*, 271–283.
- Kahneman, D. (2012). *Thinking, Fast and Slow* (Penguin Books).
- Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. *Science* *154*, 1583–1585.
- Kahneman, D., and Sugden, R. (2005). Experienced Utility as a Standard of Policy Evaluation. *Environ Resource Econ* *32*, 161–181.
- Kahneman, D., and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* *47*, 263.
- Kahneman, D., and Tversky, A. (1984). Choices, values, and frames. *American Psychologist* *39*, 341–350.
- Kalenscher, T., and Pennartz, C.M.A. (2008). Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Progress in Neurobiology* *84*, 284–315.
- Kamiński, J., Sullivan, S., Chung, J.M., Ross, I.B., Mamelak, A.N., and Rutishauser, U. (2017). Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nature Neuroscience* *20*, 590–601.
- Kano, M., Oudenhove, L.V., Dupont, P., Wager, T.D., and Fukudo, S. (2020). Imaging Brain Mechanisms of Functional Somatic Syndromes: Potential as a Biomarker? *Tohoku J. Exp. Med.* *250*, 137–152.
- Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., and Carter, C.S. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science* *303*, 1023–1026.
- Khamassi, M., Enel, P., Dominey, P.F., and Procyk, E. (2013). Medial prefrontal cortex and the adaptive regulation of reinforcement learning parameters. In *Progress in Brain Research*, (Elsevier), pp. 441–464.

## Bibliographie

- Khaw, M.W., Glimcher, P.W., and Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proc. Natl. Acad. Sci. U.S.A.* *114*, 12696–12701.
- Kiani, R., Corthell, L., and Shadlen, M.N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron* *84*, 1329–1342.
- Kim, H., Shimojo, S., and O’Doherty, J.P. (2006). Is Avoiding an Aversive Outcome Rewarding? Neural Substrates of Avoidance Learning in the Human Brain. *PLOS Biology* *4*, e233.
- Kim, H., Adolphs, R., O’Doherty, J.P., and Shimojo, S. (2007). Temporal isolation of neural processes underlying face preference decisions. *Proceedings of the National Academy of Sciences* *104*, 18253–18258.
- Klein-Flügge, M.C., Kennerley, S.W., Saraiva, A.C., Penny, W.D., and Bestmann, S. (2015). Behavioral Modeling of Human Choices Reveals Dissociable Effects of Physical Effort and Temporal Delay on Reward Devaluation. *PLoS Comput Biol* *11*, e1004116.
- Klein-Flügge, M.C., Kennerley, S.W., Friston, K., and Bestmann, S. (2016). Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. *Journal of Neuroscience* *36*, 10002–10015.
- Klingner, J., Kumar, R., and Hanrahan, P. (2008). Measuring the task-evoked pupillary response with a remote eye tracker. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications - ETRA '08*, (Savannah, Georgia: ACM Press), p. 69.
- Knight, F.H. (Frank H. (1921). *Risk, uncertainty and profit* (Boston, New York, Houghton Mifflin Company).
- Knutson, B., Adams, C.M., Fong, G.W., and Hommer, D. (2001a). Anticipation of Increasing Monetary Reward Selectively Recruits Nucleus Accumbens. *J. Neurosci.* *21*, RC159–RC159.
- Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., and Hommer, D. (2001b). Dissociation of reward anticipation and outcome with event-related fMRI: *Neuroreport* *12*, 3683–3687.
- Knutson, B., Rick, S., Wimmer, G.E., Prelec, D., and Loewenstein, G. (2007). Neural Predictors of Purchases. *Neuron* *53*, 147–156.
- Knutson, B., Katovich, K., and Suri, G. (2014). Inferring affect from fMRI data. *Trends in Cognitive Sciences* *18*, 422–428.
- Kobayashi, S., Pinto de Carvalho, O., and Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. *J. Neurosci.* *30*, 534–544.
- Koechlin, E., Ody, C., and Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science* *302*, 1181–1185.
- Kolling, N., and O’Reilly, J.X. (2018). State-change decisions and dorsomedial prefrontal cortex: the importance of time. *Current Opinion in Behavioral Sciences* *22*, 152–160.

## Bibliographie

Kolling, N., Behrens, T.E.J., Mars, R.B., and Rushworth, M.F.S. (2012). Neural Mechanisms of Foraging. *Science* 336, 95–98.

Kolling, N., Wittmann, M.K., Behrens, T.E.J., Boorman, E.D., Mars, R.B., and Rushworth, M.F.S. (2016a). Value, search, persistence and model updating in anterior cingulate cortex. *Nature Neuroscience* 19, 1280–1285.

Kolling, N., Behrens, T., Wittmann, M., and Rushworth, M. (2016b). Multiple signals in anterior cingulate cortex. *Current Opinion in Neurobiology* 37, 36–43.

Kool, W., and Botvinick, M. (2018). Mental labour. *Nat Hum Behav* 2, 899–908.

Kool, W., McGuire, J.T., Rosen, Z.B., and Botvinick, M.M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General* 139, 665–682.

Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *PNAS* 108, 13852–13857.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298.

Kringelbach, M. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology* 72, 341–372.

Kringelbach, M.L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nat Rev Neurosci* 6, 691–702.

Krishnan, A., Woo, C.-W., Chang, L.J., Ruzic, L., Gu, X., López-Solà, M., Jackson, P.L., Pujol, J., Fan, J., and Wager, T.D. (2016). Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *ELife* 5.

Kuchinke, L., Fritzsche, S., Hofmann, M.J., and Jacobs, A.M. (2013). Neural correlates of episodic memory: Associative memory and confidence drive hippocampus activations. *Behavioural Brain Research* 254, 92–101.

Kurniawan, I.T. (2011). Dopamine and effort-based decision making. *Frontiers in Neuroscience* 5.

Kurniawan, I.T., Guitart-Masip, M., Dayan, P., and Dolan, R.J. (2013). Effort and Valuation in the Brain: The Effects of Anticipation and Execution. *Journal of Neuroscience* 33, 6160–6169.

Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? *Evol Psychol* 8, 244–259.

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology* 7, 67–70.

Kurzban, R., Duckworth, A., Kable, J.W., and Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences* 36, 661–679.

Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E., Weisskoff, R.M., Poncelet, B.P., Kennedy, D.N., Hoppel, B.E., Cohen, M.S., and Turner, R. (1992). Dynamic magnetic

## Bibliographie

resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences* 89, 5675–5679.

Laërce, D. (1999). Livre X - Epicure. In *Vies et Doctrines Des Philosophes Illustres*, (Le Livre de Poche), p.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics* 112, 443–478.

Lange, F., and Eggert, F. (2014). Sweet delusion. Glucose drinks fail to counteract ego depletion. *Appetite* 75, 54–63.

Laplane, D., and Dubois, B. (2001). Auto-Activation deficit: A basal ganglia related syndrome: Auto-Activation Deficit. *Mov. Disord.* 16, 810–814.

Laplane, D., Baulac, M., Widlocher, D., and Dubois, B. (1984). Pure psychic akinesia with bilateral lesions of basal ganglia. *Journal of Neurology, Neurosurgery & Psychiatry* 47, 377–385.

Larsen, T., and O’Doherty, J.P. (2014). Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI–EEG study. *Phil. Trans. R. Soc. B* 369, 20130473.

Latimer, K.W., Yates, J.L., Meister, M.L.R., Huk, A.C., and Pillow, J.W. (2015). NEURONAL MODELING. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science* 349, 184–187.

Le Bon, G. (2006). *Psychologie des foules* (PUF presses universitaires de France).

Le Bouc, R., and Pessiglione, M. (2018). La motivation dans tous ses K: Approche computationnelle des pathologies neuropsychiatriques. *Med Sci (Paris)* 34, 238–246.

Le Bouc, R., Rigoux, L., Schmidt, L., Degos, B., Welter, M.-L., Vidailhet, M., Daunizeau, J., and Pessiglione, M. (2016). Computational Dissection of Dopamine Motor and Motivational Functions in Humans. *Journal of Neuroscience* 36, 6623–6633.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. *Neuron* 64, 431–439.

Lebreton, M., Bertoux, M., Boutet, C., Lehericy, S., Dubois, B., Fossati, P., and Pessiglione, M. (2013). A Critical Role for the Hippocampus in the Valuation of Imagined Outcomes. *PLoS Biology* 11, e1001684.

Lebreton, M., Abitbol, R., Daunizeau, J., and Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience* 18, 1159–1167.

Lebreton, M., Langdon, S., Slieker, M.J., Nooitgedacht, J.S., Goudriaan, A.E., Denys, D., van Holst, R.J., and Luigjes, J. (2018). Two sides of the same coin: Monetary incentives concurrently improve and bias confidence judgments. *SCIENCE ADVANCES* 14.

## Bibliographie

- LeDoux, J.E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience* 23, 155–184.
- LeDoux, J.E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences* 111, 2871–2878.
- Lee, D., and Daunizeau, J. (2019). Trading Mental Effort for Confidence: The Metacognitive Control of Value-Based Decision-Making. *BioRxiv* 837054.
- Leech, R., and Sharp, D.J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain* 137, 12–32.
- Lejarraga, T., Woike, J.K., and Hertwig, R. (2016). Description and experience: How experimental investors learn about booms and busts affects their financial risk taking. *Cognition* 157, 365–383.
- Lempert, K.M., Johnson, E., and Phelps, E.A. (2016). Emotional arousal predicts intertemporal choice. *Emotion* 16, 647–656.
- Lévêque, M. (2016). La psychochirurgie dans le traitement de la douleur. *Douleurs : Evaluation - Diagnostic - Traitement* 17, 192–199.
- Levi, V., Cordella, R., D’Ammando, A., Tringali, G., Dones, I., Messina, G., and Franzini, A. (2019). Dorsal anterior cingulate cortex (ACC) deep brain stimulation (DBS): a promising surgical option for the treatment of refractory thalamic pain syndrome (TPS). *Acta Neurochir* 161, 1579–1588.
- Levy, D.J., and Glimcher, P.W. (2011). Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. *Journal of Neuroscience* 31, 14693–14707.
- Levy, D.J., and Glimcher, P.W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology* 22, 1027–1038.
- Levy, I., Snell, J., Nelson, A.J., Rustichini, A., and Glimcher, P.W. (2010). Neural Representation of Subjective Value Under Risk and Ambiguity. *Journal of Neurophysiology* 103, 1036–1047.
- Levy, I., Lazzaro, S.C., Rutledge, R.B., and Glimcher, P.W. (2011). Choice from Non-Choice: Predicting Consumer Preferences from Blood Oxygenation Level-Dependent Signals Obtained during Passive Viewing. *Journal of Neuroscience* 31, 118–125.
- Li, Y., Sescousse, G., Amiez, C., and Dreher, J.-C. (2015). Local Morphology Predicts Functional Organization of Experienced Value Signals in the Human Orbitofrontal Cortex. *Journal of Neuroscience* 35, 1648–1658.
- Lim, S.-L., O’Doherty, J.P., and Rangel, A. (2011). The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. *Journal of Neuroscience* 31, 13214–13223.
- Lin, A., Adolphs, R., and Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience* 7, 274–281.



## Bibliographie

- van der Linden, D., Frese, M., and Meijman, T.F. (2003). Mental fatigue and the control of cognitive processes: effects on perseveration and planning. *Acta Psychologica 113*, 45–65.
- Litt, A., Plassmann, H., Shiv, B., and Rangel, A. (2011). Dissociating Valuation and Saliency Signals during Decision-Making. *Cerebral Cortex 21*, 95–102.
- Liu, X., Hairston, J., Schrier, M., and Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews 35*, 1219–1236.
- Locke, H.S., and Braver, T.S. (2008). Motivational influences on cognitive control: behavior, brain activation, and individual differences. *Cognitive, Affective, & Behavioral Neuroscience 8*, 99–112.
- Logothetis, N.K., and Wandell, B.A. (2004). Interpreting the BOLD Signal. *Annu. Rev. Physiol. 66*, 735–769.
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *412*, 8.
- Lopez-Persem, A. (2016). *The Brain Valuation System and its role in decision-making*. Sorbonne Université.
- Lopez-Persem, A., Domenech, P., and Pessiglione, M. (2016). How prior preferences determine decision-making frames and biases in the human brain. *ELife 5*.
- Lopez-Persem, A., Verhagen, L., Amiez, C., Petrides, M., and Sallet, J. (2018). The human ventromedial prefrontal cortex: sulcal morphology and its influence on functional organization (Neuroscience).
- Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., Navarro, V., Rheims, S., Kahane, P., Domenech, P., et al. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nat Neurosci 1–12*.
- Lorist, M.M., Klein, M., Nieuwenhuis, S., Jong, R. de, Mulder, G., and Meijman, T.F. (2000). Mental fatigue and task control: Planning and preparation. *Psychophysiology 37*, 614–625.
- Louie, K., and Glimcher, P.W. (2012). Efficient coding and the neural representation of value: Louie & Glimcher. *Annals of the New York Academy of Sciences 1251*, 13–32.
- Louie, K., Khaw, M.W., and Glimcher, P.W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *PNAS 110*, 6139–6144.
- Ma, L.-B., and Wu, S. (2011). Efficient coding of natural images. *Sheng Li Xue Bao 63*, 463–471.
- MacDonald, A.W., Cohen, J.D., Stenger, V.A., and Carter, C.S. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science 288*, 1835–1838.
- Mackey, S., and Petrides, M. (2009). Architectonic mapping of the medial region of the human orbitofrontal cortex by density profiles. *Neuroscience 159*, 1089–1107.

## Bibliographie

- Mackey, S., and Petrides, M. (2010). Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains: Comparable cortical areas in human and macaque. *European Journal of Neuroscience* 32, 1940–1950.
- Mackey, S., and Petrides, M. (2014). Architecture and morphology of the human ventromedial prefrontal cortex. *European Journal of Neuroscience* 40, 2777–2796.
- Mannella, F., Gurney, K., and Baldassarre, G. (2013). The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. *Front. Behav. Neurosci.* 7.
- Marc-Aurèle (2007). Livre III. In *Pensées Pour Moi-Même*, (GF-Flammarion), p.
- Marcora, S.M., and Staiano, W. (2010). The limit to exercise tolerance in humans: mind over muscle? *Eur. J. Appl. Physiol.* 109, 763–770.
- Marcora, S.M., Staiano, W., and Manning, V. (2009). Mental fatigue impairs physical performance in humans. *J Appl Physiol* 106, 8.
- Margulies, D.S., Kelly, A.M.C., Uddin, L.Q., Biswal, B.B., Castellanos, F.X., and Milham, M.P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *NeuroImage* 37, 579–588.
- Mars, R.B., Verhagen, L., Gladwin, T.E., Neubert, F.-X., Sallet, J., and Rushworth, M.F.S. (2016). Comparing brains by matching connectivity profiles. *Neuroscience & Biobehavioral Reviews* 60, 90–97.
- Martinet, A. (1955). *Économie des changements phonétiques: traité de phonologie diachronique* (Berne: Francke).
- Massar, S.A.A., Libedinsky, C., Weiyan, C., Huettel, S.A., and Chee, M.W.L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *Neuroimage* 120, 104–113.
- Matthews, G., and Davies, D.R. (2001). Individual differences in energetic arousal and sustained attention: a dual-task study. *Personality and Individual Differences* 31, 575–589.
- Mazzola, L., Isnard, J., Peyron, R., and Mauguère, F. (2012). Stimulation of the human cortex and the experience of pain: Wilder Penfield's observations revisited. *Brain* 135, 631–640.
- Mazzola, L., Mauguère, F., and Isnard, J. (2017). Electrical Stimulations of the Human Insula: Their Contribution to the Ictal Semiology of Insular Seizures. *Journal of Clinical Neurophysiology* 34, 307–314.
- Mazzola, L., Mauguère, F., and Isnard, J. (2019). Functional mapping of the human insula: Data from electrical stimulations. *Rev. Neurol. (Paris)* 175, 150–156.
- McCloy, D.R., Lau, B.K., Larson, E., Pratt, K.A.I., and Lee, A.K.C. (2017). Pupillometry shows the effort of auditory attention switching. *The Journal of the Acoustical Society of America* 141, 2440–2451.

## Bibliographie

- McClure, S.M. (2004). Separate Neural Systems Value Immediate and Delayed Monetary Rewards. *Science* 306, 503–507.
- McGinty, V.B., Rangel, A., and Newsome, W.T. (2016). Orbitofrontal Cortex Value Signals Depend on Fixation Location during Free Viewing. *Neuron* 90, 1299–1311.
- McGuire, J.T., and Botvinick, M.M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *PNAS* 107, 7922–7926.
- McKlveen, J.M., Myers, B., and Herman, J.P. (2015). The Medial Prefrontal Cortex: Coordinator of Autonomic, Neuroendocrine and Behavioural Responses to Stress. *Journal of Neuroendocrinology* 27, 446–456.
- Menon, V., and Uddin, L.Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function* 214, 655–667.
- Meyniel, F., Sergent, C., Rigoux, L., Daunizeau, J., and Pessiglione, M. (2013). Neurocomputational account of how the human brain decides when to have a break. *Proceedings of the National Academy of Sciences* 110, 2641–2646.
- Meyniel, F., Sigman, M., and Mainen, Z.F. (2015). Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron* 88, 78–92.
- Miller, H.V., Barnes, J.C., and Beaver, K.M. (2011). Self-control and Health Outcomes in a Nationally Representative Sample.
- Miller, K.J., Shenhav, A., and Ludvig, E.A. (2019). Habits without values. *Psychological Review* 126, 292–311.
- Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., and Rangel, A. (2010). The Drift Diffusion Model Can Account for the Accuracy and Reaction Time of Value-Based Choices Under High and Low Time Pressure. *SSRN Journal*.
- Ming, V.L., and Holt, L.L. (2009). Efficient coding in human auditory perception. *J. Acoust. Soc. Am.* 126, 1312–1320.
- Mischel, W. (1966). Theory and research on the antecedents of self-imposed delay of reward. *Prog Exp Pers Res* 3, 85–132.
- Mischel, W., and Ebbesen, E.B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology* 16, 329–337.
- Molden, D.C., Hui, C.M., Scholer, A.A., Meier, B.P., Noreen, E.E., D’Agostino, P.R., and Martin, V. (2012). Motivational versus metabolic effects of carbohydrates on self-control. *Psychol Sci* 23, 1137–1144.
- Molenberghs, P., Trautwein, F.-M., Böckler, A., Singer, T., and Kanske, P. (2016). Neural correlates of metacognitive ability and of feeling confident: a large-scale fMRI study. *Social Cognitive and Affective Neuroscience* nsw093.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences* 7, 134–140.

## Bibliographie

- Morgan, W.P. (1994). Psychological components of effort sense. *Med Sci Sports Exerc* 26, 1071–1077.
- Moritz, S., Gläscher, J., Sommer, T., Büchel, C., and Braus, D.F. (2006). Neural correlates of memory confidence. *NeuroImage* 33, 1188–1193.
- Morris, L.S., Kundu, P., Dowell, N., Mechelmans, D.J., Favre, P., Irvine, M.A., Robbins, T.W., Daw, N., Bullmore, E.T., Harrison, N.A., et al. (2016). Fronto-striatal organization: Defining functional and microstructural substrates of behavioural flexibility. *Cortex* 74, 118–133.
- Motoki, K., Sugiura, M., and Kawashima, R. (2019). Common neural value representations of hedonic and utilitarian products in the ventral striatum: An fMRI study. *Sci Rep* 9, 15630.
- Mouraux, A., Diukova, A., Lee, M.C., Wise, R.G., and Iannetti, G.D. (2011). A multisensory investigation of the functional significance of the “pain matrix.” *NeuroImage* 54, 2237–2249.
- Mukherjee, S., Sahay, A., Pammi, V.S.C., and Srinivasan, N. (2017). Is loss-aversion magnitude-dependent? Measuring prospective affective judgments regarding gains and losses. *Judgment and Decision Making* 12, 81–89.
- Müller, T., and Apps, M.A.J. (2019). Motivational fatigue: A neurocognitive framework for the impact of effortful exertion on subsequent motivation. *Neuropsychologia* 123, 141–151.
- Muraven, M., and Slessareva, E. (2003). Mechanisms of self-control failure: motivation and limited resources. *Pers Soc Psychol Bull* 29, 894–906.
- Murphy, P.R., O’Connell, R.G., O’Sullivan, M., Robertson, I.H., and Balsters, J.H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Hum. Brain Mapp.* 35, 4140–4154.
- Mushtaq, F., Bland, A.R., and Schaefer, A. (2011). Uncertainty and Cognitive Control. *Frontiers in Psychology* 2.
- Myerson, J., Green, L., Scott Hanson, J., Holt, D.D., and Estle, S.J. (2003). Discounting delayed and probabilistic rewards: Processes and traits. *Journal of Economic Psychology* 24, 619–635.
- Naccache, L., Dehaene, S., Cohen, L., Habert, M.-O., Guichart-Gomez, E., Galanaud, D., and Willer, J.-C. (2005). Effortless control: executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia* 43, 1318–1328.
- Nagase, A.M., Onoda, K., Foo, J.C., Haji, T., Akaishi, R., Yamaguchi, S., Sakai, K., and Morita, K. (2018). Neural Mechanisms for Adaptive Learned Avoidance of Mental Effort. *The Journal of Neuroscience* 38, 2631–2651.
- Namkung, H., Kim, S.-H., and Sawa, A. (2017). The Insula: An Underestimated Brain Area in Clinical Neuroscience, Psychiatry, and Neurology. *Trends in Neurosciences* 40, 200–207.
- Neafsey, E.J., Hull, C.D., and Buchwald, N.A. (1978). Preparation for movement in the cat. II. Unit activity in the basal ganglia and thalamus. *Electroencephalogr Clin Neurophysiol* 44, 714–723.

## Bibliographie

- Niendam, T.A., Laird, A.R., Ray, K.L., Dean, Y.M., Glahn, D.C., and Carter, C.S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn Affect Behav Neurosci* *12*, 241–268.
- Nishiyama, R. (2016). Physical, emotional, and cognitive effort discounting in gain and loss situations. *Behavioural Processes* *125*, 72–75.
- O’Doherty, J.P. (2011). Contributions of the ventromedial prefrontal cortex to goal-directed action selection: vmPFC in goal-directed action. *Annals of the New York Academy of Sciences* *1239*, 118–129.
- Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H., and Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences* *89*, 5951–5955.
- Ohmura, Y., Takahashi, T., Kitamura, N., and Wehr, P. (2006). Three-month stability of delay and probability discounting measures. *Exp Clin Psychopharmacol* *14*, 318–328.
- Olds, J., and 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*, 419–427.
- Olson, E.A., Hooper, C.J., Collins, P., and Luciana, M. (2007). Adolescents’ performance on delay and probability discounting tasks: contributions of age, intelligence, executive functioning, and self-reported externalizing behavior. *Pers Individ Dif* *43*, 1886–1897.
- O’Reilly, J.X., Schuffelgen, U., Cuell, S.F., Behrens, T.E.J., Mars, R.B., and Rushworth, M.F.S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences* *110*, E3660–E3669.
- Oudiette, D., Vinckier, F., Bioud, E., and Pessiglione, M. (2019). A Pavlovian account for paradoxical effects of motivation on controlling response vigour. *Scientific Reports* *9*.
- Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *J. Neurosci.* *29*, 14004–14014.
- Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* *34*, 333–359.
- Padoa-Schioppa, C., and Conen, K.E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron* *96*, 736–754.
- Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., and Pessiglione, M. (2009). Brain Hemispheres Selectively Track the Expected Value of Contralateral Options. *J. Neurosci.* *29*, 13465–13472.
- Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., Czernecki, V., Karachi, C., Capelle, L., Durr, A., et al. (2012). Critical Roles for Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning. *Neuron* *76*, 998–1009.

## Bibliographie

- Pals, C., Sarampalis, A., van Rijn, H., and Başkent, D. (2015). Validation of a simple response-time measure of listening effort. *The Journal of the Acoustical Society of America* *138*, EL187–EL192.
- Papageorgiou, G.K., Sallet, J., Wittmann, M.K., Chau, B.K.H., Schüffelgen, U., Buckley, M.J., and Rushworth, M.F.S. (2017). Inverted activity patterns in ventromedial prefrontal cortex during value-guided decision-making in a less-is-more task. *Nature Communications* *8*.
- Pareto, V. (1909). *Manuel d'économie politique*.
- Park, S.Q., Kahnt, T., Rieskamp, J., and Heekeren, H.R. (2011). Neurobiology of value integration: when value impacts valuation. *J. Neurosci.* *31*, 9307–9314.
- Parvizi, J., Rangarajan, V., Shirer, W.R., Desai, N., and Greicius, M.D. (2013). The Will to Persevere Induced by Electrical Stimulation of the Human Cingulate Gyrus. *Neuron* *80*, 1359–1367.
- Pauli, W.M., Nili, A.N., and Tyszka, J.M. (2017). A High-Resolution Probabilistic In Vivo Atlas of Human Subcortical Brain Nuclei. *BioRxiv* 211201.
- Paulus, M.P., and Stein, M.B. (2006). An Insular View of Anxiety. *Biological Psychiatry* *60*, 383–387.
- Pearce, J.M., and Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*.
- Pennartz, C.M.A., Lee, E., Verheul, J., Lipa, P., Barnes, C.A., and McNaughton, B.L. (2004). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation by Hippocampal Ripples. *J. Neurosci.* *24*, 6446–6456.
- Pessiglione, M., and Delgado, M.R. (2015). The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. *Curr Opin Behav Sci* *5*, 78–84.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., and Frith, C.D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* *442*, 1042–1045.
- Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R.J., and Frith, C.D. (2007). How the Brain Translates Money into Force: A Neuroimaging Study of Subliminal Motivation. *Science* *316*, 904–906.
- Pessiglione, M., Vinckier, F., Bouret, S., Daunizeau, J., and Le Bouc, R. (2018). Why not try harder? Computational approach to motivation deficits in neuro-psychiatric diseases. *Brain* *141*, 629–650.
- Peters, J., and Büchel, C. (2011). The neural mechanisms of inter-temporal decision-making: understanding variability. *Trends in Cognitive Sciences* *15*, 227–239.
- Peterson, W., Birdsall, T., and Fox, W. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory* *4*, 171–212.

## Bibliographie

- Peysakhovich, V., Causse, M., Scannella, S., and Dehais, F. (2015). Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort. *International Journal of Psychophysiology* 97, 30–37.
- Peysakhovich, V., Vachon, F., and Dehais, F. (2017). The impact of luminance on tonic and phasic pupillary responses to sustained cognitive load. *International Journal of Psychophysiology* 112, 40–45.
- Philiastides, M.G., Biele, G., and Heekeren, H.R. (2010). A mechanistic account of value computation in the human brain. *Proceedings of the National Academy of Sciences* 107, 9430–9435.
- Phillips, A.G., and Fibiger, H.C. (1978). The role of dopamine in maintaining intracranial self-stimulation in the ventral tegmentum, nucleus accumbens, and medial prefrontal cortex. *Canadian Journal of Psychology/Revue Canadienne de Psychologie* 32, 58–66.
- Phillips, P.E.M., Walton, M.E., and Jhou, T.C. (2007). Calculating utility: preclinical evidence for cost–benefit analysis by mesolimbic dopamine. *Psychopharmacology* 191, 483–495.
- Pisauro, M.A., Fouragnan, E., Retzler, C., and Philiastides, M.G. (2017). Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nature Communications* 8, 15808.
- Platon (2018). *Œuvres complètes*. Tome IX, 2e partie: Philèbe - Platon.
- Polanía, R., Woodford, M., and Ruff, C.C. (2019). Efficient coding of subjective value. *Nat Neurosci* 22, 134–142.
- Popper (2002). *Conjectures and Refutations: The Growth of Scientific Knowledge* (Psychology Press).
- Poser, B.A., Versluis, M.J., Hoogduin, J.M., and Norris, D.G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine* 55, 1227–1235.
- Poudel, R., Riedel, M.C., Salo, T., Flannery, J.S., Hill-Bowen, L.D., Eickhoff, S.B., Laird, A.R., and Sutherland, M.T. (2020). Common and distinct brain activity associated with risky and ambiguous decision-making (Neuroscience).
- Pouget, A., Drugowitsch, J., and Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nat. Neurosci.* 19, 366–374.
- Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M.-L., and Dreher, J.-C. (2010). Separate Valuation Subsystems for Delay and Effort Decision Costs. *Journal of Neuroscience* 30, 14080–14090.
- Protopapa, F., Hayashi, M.J., Kulashekhar, S., van der Zwaag, W., Battistella, G., Murray, M.M., Kanai, R., and Buetti, D. (2019). Chronotopic maps in human supplementary motor area. *PLoS Biol* 17, e3000026.

## Bibliographie

- Qin, S., Marle, H.J.F. van, Hermans, E.J., and Fernández, G. (2011). Subjective Sense of Memory Strength and the Objective Amount of Information Accurately Remembered Are Related to Distinct Neural Correlates at Encoding. *J. Neurosci.* *31*, 8920–8927.
- Rachlin, H., Raineri, A., and Cross, D. (1991). Subjective Probability and Delay. *Journal of the Experimental Analysis of Behavior* *55*, 233–244.
- Radua, J., Pozo, N.O. del, Gómez, J., Guillen-Grima, F., and Ortuño, F. (2014). Meta-analysis of functional neuroimaging studies indicates that an increase of cognitive difficulty during executive tasks engages brain regions associated with time perception. *Neuropsychologia* *58*, 14–22.
- Rahnev, D., Desender, K., Lee, A.L.F., Adler, W.T., Aguilar-Lleyda, D., Akdoğan, B., Arbuzova, P., Atlas, L.Y., Balci, F., Bang, J.W., et al. (2020). The Confidence Database. *Nat Hum Behav* 1–8.
- Raichle, M.E. (2015). The Brain's Default Mode Network. *Annu. Rev. Neurosci.* *38*, 433–447.
- Rangel, A., and Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology* *20*, 262–270.
- Rangel, A., Camerer, C., and Montague, P.R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience* *9*, 545–556.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev* *85*, 59–108.
- Ratcliff, R., Smith, P.L., Brown, S.D., and McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences* *20*, 260–281.
- Ravizza, S.M., and Carter, C.S. (2008). Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia* *46*, 2924–2935.
- Rescorla, R., and Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. p.
- Reynolds, B., Richards, J.B., Horn, K., and Karraker, K. (2004). Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behavioural Processes* *65*, 35–42.
- Ricardo, D. (2002). *Des principes de l'économie politique et de l'impôt* (Chicoutimi: J.-M. Tremblay).
- Richter, M., Friedrich, A., and Gendolla, G.H.E. (2008). Task difficulty effects on cardiac activity. *Psychophysiology* *45*, 869–875.
- Richter, M., Gendolla, G.H.E., and Wright, R.A. (2016). Chapter Five - Three Decades of Research on Motivational Intensity Theory: What We Have Learned About Effort and What We Still Don't Know. In *Advances in Motivation Science*, A.J. Elliot, ed. (Elsevier), pp. 149–186.
- Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., and Nieuwenhuis, S. (2004). The Role of the Medial Frontal Cortex in Cognitive Control. *Science* *306*, 443–447.



## Bibliographie

- Riggs, N.R., Spruijt-Metz, D., Sakuma, K.-L., Chou, C.-P., and Pentz, M.A. (2010). Executive cognitive function and food intake in children. *J Nutr Educ Behav* 42, 398–403.
- Riggs, N.R., Spruijt-Metz, D., Chou, C.-P., and Pentz, M.A. (2012). Relationships between executive cognitive function and lifetime substance use and obesity-related behaviors in fourth grade youth. *Child Neuropsychol* 18, 1–11.
- Roberts, K.L., and Hall, D.A. (2008). Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory stroop tasks. *J Cogn Neurosci* 20, 1063–1078.
- Robinson, M.D., Johnson, J.T., and Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology* 82, 416–425.
- Rolls, E.T. (2018). *The Brain, Emotion, and Depression* (Oxford, New York: Oxford University Press).
- Rolls, E.T., Grabenhorst, F., and Deco, G. (2010). Choice, difficulty, and confidence in the brain. *NeuroImage* 53, 694–706.
- Rotter, J.B. (1954). *Social learning and clinical psychology* (Englewood Cliffs, NJ, US: Prentice-Hall, Inc).
- Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A.R., Schulz, J.B., Fox, P.T., and Eickhoff, S.B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage* 60, 830–846.
- Russo, J.F., and Sheth, S.A. (2015). Deep brain stimulation of the dorsal anterior cingulate cortex for the treatment of chronic neuropathic pain. *FOC* 38, E11.
- Russo, S.J., and Nestler, E.J. (2013). The brain reward circuitry in mood disorders. *Nat Rev Neurosci* 14, 609–625.
- Rutledge, R.B., Skandali, N., Dayan, P., and Dolan, R.J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences* 111, 12252–12257.
- Sade, D. (marquis de) (2012). *La philosophie dans le boudoir: Dialogues destinés à l'éducation des jeunes Demoiselles. La mère en prescrira la lecture à sa fille.* (publie.net).
- Samanez-Larkin, G.R., Hollon, N.G., Carstensen, L.L., and Knutson, B. (2008). Individual Differences in Insular Sensitivity During Loss Anticipation Predict Avoidance Learning. *Psychol Sci* 19, 320–323.
- Samuelson, P.A. (1937). A Note on Measurement of Utility. *Rev Econ Stud* 4, 155–161.
- Samuelson, P.A. (1938). A Note on the Pure Theory of Consumer's Behaviour. *Economica* 5, 61–71.
- Savage, L.J. (1954). The foundations of statistics. *Naval Research Logistics Quarterly* 1, 236–236.

## Bibliographie

Schlaepfer, T.E., Cohen, M.X., Frick, C., Kosel, M., Brodessa, D., Axmacher, N., Joe, A.Y., Kreft, M., Lenartz, D., and Sturm, V. (2008). Deep Brain Stimulation to Reward Circuitry Alleviates Anhedonia in Refractory Major Depression. *Neuropsychopharmacol* 33, 368–377.

Schmidt, B., Mussel, P., and Hewig, J. (2013). I'm too calm--let's take a risk! On the impact of state and trait arousal on risk taking. *Psychophysiology* 50, 498–503.

Schmidt, L., d'Arc, B.F., Lafargue, G., Galanaud, D., Czernecki, V., Grabli, D., Schüpbach, M., Hartmann, A., Lévy, R., Dubois, B., et al. (2008). Disconnecting force from money: effects of basal ganglia damage on incentive motivation. *Brain* 131, 1303–1310.

Schmidt, L., Cléry-Melin, M.-L., Lafargue, G., Valabregue, R., Fossati, P., Dubois, B., and Pessiglione, M. (2009). Get Aroused and Be Stronger: Emotional Facilitation of Physical Effort in the Human Brain. *Journal of Neuroscience* 29, 9450–9457.

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., and Pessiglione, M. (2012). Neural Mechanisms Underlying Motivation of Mental Versus Physical Effort. *PLoS Biology* 10, e1001266.

Schneider, B., and Koenigs, M. (2017). Human lesion studies of ventromedial prefrontal cortex. *Neuropsychologia* 107, 84–93.

Schneider, M., Hathway, P., Leuchs, L., Sämann, P.G., Czisch, M., and Spoormaker, V.I. (2016). Spontaneous pupil dilations during the resting state are associated with activation of the salience network. *NeuroImage* 139, 189–201.

Schneider, M., Leuchs, L., Czisch, M., Sämann, P.G., and Spoormaker, V.I. (2018). Disentangling reward anticipation with simultaneous pupillometry / fMRI. *NeuroImage* 178, 11–22.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A Neural Substrate of Prediction and Reward. *Science* 275, 1593–1599.

Seaman, K.L., Brooks, N., Karrer, T.M., Castellon, J.J., Perkins, S.F., Dang, L.C., Hsu, M., Zald, D.H., and Samanez-Larkin, G.R. (2018). Subjective value representations during effort, probability and time discounting across adulthood. *Social Cognitive and Affective Neuroscience* 13, 449–459.

Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., and Greicius, M.D. (2007). Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *Journal of Neuroscience* 27, 2349–2356.

Sellitto, M., Ciaramelli, E., and di Pellegrino, G. (2010). Myopic Discounting of Future Rewards after Medial Orbitofrontal Damage in Humans. *Journal of Neuroscience* 30, 16429–16436.

Sénèque (2005). *La Vie heureuse* (GF-Flammarion).

Sescousse, G., Caldú, X., Segura, B., and Dreher, J.-C. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews* 37, 681–696.

## Bibliographie

- Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., and Davidson, R.J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat Rev Neurosci* *12*, 154–167.
- Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A. (2019). Movement Vigor as a Reflection of Subjective Economic Utility. *Trends in Neurosciences* *42*, 323–336.
- Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical Journal* *27*, 379–423.
- Shapiro, A.D., and Grafton, S.T. (2020). Subjective value then confidence in human ventromedial prefrontal cortex. *PLoS ONE* *15*, e0225617.
- Shen, K.-K., Welton, T., Lyon, M., McCorkindale, A.N., Sutherland, G.T., Burnham, S., Fripp, J., Martins, R., and Grieve, S.M. (2019). Structural core of the executive control network: A high angular resolution diffusion MRI study. *Hum Brain Mapp*.
- Shenhav, A., Botvinick, M.M., and Cohen, J.D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron* *79*, 217–240.
- Shenhav, A., Straccia, M.A., Cohen, J.D., and Botvinick, M.M. (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature Neuroscience* *17*, 1249–1254.
- Shenhav, A., Cohen, J.D., and Botvinick, M.M. (2016a). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience* *19*, 1286–1291.
- Shenhav, A., Straccia, M.A., Botvinick, M.M., and Cohen, J.D. (2016b). Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice.
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T.L., Cohen, J.D., and Botvinick, M.M. (2017). Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience* *40*, 99–124.
- Silvestrini, N. (2017). Psychological and neural mechanisms associated with effort-related cardiovascular reactivity and cognitive control: An integrative approach. *International Journal of Psychophysiology* *119*, 11–18.
- Silvetti, M., Seurinck, R., and Verguts, T. (2013). Value and prediction error estimation account for volatility effects in ACC: A model-based fMRI study. *Cortex* *49*, 1627–1635.
- Simmel, G. (2009). Partie analytique, 2e chapitre, sections 1 et 2. In *Philosophie de l'argent*, (GF-Flammarion), p.
- Skvortsova, V., Palminteri, S., and Pessiglione, M. (2014). Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates. *Journal of Neuroscience* *34*, 15621–15630.
- Smit, A.S., Eling, P.A.T.M., and Coenen, A.M.L. (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica* *115*, 35–42.

## Bibliographie

- Smit, A.S., Eling, P.A.T.M., Hopman, M.T., and Coenen, A.M.L. (2005). Mental and physical effort affect vigilance differently. *International Journal of Psychophysiology* 57, 211–217.
- Smith, A. (1776). *Recherches sur la nature et les causes de la richesse des nations*. 237.
- Smith, D.V., Hayden, B.Y., Truong, T.-K., Song, A.W., Platt, M.L., and Huettel, S.A. (2010a). Distinct Value Signals in Anterior and Posterior Ventromedial Prefrontal Cortex. *Journal of Neuroscience* 30, 2490–2495.
- Smith, K.S., Mahler, S.V., Peciña, S., and Berridge, K.C. (2010b). Part I: Hedonic Hotspots: Generating Sensory Pleasure in the Brain. In *Pleasures of the Brain*, (Oxford University Press: New York), p.
- Spinoza, B. (1965). *Ethique* (GF-Flammarion).
- Spooner, J., Yu, H., Kao, C., Sillay, K., and Konrad, P. (2007). Neuromodulation of the cingulum for neuropathic pain after spinal cord injury. Case report. *J. Neurosurg.* 107, 169–172.
- Spraker, M.B., Yu, H., Corcos, D.M., and Vaillancourt, D.E. (2007). Role of individual basal ganglia nuclei in force amplitude generation. *J. Neurophysiol.* 98, 821–834.
- Sridharan, D., Levitin, D.J., and Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *PNAS* 105, 12569–12574.
- Stanovich, K.E., and West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences* 23, 645–665.
- Starr, C.J., Sawaki, L., Wittenberg, G.F., Burdette, J.H., Oshiro, Y., Quevedo, A.S., and Coghill, R.C. (2009). Roles of the insular cortex in the modulation of pain: insights from brain lesions. *J. Neurosci.* 29, 2684–2694.
- Stephani, C., Fernandez-Baca Vaca, G., Maciunas, R., Koubeissi, M., and Lüders, H.O. (2011). Functional neuroanatomy of the insular lobe. *Brain Struct Funct* 216, 137–149.
- Stern, E.R., Gonzalez, R., Welsh, R.C., and Taylor, S.F. (2010). Updating Beliefs for a Decision: Neural Correlates of Uncertainty and Underconfidence. *Journal of Neuroscience* 30, 8032–8041.
- Stevens, J.C., and Mack, J.D. (1959). Scales of Apparent Force. *Journal of Experimental Psychology* 58, 405.
- Strait, C.E., Blanchard, T.C., and Hayden, B.Y. (2014). Reward Value Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex. *Neuron* 82, 1357–1366.
- Strait, C.E., Slezzer, B.J., and Hayden, B.Y. (2015). Signatures of Value Comparison in Ventral Striatum Neurons. *PLOS Biology* 13, e1002173.
- Sugiwaka, H., and Okouchi, H. (2004). Reformative self-control and discounting of reward value by delay or effort<sup>1</sup>. *Japanese Psychol Res* 46, 1–9.

## Bibliographie

- Sun, S., Zhen, S., Fu, Z., Wu, D.-A., Shimojo, S., Adolphs, R., Yu, R., and Wang, S. (2017). Decision ambiguity is mediated by a late positive potential originating from cingulate cortex. *NeuroImage* *157*, 400–414.
- Suzuki, S., Cross, L., and O’Doherty, J.P. (2017). Elucidating the underlying components of food valuation in the human orbitofrontal cortex. *Nat. Neurosci.* *20*, 1780–1786.
- Takahashi, T., Ikeda, K., and Hasegawa, T. (2007). A hyperbolic decay of subjective probability of obtaining delayed rewards. *Behav Brain Funct* *3*, 52.
- Takeuchi, T., Puntous, T., Tuladhar, A., Yoshimoto, S., and Shirama, A. (2011). Estimation of Mental Effort in Learning Visual Search by Measuring Pupil Response. *PLoS ONE* *6*, e21973.
- Talmi, D., Dayan, P., Kiebel, S.J., Frith, C.D., and Dolan, R.J. (2009). How humans integrate the prospects of pain and reward during choice. *J. Neurosci.* *29*, 14617–14626.
- Tanner, W.P., and Swets, J.A. (1954). A decision-making theory of visual detection. *9*.
- Teh, Y.W., Bapst, V., Czarnecki, W.M., Quan, J., Kirkpatrick, J., Hadsell, R., Heess, N., and Pascanu, R. (2017). Distral: Robust Multitask Reinforcement Learning. *ArXiv:1707.04175 [Cs, Stat]*.
- Thorndike, E.L. (1913). *The psychology of learning* (Teachers College, Columbia University).
- Thorndike, E.L. (Edward L. (1898). *Animal intelligence: an experimental study of the associative processes in animals* (New York : Macmillan).
- Thorndike, E.L. (Edward L. (1911). *Animal intelligence; experimental studies* (New York, The Macmillan company).
- Thura, D., Beauregard-Racine, J., Fradet, C.-W., and Cisek, P. (2012). Decision making by urgency gating: theory and experimental support. *J. Neurophysiol.* *108*, 2912–2930.
- Tobler, P.N., and Weber, E.U. (2014). Chapter 9 - Valuation for Risky and Uncertain Choices. In *Neuroeconomics* (Second Edition), P.W. Glimcher, and E. Fehr, eds. (San Diego: Academic Press), pp. 149–172.
- Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science* *315*, 515–518.
- Tsai, L.S. (1932). The laws of minimum effort and maximum satisfaction in animal behavior (See Description).
- Turner, R.S., and Desmurget, M. (2010). Basal ganglia contributions to motor control: a vigorous tutor. *Current Opinion in Neurobiology* *20*, 704–716.
- Turner, R.S., Desmurget, M., Grethe, J., Crutcher, M.D., and Grafton, S.T. (2003). Motor subcircuits mediating the control of movement extent and speed. *J. Neurophysiol.* *90*, 3958–3966.
- Tusche, A., Bode, S., and Haynes, J.-D. (2010). Neural Responses to Unattended Products Predict Later Consumer Choices. *Journal of Neuroscience* *30*, 8024–8031.

## Bibliographie

Tversky, A., and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185, 1124–1131.

Ubfal, D. (2016). How general are time preferences? Eliciting good-specific discount rates. *Journal of Development Economics* 118, 150–170.

Uddin, L.Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience* 16, 55–61.

Urai, A.E., Braun, A., and Donner, T.H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications* 8, 14637.

Vaccaro, A.G., and Fleming, S.M. (2018). Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain Neurosci Adv* 2.

Valenstein, E.S., and Beer, B. (1962). Reinforcing Brain Stimulation in Competition with Water Reward and Shock Avoidance. *Science* 137, 1052–1054.

Valenza, G., Sclocco, R., Duggento, A., Passamonti, L., Napadow, V., Barbieri, R., and Toschi, N. (2019). The central autonomic network at rest: Uncovering functional MRI correlates of time-varying autonomic outflow. *NeuroImage* 197, 383–390.

Van Cutsem, J., Marcora, S., De Pauw, K., Bailey, S., Meeusen, R., and Roelands, B. (2017). The Effects of Mental Fatigue on Physical Performance: A Systematic Review. *Sports Med* 47, 1569–1588.

Vassena, E., Silvetti, M., Boehler, C.N., Achten, E., Fias, W., and Verguts, T. (2014). Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS ONE* 9, e91008.

Vassena, E., Holroyd, C.B., and Alexander, W.H. (2017). Computational Models of Anterior Cingulate Cortex: At the Crossroads between Prediction and Effort. *Front. Neurosci.* 11, 316.

Verberne, A.J., and Owens, N.C. (1998). Cortical modulation of the cardiovascular system. *Prog. Neurobiol.* 54, 149–168.

Vibert, J.-F., Sébille, A., Lavallard-Rousseau, M.-C., and Boureau, F. (2005). *Neurophysiologie: De la physiologie à l’exploration fonctionnelle* (Elsevier).

Vinckier, F., Rigoux, L., Oudiette, D., and Pessiglione, M. (2018). Neuro-computational account of how mood fluctuations arise and affect decision making. *Nature Communications* 9.

Volz, K.G., Schubotz, R.I., and von Cramon, D.Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *NeuroImage* 21, 848–857.

Volz, K.G., Schubotz, R.I., and Cramon, D.Y. von (2005). Variants of uncertainty in decision-making and their neural correlates. *Brain Research Bulletin* 67, 403–412.

Von Neumann, J., and Morgenstern, O. (1944). *Theory of games and economic behavior*.

## Bibliographie

- Wager, T.D., Sylvester, C.-Y.C., Lacey, S.C., Nee, D.E., Franklin, M., and Jonides, J. (2005). Common and unique components of response inhibition revealed by fMRI. *NeuroImage* 27, 323–340.
- Wager, T.D., Atlas, L.Y., Lindquist, M.A., Roy, M., Woo, C.-W., and Kross, E. (2013). An fMRI-Based Neurologic Signature of Physical Pain. *New England Journal of Medicine* 368, 1388–1397.
- Walasek, L., and Stewart, N. (2015). How to Make Loss Aversion Disappear and Reverse: Tests of the Decision by Sampling Origin of Loss Aversion. *J Exp Psychol Gen* 144, 7–11.
- Walasek, L., and Stewart, N. (2019). Context-dependent sensitivity to losses: Range and skew manipulations. *J Exp Psychol Learn Mem Cogn* 45, 957–968.
- Wallis, J.D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature Neuroscience* 15, 13–19.
- Walras, L. (1874). *Éléments d'économie politique pure; ou, Théorie de la richesse sociale* (Corbaz).
- Wandell, B.A., Dumoulin, S.O., and Brewer, A.A. (2007). Visual Field Maps in Human Cortex. *Neuron* 56, 366–383.
- Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.
- Waters, R.H. (1937). The Principle of Least Effort in Learning. *The Journal of General Psychology* 16, 3–20.
- Watkins, C.J.C.H., and Dayan, P. (1992). Q-learning. *Mach Learn* 8, 279–292.
- Watson, J.B., and Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology* 3, 1–14.
- Weidemann, C.T., and Kahana, M.J. (2016). Assessing recognition memory using confidence ratings and response times. *R Soc Open Sci* 3, 150670.
- Weilbacher, R., and Gluth, S. (2016). The Interplay of Hippocampus and Ventromedial Prefrontal Cortex in Memory-Based Decision Making. *Brain Sciences* 7, 4.
- Weinstein, A.W. and Y. (2014). Exercise Addiction- Diagnosis, Bio-Psychological Mechanisms and Treatment Issues.
- Weiskopf, N., Hutton, C., Josephs, O., Turner, R., and Deichmann, R. (2007). Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. *Magnetic Resonance Materials in Physics, Biology and Medicine* 20, 39–49.
- van der Wel, P., and van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review* 25, 2005–2015.

## Bibliographie

- Westbrook, A., Kester, D., and Braver, T.S. (2013). What Is the Subjective Cost of Cognitive Effort? Load, Trait, and Aging Effects Revealed by Economic Preference. *PLoS ONE* 8, e68210.
- Westbrook, A., Lamichhane, B., and Braver, T. (2019). The Subjective Value of Cognitive Effort is Encoded by a Domain-General Valuation Network. *The Journal of Neuroscience* 3071–18.
- Wheeler, R.H. (1940). *The science of psychology*, 2nd ed (Oxford, England: Crowell).
- White, T.P., Engen, N.H., Sørensen, S., Overgaard, M., and Shergill, S.S. (2014). Uncertainty and confidence from the triple-network perspective: Voxel-based meta-analyses. *Brain and Cognition* 85, 191–200.
- Wiech, K., Lin, C. -s., Brodersen, K.H., Bingel, U., Ploner, M., and Tracey, I. (2010). Anterior Insula Integrates Information about Salience into Perceptual Decisions about Pain. *Journal of Neuroscience* 30, 16324–16331.
- Will Crescioni, A., Ehrlinger, J., Alquist, J.L., Conlon, K.E., Baumeister, R.F., Schatschneider, C., and Dutton, G.R. (2011). High trait self-control predicts positive health behaviors and success in weight loss. *J Health Psychol* 16, 750–759.
- Wong, K.-F. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Frontiers in Computational Neuroscience* 1.
- Wu, S.-W., Delgado, M.R., and Maloney, L.T. (2011). The Neural Correlates of Subjective Utility of Monetary Outcome and Probability Weight in Economic and in Motor Decision under Risk. *Journal of Neuroscience* 31, 8822–8831.
- Wunderlich, K., Rangel, A., and O’Doherty, J.P. (2010). Economic choices can be made using only stimulus values. *Proceedings of the National Academy of Sciences* 107, 15005–15010.
- Yamada, H., Louie, K., Tymula, A., and Glimcher, P.W. (2018). Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nature Communications* 9.
- Yechiam, E., and Hochman, G. (2013). Losses as modulators of attention: review and analysis of the unique effects of losses over gains. *Psychol Bull* 139, 497–518.
- Yi, R., de la Piedad, X., and Bickel, W.K. (2006). The combined effects of delay and probability in discounting. *Behavioural Processes* 73, 149–155.
- Yin, H.H., and Knowlton, B.J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476.
- Yttri, E.A., and Dudman, J.T. (2016). Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature* 533, 402–406.
- Yu, A., and Dayan, P. (2003). *Expected and Unexpected Uncertainty: ACh and NE in the Neocortex.* (MIT Press), pp. 157–164.



## Bibliographie

- Zald, D.H., McHugo, M., Ray, K.L., Glahn, D.C., Eickhoff, S.B., and Laird, A.R. (2014). Meta-Analytic Connectivity Modeling Reveals Differential Functional Connectivity of the Medial and Lateral Orbitofrontal Cortex. *Cerebral Cortex* 24, 232–248.
- Zarr, N., and Brown, J.W. (2016). Hierarchical error representation in medial prefrontal cortex. *NeuroImage* 124, 238–247.
- Zénon, A. (2019). Eye pupil signals information gain. *Proceedings of the Royal Society B: Biological Sciences* 286, 20191593.
- Zénon, A., Sidibé, M., and Olivier, E. (2014). Pupil size variations correlate with physical effort perception. *Frontiers in Behavioral Neuroscience* 8.
- Zénon, A., Sidibe, M., and Olivier, E. (2015). Disrupting the Supplementary Motor Area Makes Physical Effort Appear Less Effortful. *Journal of Neuroscience* 35, 8737–8744.
- Zénon, A., Solopchuk, O., and Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia* 123, 5–18.
- Zhang, Y.-Y., Xu, L., Liang, Z.-Y., Wang, K., Hou, B., Zhou, Y., Li, S., and Jiang, T. (2018). Separate Neural Networks for Gains and Losses in Intertemporal Choice. *Neurosci. Bull.* 34, 725–735.
- Ziegler, J.C., Montant, M., Briesemeister, B.B., Brink, T.T., Wicker, B., Ponz, A., Bonnard, M., Jacobs, A.M., and Braun, M. (2018). Do Words Stink? Neural Reuse as a Principle for Understanding Emotions in Reading. *Journal of Cognitive Neuroscience* 30, 1023–1032.
- Zipf, G.K. (1949). *Human Behavior And The Principle Of Least Effort*.

# Appendices

- I. Article en anglais en lien avec la première étude**
- II. Enregistrement de l'activité du vmPFC à l'aide de différentes séquences d'IRMf**
- III. Utiliser un modèle computationnel pour inférer les valeurs cachées associées au comportement**
- IV. Valeur et attention**

## **I Article en anglais en lien avec la première étude**

### **Value, confidence and deliberation in preference tasks: a triple dissociation in the medial prefrontal cortex**

N. Clairis<sup>1</sup>, M. Pessiglione<sup>1</sup>

(1) Motivation, Brain and Behavior team, Institut du Cerveau (ICM), Inserm U1127, CNRS U7225, Sorbonne Université, Paris, France

#### **Abstract**

Deciding about courses of action involves an estimation of costs and benefits. Decision neuroscience studies have suggested a dissociation between the ventral and dorsal medial prefrontal cortex (vmPFC and dmPFC), which would process reward value and effort cost, respectively. However, several results appeared inconsistent with this general idea of opponent reward and effort systems. These contradictions might reflect the diversity of tasks used to investigate the trade-off between effort cost and reward value. They might also reflect the confusion with a meta-decision process about the amount of cognitive control to exert in order to reach a sufficient confidence in the reward/effort estimates. Here, we used fMRI to examine the neural correlates of reward and effort estimates across several preference tasks, from (dis-)likeability ratings to binary decisions involving attribute integration and option comparison. Results confirm the role of the vmPFC as a generic valuation system, across the different tasks (likeability rating or binary decision) and attributes (the activity increasing with reward value and decreasing with effort cost). However, meta-decision variables were represented in more dorsal regions, with confidence in the mPFC and deliberation time in the dmPFC. These findings suggest that assessing commonalities across preference tasks and distinguishing between decision and meta-decision variables might help reaching a unified view of how the brain chooses a course of action.

### **Introduction**

Cost-benefit arbitrage is a pervasive process involved in any decision that bears practical consequences. Standard decision theory assumes that selecting a course of action can be reduced to maximizing a net utility function, where expected benefits are discounted by expected costs. In decision neuroscience, a general debate has been raging between opponency theories, where systems pushing and blocking an action are anatomically segregated (Boureau and Dayan, 2011), and integration theories, where costs and benefits are both estimated in a single brain system (Pessiglione and Delgado, 2015).

A central dilemma that has been scrutinized in decision neuroscience is between the resource to be invested in the action (i.e., the effort) and the value of the action outcome (i.e., the reward). Many fMRI studies have documented the existence of a brain valuation system (BVS), with the ventromedial prefrontal cortex (vmPFC) as a core component, that would signal the value of various categories of rewards (Bartra et al., 2013; Levy and Glimcher, 2012). Although the generality of the vmPFC value signal across reward categories is well accepted, whether or not this value signal integrates effort costs is a controversial issue. While some studies have reported that effort costs result in a decreased vmPFC value signal (Arulpragasam et al., 2018; Seaman et al., 2018), others have suggested the existence of an opponent system that would signal effort costs with increasing activity, such as the dorsomedial prefrontal cortex (dmPFC, sometimes referred to as ‘dorsal anterior cingulate cortex’) and the anterior insula (Kurniawan et al., 2013; Skvortsova et al., 2014).

One explanation for these discrepancies is the diversity of behavioral tasks employed to probe the neural correlates of reward and effort estimates. Here, we reversed the general logic of standard functional neuroimaging approach, which specifies the roles of brain regions with contrasts that isolate minimal differences between conditions. On the contrary, we intended to generalize our findings across various conditions and tasks, with the aim to reach more robust conclusions. Thus, we employed a series of preference tasks (also called ‘value-based’ tasks) that enable the investigation of 1) the reward value or effort cost of a single option, with (dis-)likeability rating tasks, 2) the comparison between reward or effort attributes, with one dimension – two options (1D-2O) choice tasks and 3) the integration of reward and effort attributes, with one option – two dimensions (1O-2D) choice tasks.

Another explanation for inconsistencies in the neural correlates of reward and effort is the possible confusion with meta-decision variables. Indeed, reward value and effort cost

## Appendices

estimation can itself be performed with various amounts of effort. Intuitively, everyone knows that some decisions are made on the fly, without thinking too much, while others demand a long and painful deliberation, in order to weigh the pros and cons. This intuition has been captured in a theory postulating that the investment of cognitive control is based on a cost-benefit arbitrage (Shenhav et al., 2013). Applied to decision-making, the cost is essentially the time that the agent spends considering the different options and their attributes, while the benefit is essentially the confidence that the agent gains in making the right choice. Therefore, in addition to the reward and effort levels that are attributes of choice options, we examined constructs related to the meta-choice: response confidence and deliberation time. Importantly, the reward and effort estimates did not imply actual and immediate implementation, as all choices were hypothetical in our tasks. On the contrary, confidence and time were variables associated to current behavior, i.e. to the ongoing estimation of reward value and/or effort cost.

Our results suggest a triple dissociation within the medial prefrontal cortex that is stable across preference tasks, with option attributes (reward and effort levels) being represented in vmPFC activity, whereas meta-decision variables (confidence and deliberation time) being represented in mPFC and dmPFC activity, respectively.

## Results

### Behavior

Participants ( $n=39$  in total) first performed a series of ratings, divided into three blocks (Fig. 1A). Each block presented 72 items one by one: reward items presented with text + image ( $R_{ti}$ ), reward items presented with text only ( $R_t$ ) and effort items presented with text only ( $E_t$ ). The reason for varying the mode of presentation was to assess the generality of the neural valuation process across different inputs that demand more or less imagination, according to previous study (Lebreton et al., 2013). The order of the  $R_{ti}$ ,  $R_t$  and  $E_t$  blocks was counterbalanced across participants. For rewards, participants were asked to rate how much they would like it, should they be given the item immediately after the experiment. Symmetrically, the instruction for efforts was to rate how much they would dislike it, should they be requested to implement it immediately after the experiment. We included both food and non-food (goodies) reward items, and both mental and physical effort items. There was no number on the scale, just labels on endpoints, and ratings were pseudo-continuous, from ‘I would (dis-)like it not at all’ to ‘enormously’. Thus, the left endpoint corresponded to indifference and the right endpoint to extreme attraction or extreme aversion (Fig. 2A).

## Appendices

The z-scored rating was taken as a proxy for stimulus value (Val) in this task, while the square of z-score rating was taken as a proxy for response confidence (Conf), following on previous studies (Lebreton et al., 2015; Lopez-Persem et al., 2020). Deliberation time (DT) was defined in this task as the time between item onset and the first button press used to move the cursor along the scale. DT was regressed against a linear model that included Val and Conf proxies, in addition to nuisance factors (such as jitter duration, stimulus luminance, text length and trial index, see methods). We found in all blocks (Fig. 2B) a significant effect of both value ( $R_{ti}$ :  $\beta_{Val} = -0.208 \pm 0.023$  (mean  $\pm$  standard error of the mean across participants),  $p = 4 \cdot 10^{-11}$ ;  $R_t$ :  $\beta_{Val} = -0.169 \pm 0.019$ ,  $p = 6 \cdot 10^{-11}$ ;  $E_t$ :  $\beta_{Val} = 0.261 \pm 0.028$ ,  $p = 2 \cdot 10^{-11}$ ) and confidence ( $R_{ti}$ :  $\beta_{Conf} = -0.174 \pm 0.025$ ,  $p = 3 \cdot 10^{-8}$ ;  $R_t$ :  $\beta_{Conf} = -0.188 \pm 0.028$ ,  $p = 7 \cdot 10^{-8}$ ;  $E_t$ :  $\beta_{Conf} = -0.130 \pm 0.040$ ;  $p = 0.0024$ ). Thus, participants were faster to provide their rating when the item was more appetitive (or less aversive) and when they were more confident (going towards the extremes of the rating scale). Among the nuisance factors, we observed effects of jitter duration, stimulus luminance and text length, which were therefore included as regressors in subsequent analyses. However, there was no significant effect of trial index, which discards a possible contamination of DT by habituation or fatigue.

Then participants performed a series of binary choices, either 1D-2O choices or 1O-2D choices. The choice tasks were always performed after the rating tasks because the ratings were used to control the difficulty of ratings (i.e., the difference in value between the two options). In the 1D-2O choice task (Fig. 1B), participants were asked to select the reward they would prefer to receive at the end of the experiment, if they were offered one of two options, or the effort they would prefer to exert, if they were forced to implement one of two options. Thus, the two options always pertained to the same dimension (reward or effort), and even to the same subcategory (food or good for reward, mental or physical for effort), to avoid shortcut of deliberation by general preference. The mode of presentation (text or image) was also the same for the two options, to avoid biasing the choice by a difference in salience. To obtain a same number of trials as in the rating task, each item was presented twice, for a total of three blocks ( $R_{ti}$ ,  $R_t$ ,  $E_t$ ) of 72 choices. Again, the order of the  $R_{ti}$ ,  $R_t$  and  $E_t$  blocks was counterbalanced across participants. In the 1O-2D choice task (Fig. 1C), participants were asked whether they would be willing to exert an effort in order to obtain a reward, at the end of the experiment. Only items described with text were retained for this task, each item again appearing twice, for a total of 144 choices divided into three blocks.

## Appendices

The 1D-2O choice task (Fig. 1C) was meant to assess value comparison between the two options, within a same dimension. The decision value (DV) in this task was defined as the difference in (dis-)likeability rating between the two options. We checked with a logistic regression that DV was a significant predictor of choices (Fig. 2E) in all blocks ( $R_{ti}$ :  $\beta_{DV} = 3.383 \pm 0.273$ ,  $p = 7 \cdot 10^{-15}$ ;  $R_t$ :  $\beta_{DV} = 2.669 \pm 0.155$ ,  $p = 2 \cdot 10^{-19}$ ;  $E_t$ :  $\beta_{DV} = -2.278 \pm 0.157$ ,  $p = 4 \cdot 10^{-17}$ ). The 1O-2D choice task was meant to assess value integration across two dimensions, for a single option. The decision value (or net value) in this task was defined as a linear combination of reward and effort ratings. We checked with a logistic regression that both reward and effort ratings were significant predictors of choice in this task ( $\beta_R = 1.502 \pm 0.085$ ,  $p = 6 \cdot 10^{-20}$ ;  $\beta_E = -1.117 \pm 0.079$ ,  $p = 1 \cdot 10^{-16}$ ).

To analyze DT (time between stimulus onset and button press), we defined proxies for stimulus value and response confidence, as we did for the rating task. Stimulus value (Val) was defined as a linear integration of the likeability ratings assigned to the two stimuli on screen. In the 1D-2O choice task, this is simply the sum of the two item ratings. In the 1O-2D choice task, this is the net value (sum of reward and effort ratings weighted by a scaling factor). In both cases, choice probability was calculated with the logistic regression model (softmax function of decision value). Response confidence (Conf) was defined, by analogy to the rating task, as the square of the difference between choice probability and mean choice rate. Linear regression showed that DT decreased with value (Fig. 2C) in the 1D-2O choice task ( $R_{ti}$ :  $\beta_{Val} = -0.061 \pm 0.010$ ,  $p = 3 \cdot 10^{-7}$ ;  $R_t$ :  $\beta_{Val} = -0.061 \pm 0.010$ ,  $p = 3 \cdot 10^{-7}$ ;  $E_t$ :  $\beta_{Val} = 0.048 \pm 0.013$ ,  $p = 8 \cdot 10^{-4}$ ), albeit not in the 1O-2D choice task ( $\beta_{Val} = 0.033 \pm 0.024$ ,  $p = 0.172$ ). DT also decreased with confidence (Fig. 2D) in both the 1D-2O choice task ( $R_{ti}$ :  $\beta_{Conf} = -1.738 \pm 0.201$ ,  $p = 2 \cdot 10^{-10}$ ;  $R_t$ :  $\beta_{Conf} = -1.975 \pm 0.184$ ,  $p = 4 \cdot 10^{-13}$ ;  $E_t$ :  $\beta_{Conf} = -1.731 \pm 0.221$ ,  $p = 2 \cdot 10^{-9}$ ) and the 1O-2D choice task ( $\beta_{Conf} = -1.148 \pm 0.145$ ,  $p = 1 \cdot 10^{-9}$ ). Thus, the relationship between DT and ratings was similar in rating and choice tasks: participants were faster when the options were more appetitive (or less aversive) and when they were more confident (because of a strong preference for one response or the other).

### Neural activity

The aim of fMRI data analysis was to dissociate the variables related to option valuation (reward and effort estimates) from the variables related to the meta-decision (confidence and deliberation) across value-based tasks (rating and choice). A meta-analysis of fMRI studies using Neurosynth platform (Fig. 3A) shows that value, confidence and effort keywords yield to

## Appendices

similar activation patterns with clusters in both vmPFC and dmPFC. To better dissociate the neural correlates of these constructs, we built a general linear model where stimulus onset events were modulated by three parameters of interest - Val, Conf and DT (defined as in the behavior analysis). Nuisance parameters that were found to influence DT (jitter duration, stimulus luminance, text length) were also included as modulators of stimulus onset events, before the variables of interest. Thus, due to serial orthogonalization, the variables of interest were orthogonalized with respect to nuisance factors, and deliberation time was made orthogonal to all other regressors, including stimulus value and response confidence.

After correction for multiple comparisons at the voxel level, we found only three significant clusters in the prefrontal cortex (Fig. 3B): Val was signaled in vmPFC activity, Conf in mPFC activity and DT in dmPFC activity. At this threshold, we also observed significant association with Val in other brain regions (see Table 1), such as the posterior cingulate cortex (pCC) and the primary visual cortex (V1). At a more lenient threshold ( $p < 0.05$  corrected at the cluster level), the ventral striatum (vS) was also bilaterally associated to Val. Note that vS and pCC are standard components of the brain valuation system, whereas V1 activation is likely to be an artifact of gaze position on the rating scale, as it was not observed in the choice tasks. Consistently, positive correlation with Val was found in right V1 activity, and negative correlation in left V1 activity. This was not the case in the other clusters, which were either medial or bilateral.

We further analyzed the relationship between computational variables and activity in the three medial prefrontal regions of interest (ROI) with post-hoc t-tests. We first verified that the three main associations were significant in each task (Fig. 3C): it was indeed the case for Val in vmPFC activity (rating:  $\beta_{\text{Val}} = 0.701 \pm 0.126$ ,  $p = 2 \cdot 10^{-6}$ ; 1D-2O:  $\beta_{\text{Val}} = 0.293 \pm 0.126$ ,  $p = 0.025$ ; 1O-2D:  $\beta_{\text{Val}} = 0.700 \pm 0.180$ ,  $p = 4 \cdot 10^{-4}$ ), for Conf in mPFC activity (rating:  $\beta_{\text{Conf}} = 0.768 \pm 0.111$ ,  $p = 3 \cdot 10^{-8}$ ; 1D-2O:  $\beta_{\text{Conf}} = 0.335 \pm 0.114$ ,  $p = 0.006$ ; 1O-2D:  $\beta_{\text{Conf}} = 0.313 \pm 0.133$ ,  $p = 0.024$ ) and for DT in dmPFC activity (rating:  $\beta_{\text{DT}} = 0.410 \pm 0.106$ ,  $p = 4 \cdot 10^{-4}$ ; 1D-2O:  $\beta_{\text{DT}} = 0.875 \pm 0.123$ ,  $p = 2 \cdot 10^{-8}$ ; 1O-2D:  $\beta_{\text{DT}} = 0.676 \pm 0.125$ ,  $p = 4 \cdot 10^{-6}$ ). We then verified that the three dissociations were significant (Fig. 4D): Val was better reflected in vmPFC activity ( $\beta_{\text{Val/vmPFC}} > \beta_{\text{Val/mPFC}} : p = 9 \cdot 10^{-9}$ ;  $\beta_{\text{Val/vmPFC}} > \beta_{\text{Val/dmPFC}} : p = 5 \cdot 10^{-8}$ ), Conf in mPFC activity ( $\beta_{\text{Conf/mPFC}} > \beta_{\text{Conf/vmPFC}} : p = 0.0035$ ;  $\beta_{\text{Conf/mPFC}} > \beta_{\text{Conf/dmPFC}} : p = 2 \cdot 10^{-7}$ ) and DT in dmPFC ( $\beta_{\text{DT/dmPFC}} > \beta_{\text{DT/vmPFC}} : p = 0.021$ ;  $\beta_{\text{DT/dmPFC}} > \beta_{\text{DT/mPFC}} : p = 2 \cdot 10^{-4}$ ). Thus, the triple dissociation observed in the maps corresponded to significant differences between ROI and significant associations that were robust across tasks. However, the triple dissociation does not



## Appendices

imply that the three variables of interest were solely represented in one single brain region. In particular, Conf and DT were also significantly related to vmPFC activity ( $\beta_{\text{Conf}} = 0.256 \pm 0.098$ ,  $p = 0.013$ ;  $\beta_{\text{DT}} = 0.377 \pm 0.108$ ,  $p = 0.001$ ), even if these activities were dominated by Val and DT, respectively.

We developed variants of our GLM to assess the robustness of these findings. Regarding Val, we examined whether other value-related variables employed in previous studies, such as decision value (e.g., chosen minus unchosen option value) could better account for vmPFC activity. When replacing our proxy for stimulus value by these other variables, we did not find any stronger correlation with activity in the vmPFC, even when defined anatomically or from meta-analyses of fMRI data. Regarding Conf, we assessed whether the dorsal location of the cluster (mPFC instead of vmPFC) could be related to the orthogonalization with Val, by simply removing the Val regressor, but results were unchanged. Regarding DT, we tested whether the association with dmPFC activity could arise from the delay itself, and not from a prolonged deliberation. When replacing the delta function modeling stimulus onset by a boxcar function (whose duration varies with DT), the association between DT and dmPFC activity was still significant, suggesting a modulation in amplitude and not just a prolongation of the signal.

We looked for further generalization of the valuation signal, not solely across tasks but also across stimuli. We focused on the rating task, in which the link with neural activity is easier to assess, as there is only one stimulus to value (Fig. 3D). FMRI time series were regressed against a GLM that separated stimulus categories (Rti, Rt and Et) into different onset regressors, each modulated by corresponding ratings. Results show that vmPFC activity was positively related to the value (likeability rating) of reward items, whether or not they are presented with an image, and negatively correlated to the cost (dislikeability rating) of effort items (Rti:  $\beta_{\text{Val}} = 0.630 \pm 0.224$ ,  $p = 0.008$ ; Rt:  $\beta_{\text{Val}} = 0.803 \pm 0.177$ ,  $p = 6 \cdot 10^{-5}$ ; Et:  $\beta_{\text{Val}} = -0.670 \pm 0.205$ ,  $p = 0.002$ ). Thus, the association between Val and vmPFC activity was independent of the presentation mode, and integrated costs as well as benefits. Importantly, the association with reward value or effort cost was not observed in putative opponent brain regions such as the dmPFC, whose activity tended to decrease with dislikeability rating of effort items (Et:  $\beta_{\text{Val}} = -0.225 \pm 0.116$ ,  $p = 0.060$ ).

Thus, it appeared that dmPFC activity did not reflect the effort cost attached to the option on valuation but the effort cost of the meta-decision (selecting a response). Importantly, the association with DT was observed despite the fact that DT was orthogonalized to both value and confidence, suggesting that the dmPFC represents the effort invested above and beyond

## Appendices

that induced by the difficulty of value-based judgment or decision. As DT is a very indirect proxy for the effort invested in solving the task, we investigated the relationship with another proxy that has been repeatedly related to effort: pupil size. Neural activity was extracted in each ROI by fitting a GLM containing one event (stimulus onset) per trial. Then pupil size at each time point was regressed across trials against a GLM that contained nuisance factors (luminance, jitter duration, text length), variables of interest (Val, Conf, DT) and neural activity (vmPFC, mPFC, dmPFC).

A positive association between pupil size and dmPFC activity was observed in all rating and choice tasks (Fig. 4), about one second before the response. This association was not an artifact of the trial being prolonged (and therefore the response to luminance being cut at different durations) since it was observed both when locking time courses on stimulus onset and on motor response (button press). Finally, it was specific to the dmPFC ROI, and observed even if dmPFC was made independent (through serial orthogonalization) to all other variables (notably Val, Conf and DT). In particular, there was no consistent association between vmPFC and pupil size, suggesting that the correlates of DT observed in vmPFC were not related to effort but to some other factors affecting DT, such as mind-wandering.

## Discussion

In this study, we investigated the neural correlates of variables that are common to different tasks involving valuation of stimuli and expression of preferences. We observed a triple dissociation within the medial prefrontal cortex: stimulus value, response confidence and deliberation time were best reflected in vmPFC, mPFC and dmPFC activity, respectively. These associations between regions and variables were stable across rating and choice tasks, whether they involved likeability judgment, attribute integration or option comparison. They suggest that reward value and effort cost attached to choice options are integrated in a same brain region (vmPFC), while meta-decision variables such as response confidence and deliberation time are represented in distinct brain regions (mPFC and dmPFC).

Our results confirm the role attributed to the vmPFC as a generic valuation system (Bartra et al., 2013; Levy and Glimcher, 2012). The subjective value of reward items was reflected in vmPFC activity irrespective of the category (food versus goods), as was reported in many studies (Abitbol et al., 2015; Chib et al., 2009; Lebreton et al., 2009; Lopez-Persem et al., 2020). Also, vmPFC value signals were observed whether or not reward items were presented with images, suggesting that they can be extracted from both direct perceptual input or from text-based imagination which was shown to recruit episodic memory systems (Lebreton et al., 2013). Critically, our results show that the vmPFC also reflects the effort cost (whether mental or physical) attached to potential courses of actions. Therefore, they disprove opponent systems theories that would predict separate representations, notably those assuming that the vmPFC is involved in stimulus valuation, while action costs would implicate the dmPFC (Padoa-Schioppa, 2011; Rangel and Hare, 2010; Schneider and Koenigs, 2017). They rather suggest that the vmPFC might compute a net value, integrating reward benefit and effort cost, so as to prescribe whether or not an action is worth engaging. This idea is in line with previous demonstrations that the vmPFC integrates costs such as potential loss or delay in reward delivery (Hare et al., 2009, 2011; Kable and Glimcher, 2007; Talmi et al., 2009; Tom et al., 2007). Consistent with these studies, we observed that vmPFC activity increases with potential benefit and decreases with potential cost, which is compatible with the idea of net value computation.

The other medial prefrontal clusters (mPFC and dmPFC) were not affected by reward values or effort costs attached to choice options, but by variables related to providing a response in preference tasks, i.e. confidence and deliberation. We have previously argued that

## Appendices

confidence is what participants maximize when performing these tasks (Lebreton et al., 2015). Following on instructions, they intend to give the best judgment, or to make the right choice, even if these responses have no material consequence. In a sense, confidence can be conceived as a value, but this value would be attached to the response, not to the options. By construction, our proxy for confidence, defined as absolute (squared) deviation from the mean response, was orthogonal to stimulus value, defined as the sum of (positive) likeability of reward items and (negative) likeability of effort items. This proxy was found to elicit similar neural correlates as direct confidence ratings (De Martino et al., 2017; Lopez-Persem et al., 2020).

Confidence was the only variable significantly associated to mPFC activity, but was also reflected in vmPFC activity, as previously reported (Chua et al., 2006; De Martino et al., 2013; Gherman and Piliastides, 2018; Lebreton et al., 2015). Indeed, the addition of value and confidence signals in the vmPFC is a pattern that has been already observed in both fMRI and iEEG activity (Lebreton et al., 2015; Lopez-Persem et al., 2020). It has been argued that, in a binary choice, this pattern may denote the transition from the sum of option values to the difference between chosen and unchosen values (Hunt et al., 2012), this difference being globally positive and hence related to confidence. However, this explanation would not hold for the correlates of confidence observed outside choice tasks, for instance in likeability rating tasks. An alternative interpretation would be that, as a generic valuation device, the vmPFC computes both the value of options and the value of responses (i.e., confidence). The dissociation observed here, with value being better related to vmPFC activity and confidence better related to mPFC activity, is consistent with a general ventro-dorsal gradient from value to confidence representation that has been previously described (De Martino et al., 2017). On the contrary, dmPFC activity tended to decrease with confidence, but this association did not survive significance threshold.

The variable that was robustly associated with dmPFC activity was deliberation time. This variable was not orthogonal to the others, since it decreased both with stimulus value and response confidence. The link between deliberation time and stimulus value might arise from an appetitive Pavlovian reflex, as suggested in previous studies (Oudiette et al., 2019; Shadmehr et al., 2019), since there was no reason to go faster when valuating better rewards, or slower when valuating worse efforts, in our design. The link between deliberation time and response confidence might relate to the difficulty of the task (Kiani et al., 2014), i.e. the uncertainty about which rating or choice best reflects subjective preference. In our analyses, deliberation time was post-hoc orthogonalized with respect to the other variables, meaning that the association

## Appendices

with dmPFC activity was observed above and beyond the variance explained by stimulus value and response confidence.

This association alone would not yield a clear-cut interpretation, since many factors may affect response time. However, the systematic link observed between trial-wise dmPFC activation and the increase in pupil size before the response hints that this association might reflect the effort invested in the task. Indeed, pupil size has been linked to the intensity of not only physical effort, such as handgrip squeeze (Zénon et al., 2014) but also mental effort, such as focusing attention to resolve conflict or overcome task difficulty (Alnaes et al., 2014; Kahneman and Beatty, 1966; van der Wel and van Steenbergen, 2018). By contrast, we did not observe this systematic link with pupil size during deliberation with vmPFC activity. The link between vmPFC and deliberation time might therefore reflect other sources of variance, such as mind-wandering (being slower because of some off-task periods).

Our dmPFC ROI overlaps with clusters that have been labeled dorsal anterior cingulate cortex, or sometimes pre-supplementary motor area, in previous studies (Kamiński et al., 2017; Kolling et al., 2016; Shenhav et al., 2013). The association with deliberation time is compatible with a role attributed to this region in conflict monitoring, or in signaling the need to exert cognitive control (Botvinick et al., 2001; Shenhav et al., 2013). This functional role would also be consistent with the negative association between dmPFC activity and our confidence proxy, which is opposite to a conflict or ambiguity signal. In binary choices, longer deliberation when options are close is often captured as a shallower drift rate in accumulation-to-bound models (Steverson et al., 2019). However, our results call for a more general theory, because the link between dmPFC activity and deliberation time was also observed in rating tasks.

To recapitulate, we have teased apart the neural correlates of likeability, confidence and deliberation in the medial prefrontal cortex, which have been confused in previous fMRI studies, as shown by meta-analytic maps. The key distinction operated here is perhaps between effort as an attribute of choice option and effort as a resource allocated to the decision task, or in other words, between valuation applied to effort (implicating the vmPFC) and effort invested in valuation (implicating the dmPFC). This dissociation is consistent with the idea that the vmPFC anticipates the aversive value of a potential effort, while the dmPFC represents the intensity of effort when it must be exerted (Hogan et al., 2019). At a meta-decisional level, our results could be interpreted in the frame of a resource allocation model, where the effort or time invested in the deliberation is meant to increase confidence in the response, whether a rating or a choice (Lee and Daunizeau, 2019). This model would predict that increasing dmPFC activity

## Appendices

is meant to increase mPFC activity, which we could not test here because our correlational approach precludes any inference about causality. It is possible that dmPFC activation is not about demand for control but about estimating the amount of control being invested, on the request of other brain systems. It could even be that dmPFC activation corresponds to the aversive feeling induced by effort exertion, without any implication in meta-decisional regulation.

Finally, we have shown that the three associations hold across rating and choice tasks, and thus cannot be captured by models narrowly applied to the case of binary choice. Yet this approach (looking for stable associations across tasks) also bears limitations. Notably, our design would not allow comparing between conditions, as is traditionally done in neuroimaging studies. One may want for instance to compare between tasks and test whether brain regions are more involved in one or the other, but this would be confounded by several factors, such as the order (choice tasks being performed after rating tasks). A significant contrast would not be interpretable anyway, because there is more than one minimal difference between tasks. Thus, the aim to generalize the role of brain regions across tasks carries the inherent drawback of a limited specificity, but also the promises of a more robust understanding of anatomo-functional relationships. We hope this study will pave the way to further investigations following a similar approach, assessing a same concept across several tasks in a single study, instead of splitting tasks over separate reports, with likely inconsistent conclusions.

## Methods

### Subjects

In total, 40 right-handed volunteers participated in this fMRI study. Participants were recruited through the online RISC (Relais d'Information en Sciences de la Cognition) online platform (<https://www.risc.cnrs.fr/>). All participants were screened for the use of psychotropic medications, alcohol and drug use, and history of psychiatric disorders, cognitive/neurological disorders, and traumatic brain injury. One participant was excluded from all analyses because of a clear misunderstanding about task instructions, leaving  $n=39$  participants for behavioral data analysis (22 females / 17 males, aged  $25.4 \pm 4.1$  years). Another participant was excluded from the fMRI analysis due excessive movement inside the scanner ( $>3\text{mm}$  within-session per direction). Seven to nine additional participants were excluded from pupil size analysis, depending on the task, due to poor signal detection.

All participants gave informed consent and were paid a fixed amount for their participation. The 15 first subjects were paid 60€ and the 25 other subjects were paid 75€. The difference in payoff corresponds to a difference in scanning protocols, although all participants performed the same tasks. The first protocol ( $n=15$ ) aimed at comparing scanning sequences. Each task was subdivided into 3 sessions. Each session was scanned through a different scanning protocol using regular EPI, EPI with multiband acceleration, EPI with multiband + multi-echo acquisition. The main analysis only includes fMRI data recorded during the first session using regular EPI acquisition. The other participants were scanned with regular EPI during the nine sessions, which were all included in the analysis.

### Behavioral tasks

All tasks were programmed using Psychtoolbox (<http://psychtoolbox.org/>) in Matlab 2012 (The MathWorks, Inc., USA). Participants were given a 4-button box (fORP 932, Current Designs Inc, Philadelphia, USA) placed under their right to provide their responses. For further data analyses, stimulus luminance was calculated using standard function of red-green-blue composition  $0.299 \cdot \text{red} + 0.587 \cdot \text{green} + 0.114 \cdot \text{blue}$  (<http://www.w3.org/TR/AERT#color-contrast>), which was estimated through the Screen('GetImage') built-in psychtoolbox command. Stimuli comprised 144 reward items (72 food and 72 goods) and 72 effort items (36 mental and 36 physical). Half the reward items were presented with text only ( $R_t$  items), and the other half was presented with both text and image ( $R_{ti}$  items). All effort items were only described with text ( $E_t$ ). For each task, fMRI sessions were preceded by a short training session

## Appendices

(not included in the analysis), for participants to familiarize with the sort of items they would have to evaluate and with the button pad they would use to express their preferences. As trial timings were self-paced, we added a small incentive to push subjects to not take too much time in answering for each trial in each task. If the subjects had not provided any answer after a certain period of time (10 seconds for the rating task, 5 seconds for the choice tasks), a message with the words “Too slow!” appeared on screen and then the current trial restarted.

Participants all started with a (dis-)likeability rating task (Fig. 1A). Each of the three fMRI sessions included three blocks of 24  $R_{ii}$ , 24  $R_t$  and 24  $E_t$  trials, the order of blocks being counterbalanced across participants. The items were presented one by one, and participants rated them by moving a cursor along an analog scale. They used their index and middle fingers to press buttons corresponding to left and right movements, and validated the final position of the cursor by pressing a third button, which triggered the new trial. The initial position of the cursor, at the beginning of each trial, was randomly placed between 25 and 75% of the 0-100 rating scale. There was no mark on the scale, giving the impression of a continuous rating, although it was discretized into 100 steps. The left and right extremes of the scale were labeled “I would not care” and “I would like it enormously” for reward items, “I would not mind” and “I would dislike it enormously” for effort items. In any case, the situations to be rated were hypothetical: the question was about how much they would like the reward (should it be given to them at the end of the experiment) and how much they would dislike the effort (should it be imposed to them at the end of the experiment).

After the three rating sessions, participants performed a series of binary choices. The 1D-2O left/right choice task (Fig. 1B) involved expressing a preference between two options of a same dimension, presented on the left and right of the screen. The two options were items presented in the rating task, drawn from the same category, regarding both the presentation mode ( $R_{ii}$  vs  $R_{ii}$ ,  $R_t$  vs  $R_t$ ,  $E_t$  vs  $E_t$ ) and type of items (food vs. food, goods vs. goods, mental vs mental, physical vs physical). Each item was presented twice, following two inter-mixed pairing schedules: one varied the mean rating (i.e., stimulus value) while controlling for distance (i.e., decision value or choice difficulty), whereas the other varied the distance in rating while controlling the mean. Participants selected the reward they would most like to obtain, or the effort they would least dislike to exert, by pressing the left or right button with their middle or index finger. The chosen option was then highlighted with a red frame, so participants could check that their choice was correctly recorded. The three sessions of the 1D-2O choice task included each three 24-trial blocks presenting the three types of options ( $R_{ii}$ ,  $R_t$ ,  $E_t$ ), the order



## Appendices

of blocks being counterbalanced across participants. If the subjects had still not provided any answer after 5 seconds, a message with the words “Too slow!” appeared on screen and then the current trial restarted.

Then participants performed the 1O-2D yes/no choice task (Fig. 1C), which involved deciding whether to accept exerting a given effort in order to get a given reward. Thus, every trial proposed one option combining two dimensions (one  $R_t$  and  $E_t$ ). Each item was presented twice, following two inter-mixed pairing schedules: one associating more pleasant reward with more painful effort (thus controlling for decision value or choice difficulty), the other associating more pleasant reward with less painful effort (this varying choice difficulty). Participants selected their response by pressing the button corresponding to ‘yes’ or ‘no’ with their index or middle finger. The left/right position of yes/no responses was counterbalanced across trials. To give participants a feedback on their choice, the selected option was highlighted with a red frame. The three sessions of the 1O-2D choice task contained 48 trials each.

### Behavioral data analysis

All data were analyzed using Matlab 2017a (The MathWorks, Inc., USA)

Choices were fitted with logistic regression models with intercept and decision value weighted by a free parameter.

For 1D-2O choices, the model was:

$$P(\text{choice} = \text{left}) = \frac{1}{1 + e^{\beta_0 + \beta_1 \cdot (\Delta V)}}$$

where  $\Delta V$  is the decision value, i.e. the difference in rating between left and right options.

For 1O-2D choices, the model was:

$$P(\text{choice} = \text{accept}) = \frac{1}{1 + e^{\beta_0 + \beta_R \cdot V_R + \beta_E \cdot V_E}}$$

where  $V_R$  and  $V_E$  are the ratings provided for the reward and effort items. Thus, the decision value (or net value) here is a weighted sum reward likeability and effort dislikeability, with an additional parameter scales the two dimensions.

Deliberation time (DT) was defined across tasks as the time between stimulus onset and first button press. Trial-wise variations in DT were fitted with linear regressions models, with a session-specific intercept, nuisance factors (fixation cross display duration (Jitter), stimulus

## Appendices

luminance (Lum), and text length in number of words (Length)) and factors of interest (stimulus value (Val) and response confidence (Conf), see below). Thus, the model was:

$$DT = \beta_{s1} + \beta_{s2} + \beta_{s3} + \beta_{jit} \cdot Jitter + \beta_{lum} \cdot Luminance + \beta_{len} \cdot Length + \beta_{val} \cdot Val + \beta_{conf} \cdot Conf$$

The Val and Conf regressors represented stimulus value (reward likeability minus effort dislikeability) and response confidence (squared distance from mean response). They were adapted to the task, as follows:

	Rating task	1D-2O choice task	1O-2D choice task
Val	V	$V_{left} + V_{right}$	$\beta_R \cdot V_R + \beta_E \cdot V_E$
Conf	$[V - \text{mean}(V)]^2$	$[P_{left} - \text{mean}(P_{left})]^2$	$[P_{accept} - \text{mean}(P_{accept})]^2$

In each case, P is probability generated with the logistic regression and V is either reward likeability or effort dislikeability, provided by z-scored individual ratings.

### fMRI data acquisition

Functional and structural brain imaging data was collected using a Siemens Magnetom Prisma 3-T scanner equipped with a Siemens 64 channel Head/Neck coil. Structural T1-weighted images were coregistered to the mean echo planar image (EPI), segmented and normalized to the standard T1 template and then averaged across subjects for anatomical localization of group-level functional activation. Functional T2\*-weighted EPIs were acquired with BOLD contrast using the following parameters: repetition time TR = 2.01 seconds, echo time TE = 25ms, flip angle = 78°, number of slices = 37, slice thickness = 2.5mm, field of view = 200mm. A tilted-plane acquisition sequence was used to optimize sensitivity to BOLD signal in the orbitofrontal cortex (44).

Most subjects (n=25) performed 9 sessions (3 per task) using this standard EPI sequence, but a subgroup (n=15) performed only 3 sessions (1 per task). In this subgroup, 3 sessions were scanned using a sequence multi-band acceleration factor, and 3 sessions using multi-band + multi-echo acquisition. However, functional data collected with these two other sequences are not analyzed in this paper. Note that the number of volumes per session was not predefined, because all responses were self-paced. Volume acquisition was just stopped when the task was completed.

### **fMRI data analysis**

Functional MRI data were preprocessed and analyzed with the SPM12 toolbox (Wellcome Trust Center for NeuroImaging, London, UK) running in Matlab 2017a. Preprocessing consisted of spatial realignment, normalization using the same transformation as anatomical images, and spatial smoothing using a Gaussian kernel with a full width at a half-maximum of 8 mm.

Preprocessed data were analyzed with a standard general linear model (GLM) approach at the first (individual) level and then tested for significance at the second (group) level. All GLM included the six movement regressors generated during realignment of successive scans. In our main GLM, stimulus onset was modeled by a stick function, modulated by the following regressors: 1) fixation cross duration, 2) luminance, 3) text length, 4) Val, 5) Conf, 6) DT. The first three were nuisance factors that were found to significantly impact DT in the linear regression described above. The regressors of interest were defined as explained in the behavioral data analysis section. All regressors of interest were z-scored and convolved with the canonical hemodynamic response function and its first temporal derivative. All parametric modulators were serially orthogonalized. At the second level, correlates of Val, Conf and DT were obtained with contrasts across tasks of corresponding regression estimates against zero. Note that dislikeability ratings obtained for effort items were negatively weighted in all regressors (meaning that they can only decrease stimulus value).

Regions of interest (ROI) were defined as clusters in group-level statistical maps that survived significance threshold of  $p < 0.05$  after family-wise error correction for multiple comparisons at the voxel level. Parameter estimates and t-values were extracted from each voxel within these clusters and then averaged across voxels. Finally, a last GLM was built with one event per trial, modeled with a stick function, at the time of stimulus onset. It was used to extract trial-by-trial activity levels in the clusters of interest, which then served as regressors to explain pupil size data (see below).

### **Meta-analysis of fMRI studies**

The meta-analytic maps were extracted from the online platform Neurosynth (<https://www.neurosynth.org/>), using the keywords “value” (470 studies), “confidence” (79 studies) and “effort” (204 studies) for “uniformity test”, which displays brain regions that are

## Appendices

consistently activated in paper mentioning the keyword. Each map was binarized to visualize clusters surviving a significance threshold of  $p < 0.01$  after false discovery rate (FDR) correction for multiple comparisons.

### **Pupil size**

Pupil diameter was recorded at a sampling rate of 1000Hz, using an EyeLink 1000 plus (SR Research) eye-tracker, after calibration before fMRI sessions, once the subject was positioned inside the scanner. A cubic interpolation was performed to compensate for any period of time when the pupil signal was lost due to blinking or any other reason. The pupil size time series were subsequently band-pass filtered (1/128 to 1Hz) and zscored per session.

Within-trial variations in pupil size was baseline-corrected (by removing the mean signal over the 200 ms preceding stimulus onset) and time-locked either to stimulus onset or button press. Then trial-wise variations in pupil size were then fitted with a linear regression model that included nuisance factors (an intercept per block, jitter duration, stimulus luminance and text length), variables of interest (Val, Conf and DT defined as in behavioral data analysis) and neural activity (extracted from vmPFC, mPFC and dmPFC ROI clusters). Within-trial individual time series of regression estimates were then smoothed using a 100ms kernel. Group-level significant time clusters were identified after correction for multiple comparisons estimated according to random field theory, using the RFT\_GLM\_contrast.m function of the VBA\_toolbox (available at <http://mbb-team.github.io/VBA-toolbox/>).

## Bibliography

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., and Pessiglione, M. (2015). Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. *Journal of Neuroscience* 35, 2308–2320.

Alnaes, D., Sneve, M.H., Espeseth, T., Endestad, T., van de Pavert, S.H.P., and Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision* 14, 1–1.

Arulpragasam, A.R., Cooper, J.A., Nuutinen, M.R., and Treadway, M.T. (2018). Corticoinsular circuits encode subjective value expectation and violation for effortful goal-directed behavior. *Proceedings of the National Academy of Sciences* 115, E5233–E5242.

Bartra, O., McGuire, J.T., and 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.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., and Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychol Rev* 108, 624–652.

Boureau, Y.-L., and Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36, 74–97.

Chib, V.S., Rangel, A., Shimojo, S., and O’Doherty, J.P. (2009). Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *Journal of Neuroscience* 29, 12315–12320.

Chua, E.F., Schacter, D.L., Rand-Giovannetti, E., and Sperling, R.A. (2006). Understanding metamemory: neural correlates of the cognitive process and subjective level of confidence in recognition memory. *Neuroimage* 29, 1150–1160.

De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Confidence in value-based choice. *Nature Neuroscience* 16, 105–110.

De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., and Love, B.C. (2017). Social Information Is Integrated into Value and Confidence Judgments According to Its Reliability. *The Journal of Neuroscience* 37, 6066–6074.

Deichmann, R., Gottfried, J.A., Hutton, C., and Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19, 430–441.

Gherman, S., and Philiastides, M.G. (2018). Human VMPFC encodes early signatures of confidence in perceptual decisions. *ELife* 7, e38293.

Hare, T.A., Camerer, C.F., and Rangel, A. (2009). Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. 324, 4.

## Appendices

Hare, T.A., Malmaud, J., and Rangel, A. (2011). Focusing Attention on the Health Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary Choice. *Journal of Neuroscience* *31*, 11077–11087.

Hogan, P.S., Galaro, J.K., and Chib, V.S. (2019). Roles of Ventromedial Prefrontal Cortex and Anterior Cingulate in Subjective Valuation of Prospective Effort. *Cerebral Cortex* *29*, 4277–4290.

Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F.S., and Behrens, T.E.J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience* *15*, 470–476.

Kable, J.W., and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* *10*, 1625–1633.

Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. *Science* *154*, 1583–1585.

Kamiński, J., Sullivan, S., Chung, J.M., Ross, I.B., Mamelak, A.N., and Rutishauser, U. (2017). Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nature Neuroscience* *20*, 590–601.

Kiani, R., Corthell, L., and Shadlen, M.N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron* *84*, 1329–1342.

Kolling, N., Wittmann, M.K., Behrens, T.E.J., Boorman, E.D., Mars, R.B., and Rushworth, M.F.S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nature Neuroscience* *19*, 1280–1285.

Kurniawan, I.T., Guitart-Masip, M., Dayan, P., and Dolan, R.J. (2013). Effort and Valuation in the Brain: The Effects of Anticipation and Execution. *Journal of Neuroscience* *33*, 6160–6169.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. *Neuron* *64*, 431–439.

Lebreton, M., Bertoux, M., Boutet, C., Lehericy, S., Dubois, B., Fossati, P., and Pessiglione, M. (2013). A Critical Role for the Hippocampus in the Valuation of Imagined Outcomes. *PLoS Biology* *11*, e1001684.

Lebreton, M., Abitbol, R., Daunizeau, J., and Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience* *18*, 1159–1167.

Lee, D., and Daunizeau, J. (2019). Trading Mental Effort for Confidence: The Metacognitive Control of Value-Based Decision-Making. *BioRxiv* 837054.

Levy, D.J., and Glimcher, P.W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology* *22*, 1027–1038.

Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., Navarro, V., Rheims, S., Kahane, P., Domenech, P., et al. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nat Neurosci* 1–12.

## Appendices

Oudiette, D., Vinckier, F., Bioud, E., and Pessiglione, M. (2019). A Pavlovian account for paradoxical effects of motivation on controlling response vigour. *Scientific Reports* 9.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359.

Pessiglione, M., and Delgado, M.R. (2015). The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. *Curr Opin Behav Sci* 5, 78–84.

Rangel, A., and Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology* 20, 262–270.

Schneider, B., and Koenigs, M. (2017). Human lesion studies of ventromedial prefrontal cortex. *Neuropsychologia* 107, 84–93.

Seaman, K.L., Brooks, N., Karrer, T.M., Castellon, J.J., Perkins, S.F., Dang, L.C., Hsu, M., Zald, D.H., and Samanez-Larkin, G.R. (2018). Subjective value representations during effort, probability and time discounting across adulthood. *Social Cognitive and Affective Neuroscience* 13, 449–459.

Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A. (2019). Movement Vigor as a Reflection of Subjective Economic Utility. *Trends in Neurosciences* 42, 323–336.

Shenhav, A., Botvinick, M.M., and Cohen, J.D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron* 79, 217–240.

Skvortsova, V., Palminteri, S., and Pessiglione, M. (2014). Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates. *Journal of Neuroscience* 34, 15621–15630.

Steverson, K., Chung, H.-K., Zimmermann, J., Louie, K., and Glimcher, P. (2019). Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. *Sci Rep* 9, 20053.

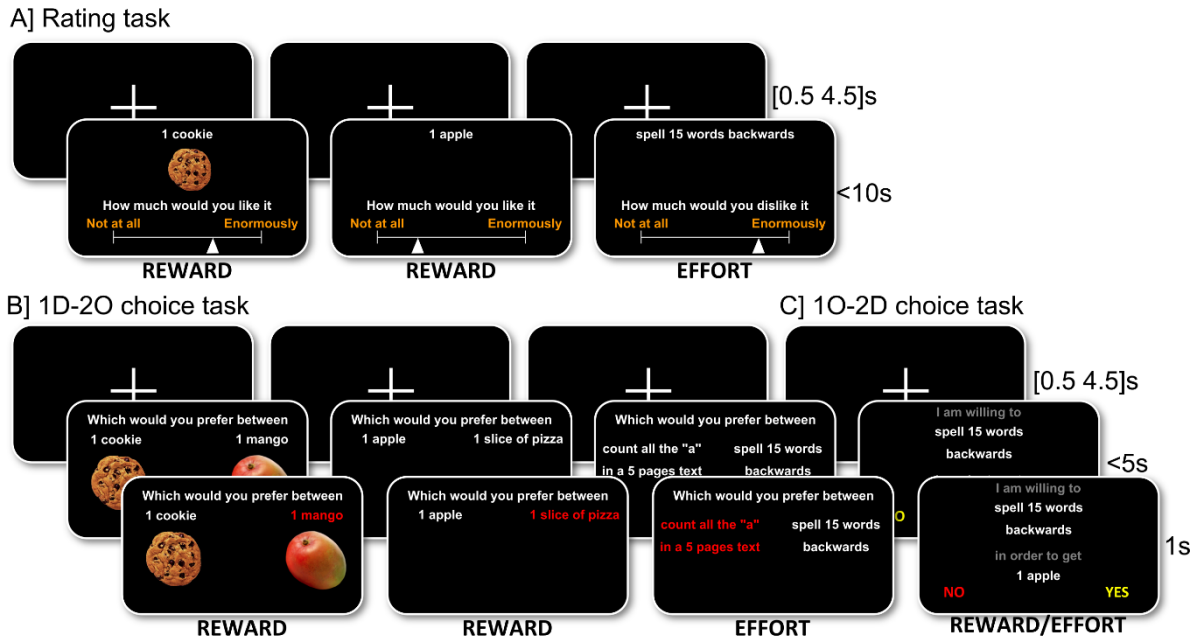
Talmi, D., Dayan, P., Kiebel, S.J., Frith, C.D., and Dolan, R.J. (2009). How humans integrate the prospects of pain and reward during choice. *J. Neurosci.* 29, 14617–14626.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science* 315, 515–518.

Weiskopf, N., Hutton, C., Josephs, O., Turner, R., and Deichmann, R. (2007). Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. *Magnetic Resonance Materials in Physics, Biology and Medicine* 20, 39–49.

van der Wel, P., and van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review* 25, 2005–2015.

Zénon, A., Sidibé, M., and Olivier, E. (2014). Pupil size variations correlate with physical effort perception. *Frontiers in Behavioral Neuroscience* 8.



**Figure 1. Behavioral tasks.**

Example trials are illustrated as a succession of screenshots from top to bottom, with durations in seconds. Only the fixation cross display at the beginning of trials is jittered. The duration of the response screen depends on deliberation time, as both rating and choice are self-paced.

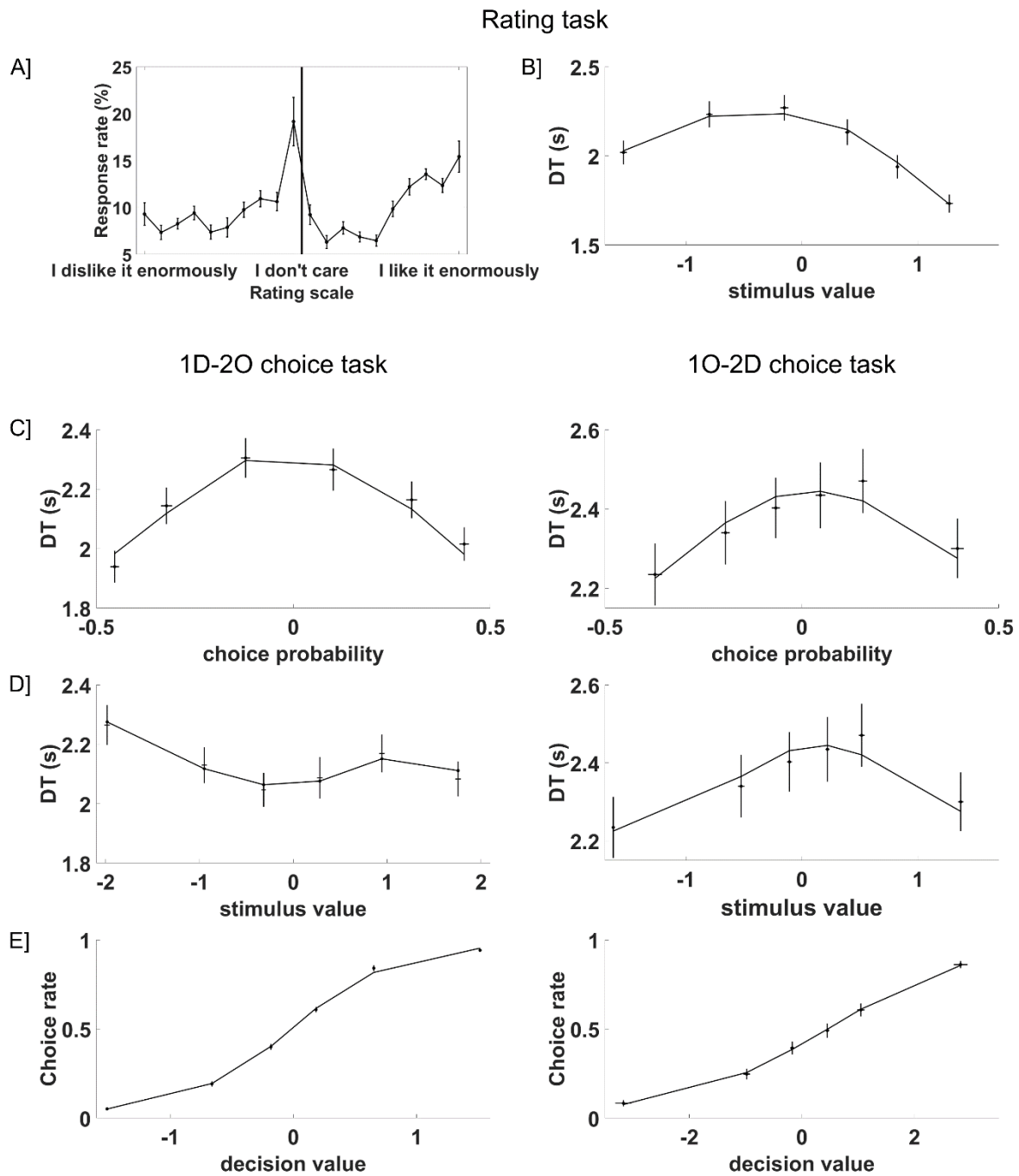
A) Rating task. In every trial, subjects are shown an item that can be a reward described with both text and image ( $R_{ti}$ ), a reward described with text only ( $R_t$ ) or an effort described with text only ( $E_t$ ). The task for subjects is to rate how much they would like receiving the proposed reward or dislike performing the proposed effort, should it occur, hypothetically, at the end of the experiment. They first move the cursor using left and right buttons on a pad to the position that best reflect their (dis)-likeability estimate, then validate their response with a third button and proceed to the next trial.

B) 1D-2O choice task. In every trial, two options belonging to the same category are shown on screen and subjects are asked to select their favorite option, i.e. which reward they would prefer to receive if they were offered the two options (hypothetically) or the effort they would prefer to exert if they were forced to implement one of the two options (hypothetically). The choice is expressed by selecting between left and right buttons with the index or middle finger. The chosen option is then highlighted in red, and subjects proceed to the next trial.



## Appendices

C] 1O-2D choice task. In every trial, one option combining the two dimensions is shown on screen and subjects are asked to state whether they would be willing to exert the effort in order to receive the reward, if they were given the opportunity at the end of the experiment (hypothetically). They select their response ('yes' or 'no', positions counterbalanced across trials) by pressing the left or right button, with their index or middle finger.



**Figure 2: Behavioral results.**

A] Distribution of ratings. Bars show the average response rate for each bin of ratings. Effort items (on the left) are rated between bin 0 ('I would not mind') and bin 10 ('I would dislike it enormously'). Reward items (on the right) are rated between bin 0 ('I would not care') and bin 10 ('I would like it enormously').

B] Deliberation time in the rating task, as a function of the centered item value (likeability rating). Positive value means higher likeability (for rewards) or lower dislikeability (for efforts), 0 is the mean rating across trials.

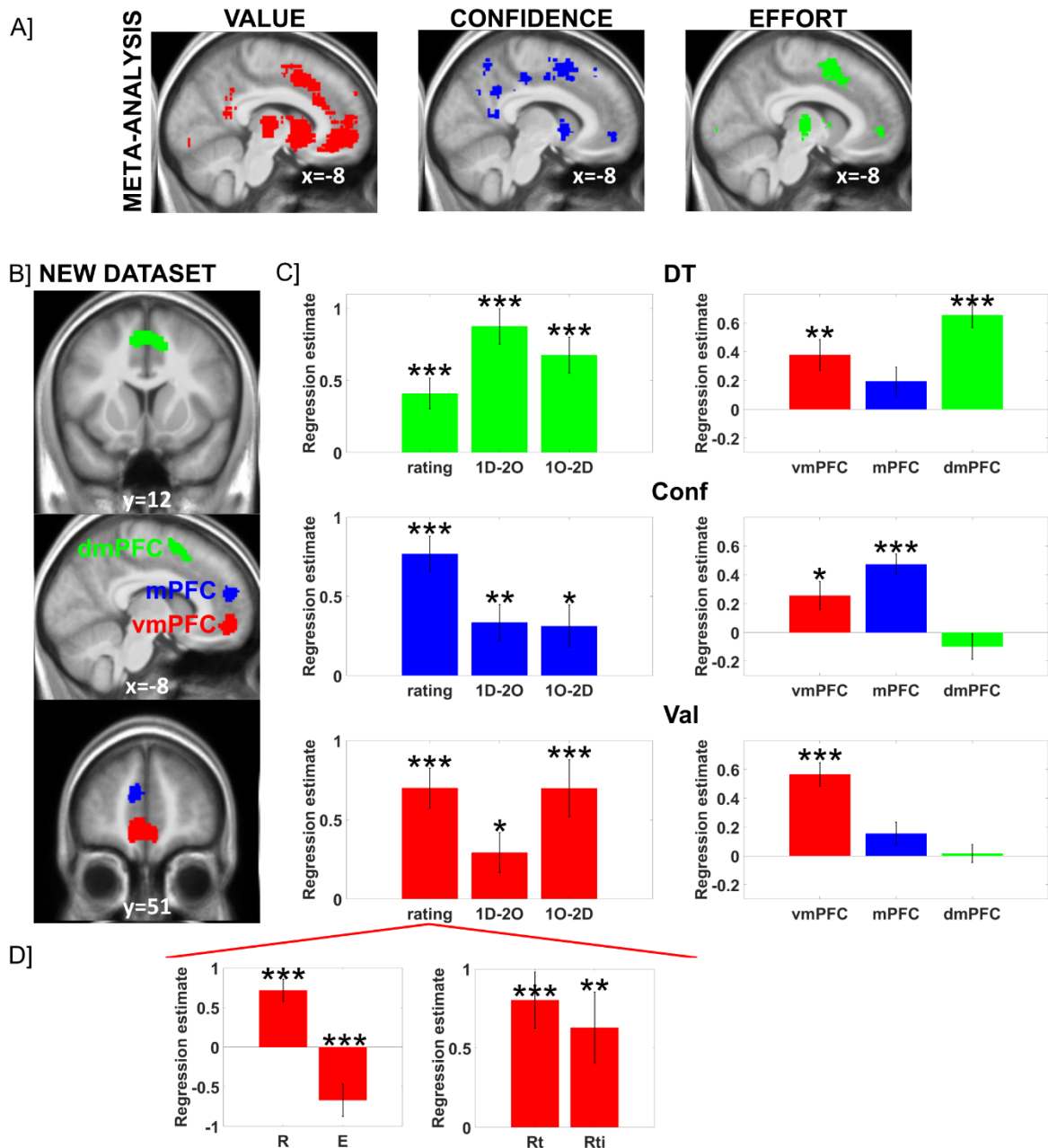
## Appendices

C] Deliberation time as a function of stimulus value. In the 1D-2O choice task (left graph), stimulus value is the sum of likeability ratings for left and right options ( $V_{\text{left}} + V_{\text{right}}$ ). In the 1O-2D choice task (right graph), stimulus value is the likeability of reward plus the dislikeability of effort. Thus, in this yes/no choice task, stimulus value is equivalent to decision value ( $\beta_R \cdot V_R + \beta_E \cdot V_E$ ). In any case, stimulus value was centered, such that 0 is the mean across trials.

D] Deliberation time as a function of choice probability. In the 1D-2O choice task (left graph), choice probability is the output of the softmax function for the left option, centered such that 0 is the mean across trials. In the 1D-2O choice task (right graph), choice probability is the output of the softmax function for the yes option, centered such that 0 is the mean across trials.

E] Choice rate as a function of decision value. Decision value is simply the difference between left and right option ratings in the 1D-2O choice task, and the weighted sum of reward and effort ratings in the 1O-2D choice task. The set of decision values across trials were distributed over 6 bins. Dots are mean choice rate ('left' response in the 1D-2O and 'yes' response in the 1O-2D task). Lines show binned logistic regression fits.

Dots represent mean across participants, error bars are inter-participant standard errors.



**Figure 3: Neural results.**

A] Meta-analysis of fMRI studies. Statistical maps (sagittal slices) were extracted from the Neurosynth platform with the ‘value’, ‘confidence’ and ‘effort’ keywords. Significant clusters in the medial prefrontal cortex are similar across keywords.

B] Neural correlates of value, confidence and deliberation constructs in the present dataset. Statistical maps were obtained with a GLM including the different variables as parametric modulators of stimulus onset, across rating and choice tasks. Sagittal slice was taken at the same coordinates as the Neurosynth output, and superimposed on the average anatomical scan

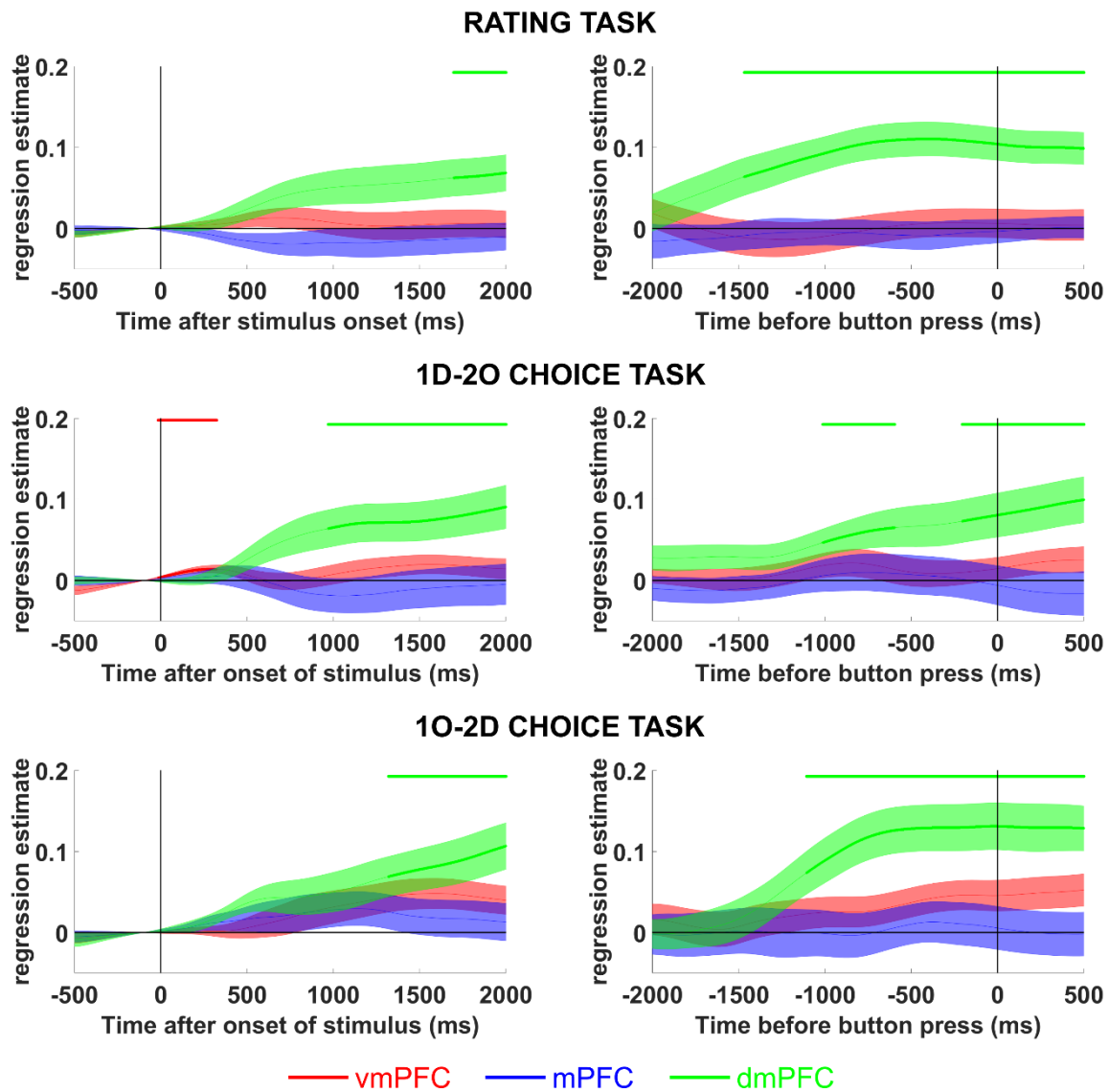
## Appendices

normalized to canonical (MNI) template. Coronal slices show the extent of the different medial prefrontal clusters. Statistical threshold was set at  $p < 0.05$  after family-wise error for multiple comparisons at the voxel level. For clusters outside the medial prefrontal cortex, surviving a more tolerant statistical threshold, see activations in Table 1

C] Decomposition of regression estimates obtained for each variable of interest, per task (rating, 1D-2O and 1O-2D choice) on the left, and per ROI (vmPFC, mPFC, dmPFC) on the right.

D] Decomposition of regression estimates, obtained for Val in the vmPFC during rating, per stimulus category (reward R versus effort E, and reward presented as text + image  $R_{ti}$  versus text only  $R_t$ ).

Bars show mean across participants, error bars show inter-participant standard errors. Stars indicate significance of t-test against zero (\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ ).



**Figure 4: Pupillometric results.**

Plots show the time course of regression estimates, obtained with a GLM built to explain pupil size. The GLM included nuisance regressors (jitter duration, stimulus luminance, text length), variables of interest (Val, Conf, DT) and activities in main ROI (vmPFC, mPFC, dmPFC, corresponding to red, blue and green traces, respectively). Each row corresponds to a different task (likeability rating, 1D-2O and 1O-2D choice tasks). Left and right columns show time courses aligned on stimulus onset and button press, respectively. Lines represent means across

## Appendices

participants and shaded areas inter-participant standard errors. Horizontal bars indicate significant time clusters after correction for multiple comparisons.

**Table 1: Brain activity signaling Val across rating and choice tasks.**

Regions survived a significance threshold of  $P < 0.05$  after FWE correction for multiple comparisons the voxel level. Clusters smaller than 12 voxels, corresponding to the size of our smoothing kernel, were excluded from the table. Coordinates refer to the MNI space. Region labels were extracted with the AAL atlas. The p.value reported is the p.value of the cluster after a FWE correction at the cluster level.

Region	P cluster	Peak x	Peak y	Peak z	No. of Voxels
Superior Frontal Gyrus, Medial Orbital, left	$3 \cdot 10^{-10}$	-10	48	-12	364
Middle Frontal Gyrus, Orbital, left	$2 \cdot 10^{-4}$	-28	36	-14	57
Lingual Gyrus, right	$1 \cdot 10^{-4}$	16	-70	-6	64
Cingulate Gyrus, Anterior part, left	0.005	-8	38	6	16
Precuneus, left	0.003	-6	-54	14	22



**Table 2: Brain activity signaling Conf across rating and choice tasks.**

Regions survived a significance threshold of  $P < 0.05$  after FWE correction for multiple comparisons the voxel level. Clusters smaller than 12 voxels, corresponding to the size of our smoothing kernel, were excluded from the table. Coordinates refer to the MNI space. Region labels were extracted with the AAL atlas. The p.value reported is the p.value of the cluster after a FWE correction at the cluster level.

Region	P FWE cluster	Peak x	Peak y	Peak z	No. of Voxels
Superior Frontal Gyrus, Medial left	$5 \cdot 10^{-6}$	-8	52	18	128
Temporal Inferior pole left	0.007	-46	2	-36	13
Supramarginal Gyrus left	$3 \cdot 10^{-4}$	-62	-40	32	56
Temporal Mid left	$2 \cdot 10^{-4}$	-56	-26	-10	63
Temporal Mid pole left	0.003	-46	-64	12	22
Caudate Nucleus left	0.006	-12	14	-12	14

**Table 3: Brain activity signaling DT across rating and choice tasks.**

Regions survived a significance threshold of  $P < 0.05$  after FWE correction for multiple comparisons the voxel level. Clusters smaller than 12 voxels, corresponding to the size of our smoothing kernel, were excluded from the table. Coordinates refer to the MNI space. Region labels were extracted with the AAL atlas. The p.value reported is the p.value of the cluster after a FWE correction at the cluster level.

Region	P FWE cluster	x	y	z	No. of Voxels
Insula left	$2 \cdot 10^{-5}$	-30	26	4	110
Supplementary Motor Area right	$1 \cdot 10^{-9}$	10	12	48	365
Insula right	$4 \cdot 10^{-5}$	32	26	4	95
Inferior Frontal Gyrus, Triangular left	$7 \cdot 10^{-8}$	-40	22	24	242
Lingual Gyrus left	0.003	-18	-88	-10	23

## II Enregistrement de l'activité du vmPFC à l'aide de différentes séquences d'IRMf

### II.1 Introduction

Le cortex préfrontal ventromédian, du fait de sa localisation proche des orbites, est soumis à des artefacts au moment de l'acquisition des données d'imagerie fonctionnelle (Deichmann et al., 2003; Domsch et al., 2013; Weiskopf et al., 2007). Nous avons souhaité voir si le fait d'utiliser des séquences légèrement différentes aux séquences classiquement utilisées en imagerie par résonance magnétique fonctionnelle (IRMf) allait changer les résultats observés au niveau du cortex préfrontal ventral en lien avec la valeur. Nous espérons notamment ainsi réduire le bruit inhérent aux méthodes classiques d'acquisition et éventuellement ainsi mieux capter la sensibilité du vmPFC pour la valeur subjective. Nous avons acquis les données au cours de 3 sessions séparées pour chacune de nos tâches. Les participants de la deuxième cohorte ont effectué toutes les sessions avec la même séquence, la séquence A (voir *IRM*). Cependant, grâce à ce design expérimental, nous avons pu acquérir les données, pour chacune de nos tâches et chacun de nos participants, avec une séquence différente pour chacune des sessions effectuées dans le scanner. L'expérience initiale visait à établir des corrélats stables de la valeur subjective, de la confiance subjective et du temps de délibérations au travers de 3 tâches différentes : une tâche d'évaluation, une tâche de choix entre 2 options au sein de la même dimension (tâche de choix 1D-2O) et une tâche de choix d'accepter ou de rejeter une offre nécessitant d'intégrer deux dimensions : les bénéfices d'une récompense et les coûts d'une action (tâche de choix 1O-2D).

### II.2 Méthodes

Pour la séquence A et B, nous avons traité les données comme détaillé dans la première partie (voir *IRM*). Le GLM utilisé était aussi complètement identique. La seule différence entre les données détaillées dans la première étude et l'analyse présente concerne le fait que l'acquisition de données avec plusieurs temps d'échos de la séquence C qui requiert de regrouper ensemble les données issues de chaque temps d'échos en une seule image. C'est d'ailleurs tout l'intérêt de cette technique en vue d'améliorer la qualité des données via le moyennage de plusieurs groupes de données acquis parallèlement.

Dans le cas de la séquence C, nous avons acquis deux séries d'images en parallèle : une pour chaque temps d'écho. Pour chaque série d'images acquise avec chaque temps d'écho, nous

## Appendices

avons d'abord effectué les étapes suivantes indépendamment pour chaque temps d'écho : 1) Nous avons réaligné les images fonctionnelles, 2) nous les avons co-enregistrées avec l'image anatomique de chaque sujet et 3) nous les avons normalisées dans l'espace MNI (*Montreal Neurological Institute*) et lissées avec un facteur de 8mm. Ensuite, nous avons calculé un ratio contraste/bruit (*contrast-to-noise ratio*, CNR) pour chaque série d'images selon la formule :

$$CNR_e = \frac{\mu}{\sigma} * TE_e$$

où  $\mu$  représente la moyenne et  $\sigma$  la déviation standard de toutes les images pour une série d'image d'un temps d'écho et d'une session d'acquisition donnés,  $TE$  représente le temps d'écho de chaque série d'images ( $TE_0 = 11.00\text{ms}$  et  $TE_1 = 29.89\text{ms}$ ) et  $e$  le temps d'écho utilisé. Enfin, nous avons regroupé les images acquises pour chaque temps d'échos dans chaque session selon la formule :

$$S_i = \frac{S_{i,E_0} \cdot CNR_{E_0} + S_{i,E_1} \cdot CNR_{E_1}}{CNR_{E_0} + CNR_{E_1}}$$

où  $S_i$  est la coupe d'index  $i$  et  $S_{i,E_e}$  la coupe d'index  $i$  pour l'écho d'index  $e$ . Cette manière de pondérer les images issues de temps d'écho différents a été établie à la suite d'autres études ayant employé des séquences similaires (Halai et al., 2014; Poser et al., 2006).

Une fois les images regroupées ainsi, nous avons pu analyser les données avec le même GLM que celui utilisé dans l'analyse de la première étude avec la séquence A (voir *Analyse des données d'IRMf : les modèles utilisés*).

Nous avons utilisé une région d'intérêt au niveau de l'aire 14m définie par Scott Mackey et Michael Petrides (Mackey and Petrides, 2014). Nous avons inclus tous les voxels dont la probabilité était supérieure à 75% d'être dans cette zone d'après les auteurs et nous avons extraits les coefficients de régression correspondants au régresseur Val dans cette zone pour chaque séquence.

### II.3 Résultats

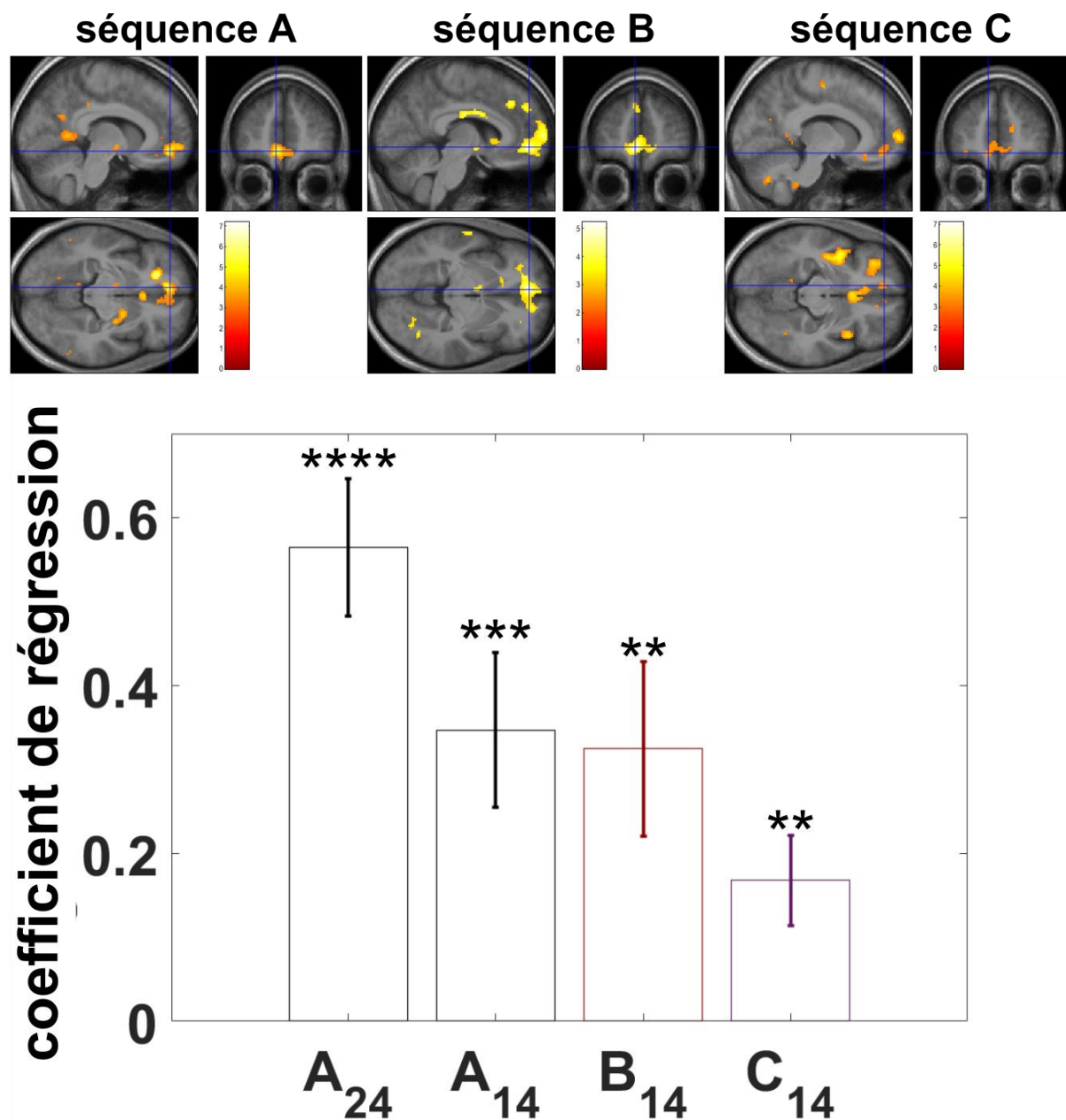
Du fait de sa localisation, le cortex orbitofrontal ventral est soumis à plusieurs artéfacts en IRM fonctionnelle (Deichmann et al., 2003; Domsch et al., 2013; Weiskopf et al., 2007). En partenariat avec le Centre de Neuro-Imagerie de Recherche ([CENIR](#), Institut du Cerveau et de la Moelle Epinière, Hôpital de la Pitié Salpêtrière, Paris 13<sup>ème</sup>, France), nous avons pu tester trois séquences d'imagerie différentes (voir *Acquisition des données d'IRM anatomique et fonctionnelle*) afin de voir si nous pouvions ainsi améliorer la sensibilité de l'IRMf pour nos

## Appendices

zones d'intérêt, en particulier au sein du vmPFC. La séquence B comportait un facteur d'accélération multi-bandes qui permet, entre autres, d'améliorer la fréquence d'échantillonnage des données, c'est-à-dire de réduire le temps de répétition (TR) tout en gardant une bonne résolution spatiale. La séquence C correspondait à l'acquisition avec un facteur d'accélération multi-bandes de deux séries de données en parallèle grâce à l'utilisation de deux temps d'échos (TE) différents (11.00 ms et 29.89 ms). Nous avons ensuite regroupé ces deux séries d'images en une seule série afin de pouvoir les analyser comme les autres séquences (voir *IRM* pour le détail). La première cohorte de participants (N = 14) a ainsi effectué 3 sessions de chaque tâche avec une séquence différente pour chaque session afin de pouvoir comparer les résultats.

A cause du nombre plus faible de participants (n = 14 dans la première cohorte) et du nombre plus faible de sessions (1 seule session par séquence) en comparaison aux résultats regroupant les deux cohortes avec la séquence A, on constate que, lorsque l'on observe les cartes de corrélation avec la valeur en cerveau entier, aucun cluster, dans aucune des cartes, ne passe les seuils après correction pour les comparaisons multiples au niveau des voxels à p.value < 0.001 non-corrigée au niveau des voxels. Malgré tout, il est possible d'identifier un cluster dans le vmPFC pour chacune des trois séquences utilisées (voir **Figure 42**) à des coordonnées relativement proches entre les trois séquences qui plus est. En effet, on peut identifier un pic d'activation pour les séquences A, B et C situé, respectivement aux coordonnées MNI (-8 ; 48 ; -10), (-8 ; 50 ; -6) et (-10 ; 48 ; -12). Malgré les distorsions différentielles qui peuvent être induites par l'utilisation de différentes séquences d'acquisition, le résultat paraît ainsi relativement stable.

De plus, même en extrayant les données sur la base du cluster identifié intégralement à l'aide de la séquence A sur les 38 participants de la première étude, on observe que l'activité du vmPFC est relativement similaire au travers des 3 séquences employées (voir **Figure 42**). Pour éviter d'effectuer du « double deeping » puisque le cluster dans le vmPFC a été établi sur la base de sa corrélation avec la valeur en regroupant toutes les données basées sur la séquence A, nous avons aussi extrait le même contraste en utilisant une aire définie anatomiquement, l'aire 14m, qui semble correspondre au cluster régulièrement identifié en lien avec la valeur subjective. Nous avons pu voir que les résultats pour chaque séquence étaient très similaires aussi dans ce cas (A<sub>14</sub> :  $\beta = 0.208 \pm 0.076$ , p = 0.017; B<sub>14</sub> :  $\beta = 0.192 \pm 0.080$ , p = 0.032 ; C<sub>14</sub> :  $\beta = 0.089 \pm 0.033$ , p = 0.019). De plus, aucune des comparaisons deux à deux des séquences n'était significative (A<sub>14</sub>>B<sub>14</sub> : p = 0.856, A<sub>14</sub>>C<sub>14</sub> : p = 0.212 ; B<sub>14</sub> > C<sub>14</sub> : p = 0.257).



**Figure 42 : Les différentes séquences d’acquisition en IRMf donnent des résultats équivalents pour l’activité du vmPFC en corrélation avec la valeur.** Les figures du haut représentent la carte d’activation, pour chaque séquence, correspondant au régresseur valeur moyenné à travers les 3 tâches. Les cartes sont seuillées à  $p.value < 0.005$  non-corrigée au niveau des voxels. Les cartes représentent donc les corrélats de la valeur avec un zoom sur un cluster dans le vmPFC pour, de gauche à droite, la séquence A ( $k = 395$  voxels, pic en  $(-8 ; 48 ; -10)$ ), la séquence B ( $k = 1148$  voxels, pic en  $(-6 ; 50 ; -6)$ ) et la séquence C ( $k = 403$  voxels, pic en  $(-6 ; 64 ; 4)$ ), mais la carte est zoomée sur un autre pic du cluster aux coordonnées  $(-10 ; 48 ; -12)$ ). Les cartes d’activation sont superposées à la moyenne des scans anatomiques des 14 cerveaux des participants de la première cohorte après normalisation dans l’espace MNI. La barre colorée accompagnant chaque carte montre la correspondance entre les couleurs de la carte et les valeurs T pour chaque voxel. La figure du bas représente la valeur des corrélations entre le régresseur Val et l’activité du vmPFC pour chaque séquence. Le vmPFC utilisé pour cette extraction a été extrait sur la base de la première étude qui regroupait les données des 38 participants ayant effectué la séquence A, les données ne sont ainsi montrées qu’à titre indicatif

pour ce qui est de la séquence A.  $A_{24}$  représente le résultat dans le cas des 24 participants ayant effectué l'ensemble des sessions avec la séquence A. Ces participants disposaient de 3 sessions au lieu d'une ce qui rend bien évidemment leurs résultats non-comparables aux autres mais ils sont montrés à titre indicatif.  $A_{14}$ ,  $B_{14}$  et  $C_{14}$  correspondent, respectivement, aux données des 14 participants pour la séquence A, la séquence B et la séquence C d'IRMf.

### **II.4 Discussion**

Nous pensons éventuellement améliorer la qualité des enregistrements dans le cortex préfrontal ventromédian (vmPFC) en lien avec la valeur subjective en utilisant différents protocoles d'imagerie voire peut-être découvrir des résultats inattendus à l'aide de ces techniques censées être moins sensibles au bruit lors de l'acquisition des données. Les résultats ont au contraire révélé une certaine stabilité que ce soit dans la valeur de la corrélation entre le vmPFC et la valeur subjective ou dans la localisation de ce cluster. Nous avons ainsi constaté que les séquences utilisées habituellement dans le domaine de la prise de décision qui sont relativement similaires à la séquence A sont déjà relativement optimales pour ce qui est de l'acquisition de données en lien avec la valeur dans le cortex préfrontal ventromédian. Rien ne nous dit pas cependant que l'évolution des techniques ne permettra pas d'améliorer encore la qualité des données à l'avenir. L'acquisition à l'aide de plusieurs temps d'échos, mais il faut en avoir un nombre conséquent, paraît notamment prometteuse pour pouvoir réduire le bruit entre autres.

## III Utiliser un modèle computationnel pour inférer les valeurs cachées associées au comportement

### III.1 Introduction

Il est généralement reconnu qu'au moment d'effectuer un choix, les individus reconstruisent une valeur subjective de manière dynamique (voir *Modèle de diffusion*). C'est là un des fondements des modèles de diffusion tels qu'ils ont été développés dans le champ perceptuel avant d'être ensuite appliqués au champ de la prise de décision basée sur la valeur subjective. Dans le cas des choix perceptuels, l'origine de l'information perceptuelle est liée à la perception, c'est-à-dire aux entrées sensorielles. Dans le cas des choix basés sur la valeurs subjective, ce type de modèle ne fait pas d'hypothèse sur l'origine de cette construction dynamique de la valeur subjective associée à chaque option évaluée en vue d'effectuer un choix (Padoa-Schioppa, 2011). L'origine de la valeur subjective qui dirige le comportement pourrait dépendre du contexte. Par exemple, il se pourrait que cette construction soit basée sur la mémoire et qu'elle implique juste une reconstruction progressive de la valeur subjective associée à un item depuis les structures cérébrales qui stockent le souvenir vers les structures qui se chargent d'opérer l'évaluation et le choix. Une autre hypothèse implique que ces valeurs soient produites par rapport aux informations externes. Par exemple, avant d'effectuer un choix entre différents aliments, il pourrait s'agir d'évaluer la composition en différents types de nutriments des différents items à disposition sur la base d'informations sensorielles liées à la vue ou à l'odorat (Suzuki et al., 2017).

Une approche originale aux tâches basées sur des préférences revient ainsi à considérer que les évaluations et les choix effectués par les individus correspondent à la transformation et au décodage d'une « vraie valeur » interne (Polanía et al., 2019). Ce décodage s'opère de manière dynamique au cours du choix et il s'accompagne de bruit. Nous avons essayé d'appliquer un modèle computationnel qui a été développé au sein de notre équipe aux données de la première étude qui regroupe les 38 participants. Ce modèle offre l'avantage de se baser sur toutes les tâches effectuées par les participants au sein de l'étude en vue d'inférer les « vraies valeurs » qui constitueraient les valeurs cachées qui dirigent le comportement et non pas seulement de la tâche d'évaluation comme nous l'avons fait dans notre étude. L'idée est que ce modèle permet ainsi de réduire le bruit autour de la valeur associée à chaque item puisqu'il est basé sur plusieurs points de données (chaque item apparaissait au cours de 2 choix dans la tâche de choix 1D-2O et de même dans la tâche de choix 1O-2D). Nous avons ainsi souhaité voir si



les résultats observés lors de la première étude étaient modifiés du fait d'utiliser ce modèle des valeurs cachées supposé mieux refléter la valeur interne associée à chaque item pour nos participants.

## III.2 Méthodes

### III.2.a Le modèle des valeurs cachées employé

Nicolas Borderies, Fabien Vinckier, Jean Daunizeau et Mathias Pessiglione ont développé un modèle dont l'idée principale est que chaque item présenté à l'écran dans nos tâches est associé à une valeur cachée qui peut différer de la note donnée par les participants au cours de la tâche d'évaluation. Pour identifier au mieux la valeur cachée associée à chaque item, le modèle prend ainsi en compte les notes attribuées lors de la tâche d'évaluation, mais aussi les choix effectués lors des tâches de choix 1D-2O et 1O-2D. Ce modèle a été opérationnalisé via la toolbox [VBA](#) (*Variational Bayesian Analysis*) développée par Jean Daunizeau dans l'environnement Matlab (Daunizeau et al., 2014). Le modèle a été inversé en utilisant l'approche variationnelle bayésienne avec une approximation de Laplace. Nous l'avons appliqué à nos propres données afin de voir s'il induisait des différences dans les corrélations en IRMf. Les détails du modèle utilisé se trouvent dans un article en préparation de Fabien Vinckier et Mathias Pessiglione, mais nous allons cependant en détailler les grandes lignes ci-dessous.

Dans le modèle, chaque item était associé à une valeur cachée. Cette valeur cachée était représentée par un paramètre à estimer par le modèle. Dans le modèle, ces valeurs cachées étaient ainsi les variables qui expliquent le comportement dans chacune des 3 tâches.

Dans la **tâche d'évaluation**, l'évaluation  $E(A)$  fournie par un participant donné pour un item A donné dépendait ainsi, dans le modèle, de la valeur cachée  $VC(A)$  en fonction de l'équation :

$$E(A) = \frac{1}{1 + e^{-\frac{(valence*VC(A)-5)}{\beta_E}} - \beta_0}$$

(Équation 13)

La variable *valence* est égale à 1 pour les récompenses et à -1 pour les efforts. Ce paramètre *valence* permettait de prendre en compte le fait que l'échelle était inversée en fonction du type de bloc : en allant de gauche à droite, de la récompense la moins plaisante vers la plus plaisante dans un cas et de l'effort le moins pénible à l'effort le plus pénible dans l'autre cas. La variable -5 permettait d'ancrer la valeur sur l'échelle en posant arbitrairement que 5 (ou -5 pour les

## Appendices

efforts) était le centre de l'échelle.  $\beta_E$  était un paramètre libre positif spécifique à la tâche d'évaluation qui permet d'ajuster quelle portion de l'échelle était utilisée et  $\beta_0$  était un paramètre libre exprimant un biais à donner une certaine note sur l'échelle indépendamment des items. Si  $\beta_E$  était élevé, cela indiquait que le participant avait tendance à n'utiliser que le centre de l'échelle. Si  $\beta_E$  était faible, cela indiquait au contraire que le participant avait tendance à se focaliser sur les extrémités de l'échelle. Le paramètre  $\beta_0$  permettait d'envisager une situation où le participant se focaliserait sur une portion de l'échelle (à droite par exemple) indépendamment des items qui lui ont été présentés.

Dans la **tâche de choix 1D-2O**, le modèle permettait de prédire que la probabilité  $p(\text{choix} = A)$  de choisir un item A en fonction des valeurs cachées  $VC(A)$  et  $VC(B)$  associées aux deux options présentées variait en fonction de la formule :

$$p(\text{choix} = A) = \frac{1}{1 + e^{-\frac{VC(A)-VC(B)}{\beta_{2O1D}} + \beta_m}}$$

(Équation 14)

$B_{2O1D}$  était un paramètre libre positif spécifique à la tâche 1D-2O représentant la stochasticité des choix, et  $\beta_m$  était un biais moteur. Le paramètre  $\beta_m$  permettait de prendre en compte le fait qu'un participant pouvait avoir une tendance à répondre plus souvent à droite (ou à gauche) indépendamment des options présentées et des valeurs qui leur étaient associées.

Dans la **tâche de choix 1O-2D**, le modèle a été établi afin que la probabilité d'accepter une offre donnée  $p(\text{choix} = \text{accepter})$  varie en fonction de la valeur cachée  $VC(R)$  associée à une récompense R donnée et de la valeur cachée  $VC(E)$  associée à un effort E donné selon la formule :

$$p(\text{choix} = \text{accepter}) = \frac{1}{1 + e^{-\frac{VC(R)+VC(E)}{\beta_{1O2D}} + \beta_m}}$$

(Équation 15)

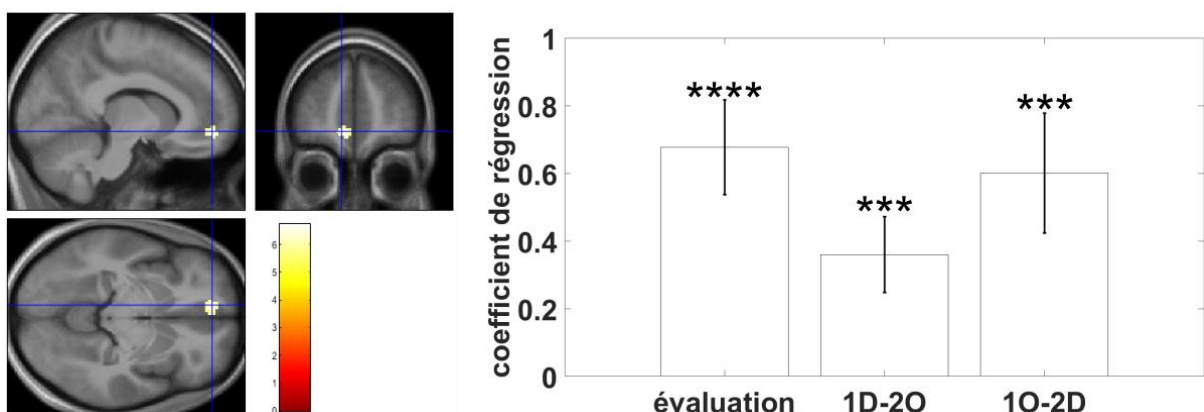
$B_{1O2D}$  était un paramètre libre positif spécifique à la tâche de choix 1O-2D représentant la stochasticité des choix, et  $\beta_m$  était un biais moteur. Le paramètre  $\beta_m$  permet de prendre en compte le fait qu'un participant pouvait avoir une tendance à répondre plus souvent à droite (ou à gauche) indépendamment des options présentées et des valeurs qui leur étaient associées.

### III.2.b Analyse des données IRMf

Nous avons utilisé le même GLM que dans la première étude (voir *Analyse des données d'IRMf: les modèles utilisés*) en remplaçant les scores issus de la tâche d'évaluation par les valeurs inférées par le modèle computationnel au sein de chaque tâche.

### III.3 Résultats

En regardant les corrélats cérébraux du contraste Val à travers les tâches, nous avons trouvé, là encore, un cluster principal de 70 voxels dans le vmPFC (voir **Figure 43**) avec un pic aux coordonnées (-10 ; 50 ; -8). A ce seuil, seuls deux autres clusters survivaient dont un de 11 voxels dans le cortex orbitofrontal latéral gauche aux coordonnées MNI (-28 ; 36 ; -14) et un autre d'un voxel aux coordonnées (-14 ; -10 ; -18). Ensuite, nous avons voulu voir si le fait d'avoir utilisé les valeurs cachées issues du modèle, plutôt que les évaluations fournies par les participants, modifiait les résultats lorsque nous regardions les corrélations entre le vmPFC et la variable Val tâche par tâche. Nous avons alors regardé dans le cluster identifié à travers les tâches si la corrélation était toujours stable à travers les tâches. Nous avons vu en effet que le vmPFC corrélait toujours significativement avec Val que ce soit dans la tâche d'évaluation ( $\beta_{\text{Val}} = 0.677 \pm 0.140$ ,  $p = 2 \cdot 10^{-5}$ ), la tâche de choix 1D-2O ( $\beta_{\text{Val}} = 0.360 \pm 0.112$ ,  $p = 0.003$ ) ou la tâche de choix 1O-2D ( $\beta = 0.601 \pm 0.177$ ,  $p = 0.002$ ). Ces valeurs étaient d'ailleurs très proches des valeurs obtenues lorsque nous avons utilisé les évaluations fournies par les participants directement (tâche d'évaluation :  $\beta_{\text{Val}} = 0.701 \pm 0.126$ ,  $p = 2 \cdot 10^{-6}$  ; tâche de choix 1D-2O :  $\beta_{\text{Val}} = 0.293 \pm 0.126$ ,  $p = 0.025$  ; tâche de choix 1O-2D :  $\beta_{\text{Val}} = 0.700 \pm 0.180$ ,  $p = 4 \cdot 10^{-4}$ ).



**Figure 43 : Corrélation entre le vmPFC et Val en utilisant les valeurs cachées issues du modèle computationnel.** A gauche, carte d'activation montrant le résultat pour le contraste Val à travers les tâches dans le GLM utilisant les valeurs cachées. La carte a été seuillée à  $p < 0.05$  corrigée pour les comparaisons multiples au niveau des voxels. A droite, activité du cluster

identifié à gauche pour chacune des tâches avec, dans l'ordre, la tâche d'évaluation (rating), la tâche de choix 1D-2O et la tâche de choix 1O-2D. Le sens des étoiles est le suivant : \* $p < 0.05$  ; \*\* $p < 0.01$  ; \*\*\* $p < 0.005$  ; \*\*\*\* $p < 0.001$ .

### III.4 Discussion

Nous avons essayé de modéliser la valeur subjective associée à chaque item comme une variable interne cachée de laquelle dériveraient les comportements des participants. Nous espérons ainsi améliorer les cartes d'activation cérébrales en lien avec la valeur subjective, mais les résultats obtenus avec cette approche étaient finalement très proches des résultats obtenus en utilisant les évaluations directement fournies par les participants (voir **Figure 43**). Cette approche, si elle est limitée dans le cadre de cette étude au vu de ces résultats, garde cependant un intérêt primordial lorsqu'il s'agit d'estimer une sensibilité globale des participants aux récompenses et aux efforts au travers de nos différentes tâches. Si ce n'est pas directement le sujet de cette thèse, il faut souligner en effet ici que ce modèle permet d'obtenir des scores individuels de sensibilité aux récompenses et de sensibilité aux efforts à travers l'ensemble des tâches présentées dans cette thèse, ainsi que dans d'autres qui n'ont pas été incluses ici. Le but de ces scores est d'avoir ensuite une utilisation clinique. Ces scores permettent ainsi d'établir la sensibilité aux récompenses et aux efforts de différentes catégories de patients neurologiques et psychiatriques en vue d'améliorer leur prise en charge et leurs traitements dans le cadre d'une étude plus large de notre équipe menée par Mathias Pessiglione et Fabien Vinckier. À défaut d'améliorer nos résultats ou de révéler d'autres corrélations intéressantes, cette étude nous aura donc permis de valider l'utilisation de ce modèle de valeurs cachées en imagerie en vue d'une utilisation ultérieure dans un but de diagnostic clinique.

## IV Valeur et attention

### IV.1 Introduction

Etant donné que nous avons pu enregistrer la position du regard dans chacune de nos tâches, en particulier dans les tâches de la première et de la troisième étude, nous avons voulu étudier l'interaction entre l'objet de l'attention et la valeur subjective. Il ne s'agissait pas d'une question centrale de notre paradigme aussi nos tâches n'ont certainement pas été optimisées pour faire cette étude. Cependant, nous pensons que les résultats de ces études préliminaires pourront intéresser le lecteur, notamment pour ce qui a trait au lien entre la valeur et l'attention. Nous avons vu en effet qu'une théorie, basée sur un modèle de diffusion attentionnel, propose qu'il existe un lien direct entre l'objet de l'attention et la valeur qui lui est attribué (voir *Modèles de choix alternatifs*). Cette théorie pose un lien direct entre la durée pendant laquelle un item fait l'objet de l'attention et la valeur subjective qui va lui être associée. On sait déjà qu'un item de valeur plus grande a tendance à attirer l'attention (Anderson, 2016, 2019; Anderson et al., 2011). Cette théorie propose qu'un lien de causalité existe aussi dans l'autre direction. Ainsi, un item qui serait observé plus longtemps aurait aussi tendance à être perçu comme possédant une valeur plus élevée, indépendamment de la raison qui fait qu'il serait observé plus longtemps (Krajbich and Rangel, 2011; Krajbich et al., 2010). Cette théorie implique donc qu'un item qui aurait une valeur plus élevée devrait aussi nécessairement être observé plus longtemps puisque le temps passé à observer un stimulus et la valeur associée à chaque stimulus devraient, de ce fait, être très fortement corrélés.

Pour tester cette prédiction, nous avons regardé les données oculaires dans les tâches où les participants de nos études devaient effectuer un choix. Nous avons ensuite étudié la dynamique moyenne du regard à travers les essais et les participants en fonction du sens de lecture et de l'option finalement choisie à chaque essai. De plus, nous avons ainsi aussi pu tester le temps passé à observer chacune des options proposées au cours du choix.

### IV.2 Méthodes d'acquisition dans chaque étude

#### IV.2.a Analyse de la position du regard dans la première étude

- Acquisition des données oculaires

La position du regard de l'œil droit a été enregistrée via l'appareil EyeLink 1000 plus (SR Research). Les données ont été acquises à une fréquence de 1000Hz. L'appareil a été calibré avant de démarrer l'expérience, une fois que le sujet était installé dans le scanner.

- **Analyse**

Pour la tâche de choix 1D-2O de la première étude et la tâche de choix 1O-2D de la première étude, nous avons défini une fenêtre assez large correspondant à chacune des options présentées à l'écran. Pour la tâche de choix 1D-2O, il s'agissait ainsi de deux rectangles, un à gauche et un à droite, autour de la zone où chacune des options était présentée. Les dimensions du rectangle défini étaient fixes à travers les essais pour un sujet donné mais nous avons parfois dû les adapter d'un sujet à l'autre à cause de décalages de l'eye-tracker. Pour la tâche de choix 1O-2D, nous avons défini de même 4 rectangles : un autour de la zone où la récompense était affichée à l'écran, un autour de la zone où l'effort était affiché à l'écran, un autour de la zone où l'option « OUI » et un autour de la zone où l'option « NON » étaient affichées. Il faut rappeler ici que la position du « OUI » et du « NON » alternait d'un essai sur l'autre. De même que pour la tâche de choix 1D-2O, les dimensions du rectangle utilisé étaient fixes pour un participant donné mais nous les avons adaptés d'un participant à l'autre pour compenser les décalages de l'eye-tracker.

Nous avons regroupé les données de tous les essais d'un bloc donné ensemble en regroupant aussi les données à travers les 3 sessions. Toutes les données en dehors de nos rectangles d'intérêt étaient ignorées. Les données ont été regroupées de 3 manières : soit 1) en les calant par rapport au temps d'apparition des stimuli, soit 2) en les calant sur le moment de la réponse, soit 3) en les normalisant du moment d'apparition des stimuli jusqu'à la réponse. Dans ce dernier cas, comme la durée variait d'un essai à l'autre, nous avons réparti les données en 1000 échantillons pour chaque essai (1 échantillon correspond donc à 0.1% du total) en moyennant les données pour chaque bin.

Pour chaque point temporel d'un essai donné d'un bloc de la tâche de choix 1D-2O, nous avons effectué deux types d'analyse :

- 1) Nous avons mesuré, pour chaque point temporel, pour un bloc donné, le pourcentage d'essais où le participant regardait l'option de gauche par rapport au nombre d'essais où le participant regardait l'une des deux options. Pour un essai donné, il nous suffisait ainsi d'encoder les données de manière binaire avec un 1 si le participant regardait l'option de gauche et un 0 si le participant regardait l'option de droite.
- 2) Nous avons mesuré, pour chaque point temporel, pour un bloc donné, le pourcentage d'essais où le participant regardait l'option choisie à la fin de l'essai, par rapport au nombre d'essais où le participant regardait l'une des deux options. Pour un essai

donné, il nous suffisait ainsi d'encoder les données de manière binaire avec un 1 si le participant regardait l'option finalement choisie et un 0 si le participant regardait l'option non-choisie à la fin de l'essai.

De plus, nous avons calculé la proportion totale du temps passé à regarder l'option choisie par rapport au temps passé à regarder les deux options pour chaque bloc et nous avons moyenné cette valeur à travers les participants. Nous avons ensuite fait un t.test pour voir si ce pourcentage était significativement différent de 50% pour chaque bloc.

Pour la tâche de choix 1O-2D, nous avons appliqué la même logique afin de voir, à chaque point temporel, le pourcentage d'essais où le participant regardait, soit la récompense, soit l'effort, soit l'option « OUI », soit l'option « NON » par rapport au nombre d'essais où le participant regardait l'une de ces 4 options. Pour un essai donné, nous avons ainsi 4 variables nous permettant de savoir laquelle des 4 options était observée (ou si aucune d'entre elles n'était observée). Nous avons, de plus, séparé les essais en fonction de la réponse finale des participants, c'est-à-dire que nous avons analysé séparément les essais où les participants ont fini par accepter l'offre présentée des essais où les participants ont rejeté l'offre. Nous voulions voir ainsi si le temps passé à observer la récompense par rapport à l'effort de chaque option était corrélé à la réponse finale donnée par les participants à chaque essai.

Pour les deux tâches, si le regard était situé en dehors de nos rectangles d'intérêt, l'essai n'était pas comptabilisé dans l'analyse. Ensuite, nous avons moyenné les résultats, pour chacune de ces analyses pour chaque bloc, à travers les participants et nous avons aussi calculé l'erreur type de la moyenne (*Standard Error of the Mean*, SEM) à travers les participants. Nous avons finalement lissé la moyenne et la SEM obtenues avec une fenêtre de 100ms pour les données brutes et une fenêtre de 10 échantillons pour les données normalisées.

### **IV.2.b Analyse de la position du regard dans la troisième étude**

- **Acquisition**

La position du regard de l'œil droit a été enregistrée via l'appareil EyeLink 1000 plus (SR Research). Les données ont été acquises à une fréquence de 1000Hz. L'appareil a été calibré au début de l'expérience une fois que le sujet était installé dans le scanner.

- **Analyse**

Pour chaque session de chaque individu, nous avons défini manuellement deux rectangles, un à gauche et un à droite, autour de la zone où chacune des options était présentée.

## Appendices

Les dimensions du rectangle défini étaient fixes à travers les essais pour un sujet donné mais nous avons parfois dû les adapter d'un sujet à l'autre à cause de décalages légers de l'eye-tracker.

Nous avons regroupé les données de tous les essais d'un bloc donné ensemble en regroupant aussi les données à travers les 3 sessions. Toutes les données en dehors de nos rectangles d'intérêt étaient ignorées. Les données ont été regroupées en les calant par rapport au temps d'apparition des stimuli.

Pour chaque point temporel d'un essai donné, nous avons effectué deux types d'analyse :

- 1) Nous avons mesuré, pour chaque point temporel, le pourcentage d'essais où le participant regardait l'option de gauche par rapport au nombre d'essais où le participant regardait l'une des deux options. Pour un essai donné, il nous suffisait ainsi d'encoder les données de manière binaire avec un 1 si le participant regardait l'option de gauche et un 0 si le participant regardait l'option de droite.
- 2) Nous avons mesuré, pour chaque point temporel, le pourcentage d'essais où le participant regardait l'option choisie à la fin de l'essai, par rapport au nombre d'essais où le participant regardait l'une des deux options. Pour un essai donné, il nous suffisait ainsi d'encoder les données de manière binaire avec un 1 si le participant regardait l'option finalement choisie et un 0 si le participant regardait l'autre option.

Si le regard était situé en dehors de nos rectangles d'intérêt, l'essai n'était pas comptabilisé dans l'analyse. Ensuite, nous avons moyenné les résultats à travers les participants et nous avons aussi calculé l'erreur type de la moyenne (SEM) à travers les participants. Nous avons finalement lissé la moyenne et la SEM obtenues avec une fenêtre de 100ms.

Nous avons inclus 23 participants dans cette analyse étant donné que même chez le participant qui a été exclu de l'IRMf à cause du mouvement, la pupille a été correctement enregistrée. Seules 3 sessions de 3 participants différents ont ainsi dû être exclues de l'analyse du diamètre pupillaire car la pupille n'y a pas été correctement enregistrée.

### IV.3 Résultats

#### IV.3.a Position du regard et option choisie dans la première étude

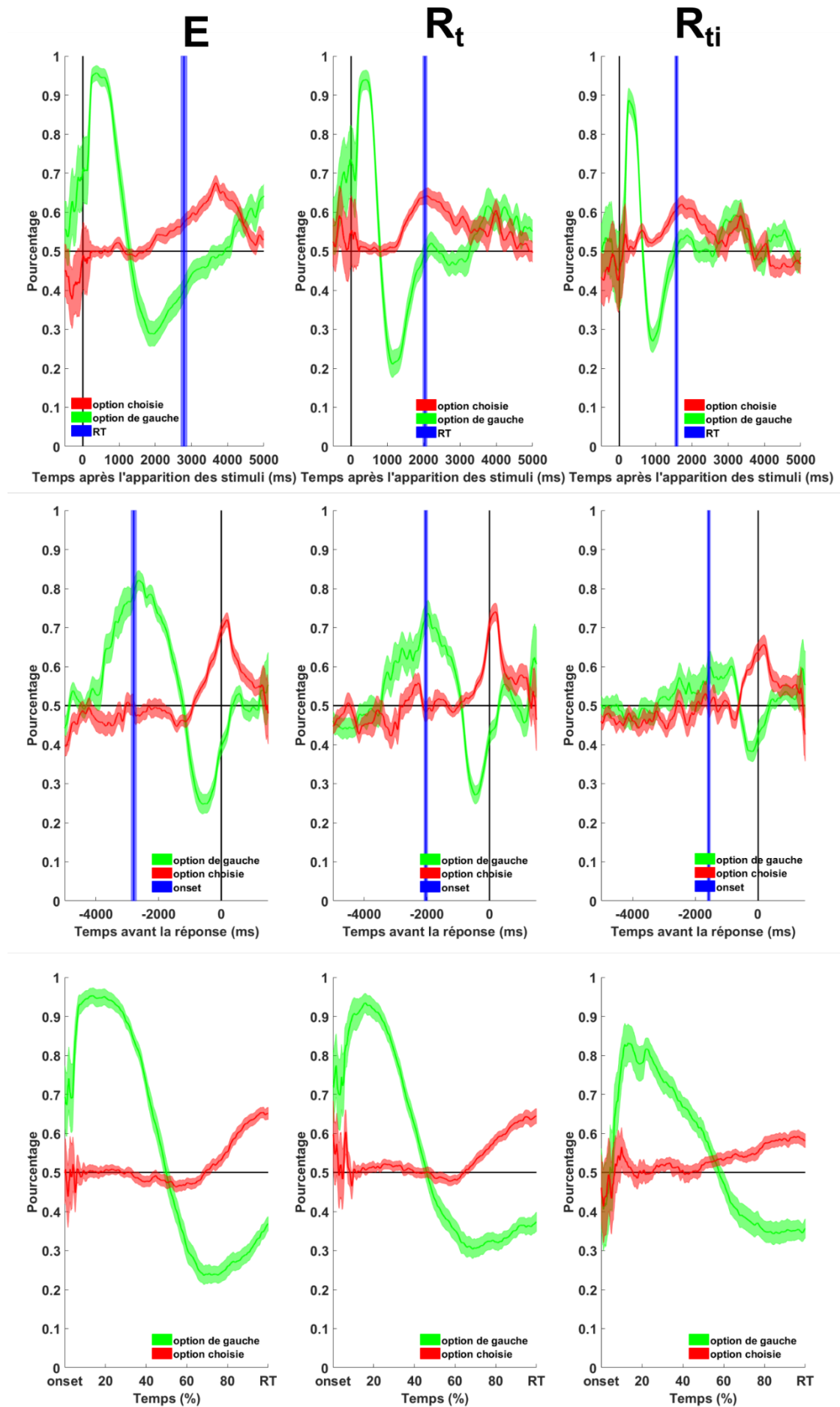
Comme on peut le voir sur la **Figure 44**, peu importe le bloc, les participants avaient tendance à regarder les options dans le sens de lecture, en français, dans la tâche de choix 1D-



## Appendices

2O, c'est-à-dire de gauche à droite. Puis, une fois qu'ils avaient pris connaissance des options à disposition, leur regard s'orientait généralement vers l'option choisie quelques centaines de millisecondes avant de donner leur réponse. On peut d'ailleurs constater que les participants ont généralement passé plus de temps à observer l'option finalement choisie que l'autre option dans les blocs d'effort (51.83% ; p.value =  $7 \cdot 10^{-4}$ ), dans les blocs de récompenses présentées sous forme de texte (53.07% ; p.value =  $2 \cdot 10^{-8}$ ) et dans les blocs de récompenses présentées sous forme de texte accompagné d'une image (54.62% ; p.value =  $3 \cdot 10^{-5}$ ).

Appendices



**Figure 44 : Position du regard au cours du temps dans la tâche de choix 1D-2O.** Dans chaque bloc, on voit que le regard s'oriente d'abord dans le sens de la lecture, c'est-à-dire vers l'option de gauche puis vers l'option de droite. Ce n'est qu'à la fin de l'essai qu'on constate qu'il se dirige, en moyenne, vers l'option choisie. Cette figure représente le pourcentage moyen à travers les participants, ainsi que l'erreur type de la moyenne, des essais où, pour un point temporel donné, l'option observée est l'option présentée à gauche (courbes vertes) par rapport à l'une des deux options, ainsi que le pourcentage d'essais où l'option observée est l'option finalement choisie à la fin de l'essai (courbes rouges). Les figures sont organisées en fonction du type de bloc avec, de gauche à droite, les blocs d'efforts (E), les blocs de récompenses présentées sous forme de texte ( $R_t$ ) et les blocs de récompenses présentées sous forme de texte accompagné d'une image ( $R_{ti}$ ). Les données ont été calées selon 3 méthodes avec, en haut, un calage par rapport au moment d'apparition des stimuli à l'écran, au milieu un calage par rapport au moment de la réponse et, en bas, les données sont normalisées entre le moment d'apparition des stimuli et le moment de la réponse. Sur les figures du haut et du milieu, le temps de réaction moyen ainsi que l'erreur type de la moyenne de ce temps de réaction à travers les participants est représenté en bleu.  $N = 29$  participants.

Dans la tâche de choix 1O-2D, on constate que, l'effort étant systématiquement au-dessus de la récompense (voir **Figure 25**) et le sens de lecture, en France, étant de la gauche vers la droite et de haut vers le bas, les participants observaient, en moyenne, les options dans le sens de lecture (voir **Figure 45**). En d'autres termes, indépendamment de leur choix final, les participants regardaient d'abord l'effort proposé, puis la récompense proposée. Dans les dernières centaines de millisecondes avant de donner leur réponse, leur regard se concentrait sur l'option (« OUI » ou « NON ») qu'ils allaient finir par donner en fonction du calcul bénéfices/coûts effectué (voir **Figure 45**). Les participants avaient tendance à passer plus de temps à observer l'effort présenté (59.25%) que la récompense présentée (38.02%) ( $p.value = 9 \cdot 10^{-5}$ ). Cependant, le texte associé aux efforts ( $n = 6 \pm 2$  mots, moyenne  $\pm$  écart type) était aussi généralement plus long ( $p.value = 2 \cdot 10^{-16}$ ) que celui associé aux récompenses ( $n = 4 \pm 1$  mots, moyenne  $\pm$  écart type) ce qui rend toute comparaison difficile. De plus, nous n'avons pas constaté de différence significative ( $p.value = 0.759$ ) dans le temps passé à observer les récompenses entre les essais où l'offre était finalement acceptée (38.43%) et les essais où l'offre était finalement rejetée (38.07%). Il n'y avait pas non plus de différence significative ( $p.value = 0.894$ ) dans le temps passé à observer les efforts entre les essais où l'offre était finalement acceptée (58.95%) et les essais où l'offre était finalement rejetée (59.12%). La seule différence entre ces deux types d'essais était le temps passé à observer l'option « OUI » ou l'option « NON » quelques centaines de millisecondes avant la réponse qui prédisait la réponse finalement donnée (voir **Figure 45**).

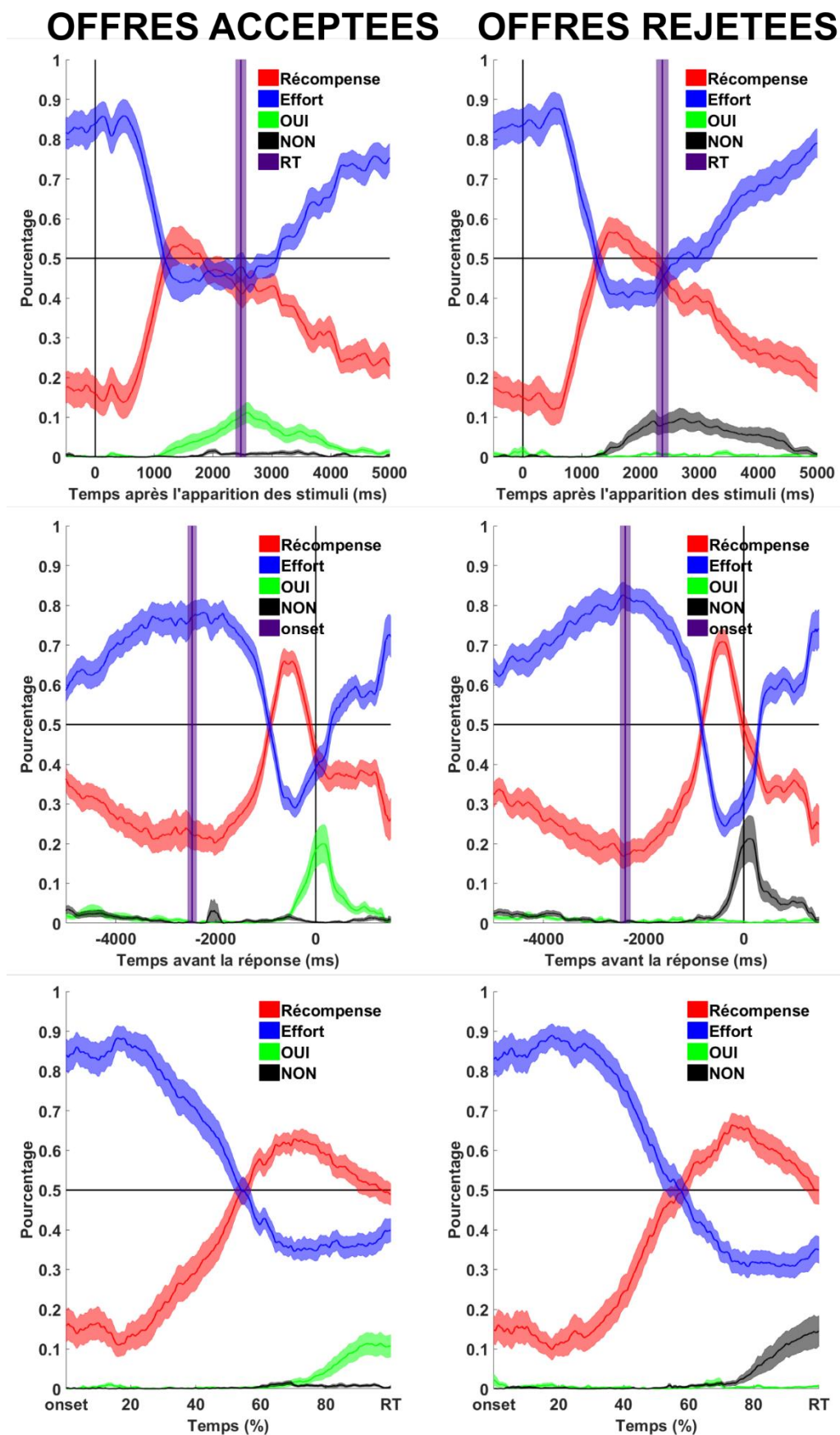
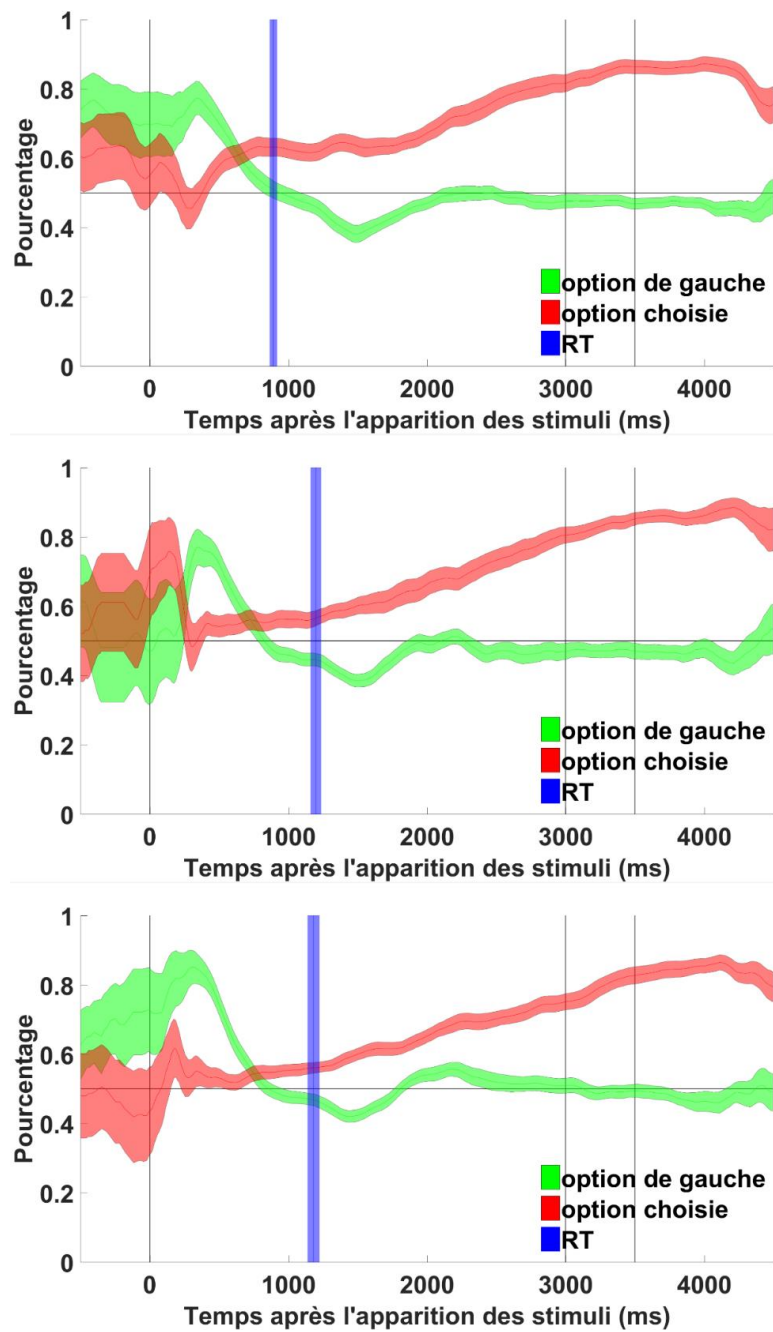


Figure 45 : Position du regard au cours du temps dans la tâche de choix 10-2D. On constate que le regard s'oriente d'abord dans le sens de la lecture, c'est-à-dire du haut vers le

bas en regardant d'abord vers l'effort puis vers la récompense. Ce n'est qu'à la fin de l'essai qu'on constate qu'il se dirige, en moyenne, vers l'option choisie : « OUI » lorsque l'offre est acceptée (figures de gauche) et « NON » lorsqu'elle est refusée (figures de droite). Cette figure représente le pourcentage moyen à travers les participants, ainsi que l'erreur type de la moyenne, pour chaque point temporel des essais passés à regarder la récompense, l'effort, l'option « OUI » ou l'option « NON ». Les données ont été calées selon 3 méthodes avec, en haut, un calage par rapport au moment d'apparition des stimuli à l'écran, au milieu un calage par rapport au moment de la réponse et, en bas, les données sont normalisées entre le moment d'apparition des stimuli et le moment de la réponse. Dans toutes les figures, les codes couleurs sont fixes avec, en rouge, le pourcentage d'essais où la récompense est observée, en bleu, le pourcentage d'essais où l'effort est observé, en vert, le pourcentage d'essais où l'option « OUI » est observée et enfin en noir le pourcentage d'essais où l'option « NON » est observée. Sur les figures du haut et du milieu, le temps de réaction moyen ainsi que l'erreur type de la moyenne de ce temps de réaction à travers les participants est représenté en violet. N = 28 participants.

### IV.3.a Position du regard et option choisie dans la troisième étude

Nous avons, là encore (voir **Figure 44** et **Figure 45**), répliqué un résultat classique en constatant que les participants ont tendance à d'abord observer l'option qui apparaît dans le sens de la lecture, même si en l'occurrence il s'agissait de symboles et non de texte, puis l'option qu'ils vont finir par choisir (voir **Figure 46**). Ce phénomène du regard qui s'oriente vers l'option finalement choisie se vérifie en tout cas indépendamment de la valence de la paire observée. Nous avons effectivement pu l'observer aussi bien pour la paire associée à des gains, celle associée à un résultat toujours neutre et celle associée à des pertes. Les participants ont d'ailleurs passé plus de temps à regarder l'option choisie que l'autre option aussi bien pour les paires associées à des gains ( $62.70 \pm 2.53$  %, moyenne  $\pm$  erreur type de la moyenne à travers les participants ;  $p.value = 5 \cdot 10^{-5}$ ), pour les paires associées à un résultat toujours nul ( $55.74 \pm 2.53$  % ;  $p.value = 0.033$ ) et pour les paires associées à des pertes ( $55.61 \pm 1.88$  % ;  $p.value = 0.007$ ). Le ratio était cependant plus élevé dans le cas de la paire associée à des gains par rapport à la paire associée à un résultat toujours neutre ( $p = 0.038$ ) ou la paire associée à des pertes ( $p = 0.021$ ), mais cette différence est probablement due au fait que la délibération était plus rapide dans le cas des paires associées à des gains (voir **Figure 37**) et que les participants devaient garder le doigt appuyé sur leur réponse jusqu'à ce que l'essai prenne fin. On ne constate d'ailleurs pas de différence dans ce ratio entre la paire associée à des pertes par rapport à la paire associée à un résultat toujours nul ( $p = 0.962$ ).



**Figure 46 : Position du regard au cours du temps dans la tâche d'apprentissage.** Les figures montrent l'évolution de la position du regard pour la paire associée à des gains en haut, la paire associée à un résultat toujours nul au milieu et la paire associée à des pertes en bas. Les données ont été moyennées à travers les 23 participants. Les courbes indiquent ainsi la moyenne ainsi que l'erreur type de la moyenne de la position du regard à travers les 23 participants de l'étude. Le temps de réaction (RT) moyen ainsi que l'erreur type de la moyenne de ce temps de réaction pour chacune des paires à travers les participants est représenté en bleu sur chaque figure. Au bout de 3s après l'apparition des symboles à l'écran, l'option ayant été choisie était entourée d'un rectangle rouge pendant 0.5s. Après cette période, le feedback apparaissait à l'écran. Des traits noirs verticaux ont été placés sur chaque figure pour indiquer les frontières entre les différentes parties d'un essai.

#### IV.4 Discussion

Nous avons répliqué un résultat classique en montrant que **l'attention, via la position du regard, était sensible à la valeur des options présentées**. L'attention était en effet attirée vers l'option avec la valeur subjective la plus haute qui allait être choisie à la fin de l'essai (voir **Figure 44**). Cependant, dans notre cas, il pourrait aussi s'agir d'un moyen simple pour le cerveau de vérifier la correspondance entre la position de l'option à sélectionner par l'action motrice de la main et la position de l'option qu'il souhaite sélectionner sur l'écran. On voit d'ailleurs bien dans la tâche de choix 1O-2D que l'option « OUI » ou l'option « NON » est observée à la fin de chaque essai en fonction de la réponse qui va être donnée (voir **Figure 45**) alors que ces mots ne sont pas, *a priori*, associés à une valeur subjective. De même, dans la troisième étude (voir **Figure 46**), on peut voir que le regard se dirige vers l'option choisie à la fin de l'essai même dans le cas des paires systématiquement associées à un résultat neutre, c'est-à-dire où le résultat est toujours nul indépendamment de l'item choisi. Il serait intéressant de vérifier si le lien entre l'attention et la valeur pourrait être décorrélé de l'initiation d'une action motrice bien que ce lien paraisse évolutivement pertinent.

De plus, si l'on considère le biais inverse prôné par la théorie du modèle de diffusion attentionnel, c'est-à-dire que **la valeur subjective est biaisée par ce sur quoi l'attention est dirigée** (Krajbich and Rangel, 2011; Krajbich et al., 2010), nous trouvons quelques éléments incohérents avec cette théorie dans nos données. En effet, d'après cette théorie, on s'attendrait à ce que, dans les essais où l'offre était finalement acceptée dans la tâche de choix 1O-2D, les participants passent plus de temps à regarder la récompense de l'offre proposée relativement au temps total passé à regarder l'une des options par rapport aux essais où l'offre était finalement rejetée, et réciproquement pour l'effort. Or nous n'avons pas constaté de différence dans le temps relatif passé à observer les options appétitives, c'est-à-dire les récompenses, et le temps passé à observer les options aversives, c'est-à-dire les efforts, entre les essais où l'offre était acceptée et les essais où l'offre était rejetée dans la tâche de choix 1O-2D (voir *Position du regard et option choisie*). Nous n'avons cependant pas cherché à équilibrer le nombre de mots et de syllabes à travers les options ce qui rend difficile toute interprétation du temps passé à regarder chaque option. Cependant, s'il existe un effet quelconque de l'attention vers la valeur, il semblerait qu'il ne soit pas assez prééminent, dans notre tâche tout du moins, pour que l'on puisse le détecter. La valeur associée à chaque item rend visiblement donc bien mieux compte du choix d'accepter ou de rejeter une offre que le temps passé à observer la récompense ou l'effort associés à chaque proposition.

# Liste des figures

## Chapitre 1: Introduction

<b>Figure 1 : Célèbre caricature anonyme du début du <sup>XX</sup><sup>ème</sup> siècle censée résumer la pensée de Freud..</b> .....	20
<b>Figure 2: Illustration des grands principes de la théorie des perspectives.</b> .....	33
<b>Figure 3 : Zones effectrices des neurones dopaminergiques.</b> .....	37
<b>Figure 4 : Centres principaux liés à la sensation de plaisir dans un cerveau de rongeur.</b> .....	39
<b>Figure 5 : L’erreur de prédiction dans les neurones dopaminergiques.</b> .....	40
<b>Figure 6 : Relation entre différentes mesures de l’activité électrique des neurones du cortex visuel d’un singe et le signal BOLD dans le cortex visuel.</b> .....	41
<b>Figure 7 : Corrélats neuraux de la valeur subjective.</b> .....	44
<b>Figure 8 : Correspondances anatomiques des structures fonctionnelles du réseau de la valeur.</b> .....	48
<b>Figure 9 : La transformation subjective des probabilités vues par la théorie des perspectives.</b> .....	59
<b>Figure 10 : Cluster au sein du cortex préfrontal dorso-médian corrélant à la volatilité.</b> 63	
<b>Figure 11 : Activité du striatum gauche et du striatum droit en fonction de la probabilité.</b> .....	64
<b>Figure 12 : Le réseau des valeurs est plus actif pour les récompenses immédiates.</b> .....	70
<b>Figure 13 : Arbre phylogénétique des animaux avec un focus sur les différentes innovations évolutives qui ont mené jusqu’à l’homme.</b> .....	73
<b>Figure 14 : Différents types de décompte de la valeur subjective par l’effort physique sur la base de données réelles.</b> .....	76
<b>Figure 15 : Impact des lésions aux ganglions de la base sur la capacité à doser son effort en fonction des récompenses en jeu.</b> .....	79



## Liste des figures

<b>Figure 16 : Méta-analyse montrant les zones liées aux fonctions exécutives sur la base de 193 études chez l'adulte sain..</b> .....	89
<b>Figure 17 : Aires cérébrales liées à la sensation de douleur..</b> .....	97
<b>Figure 18 : Réseaux communs entre punitions physiques et punitions psychologiques..</b> 98	
<b>Figure 19 : Le délai, l'incertitude, l'effort et les punitions sont liés à l'activité du dmPFC et de l'insula antérieure..</b> .....	104
<b>Figure 20: Différentes étapes dans les mécanismes de prise de décision d'après (Rangel et al., 2008).</b> .....	115
<b>Figure 21 : Schéma général du modèle de diffusion (<i>drift-diffusion model</i>).</b> .....	122
<b>Figure 22 : Un gradient dans le cortex préfrontal médian : de l'encodage de la valeur subjective vers l'encodage de la confiance.</b> .....	134
<b>Figure 23 : Comparaison entre clusters anatomiques dans le cortex préfrontal médian et clusters paramétriquement corrélés à la confiance.</b> .....	135
<b>Figure 24 : Dynamique du vmPFC en lien avec la valeur subjective (SV) et la confiance (<i>confidence</i>) au cours du temps.</b> .....	136

## Chapitre 2: Études Expérimentales

<b>Figure 25 : Tâches expérimentales.</b> .....	146
<b>Figure 26 : Comportement dans les tâches basées sur la valeur.</b> .....	160
<b>Figure 27 : Résultats neuraux pour Val, Conf et DT.</b> .....	166
<b>Figure 28 : Corrélations entre le diamètre pupillaire et l'activité du vmPFC, du mPFC et du dmPFC dans chacune des tâches de prise de décision basée sur la valeur.</b> .....	170
<b>Figure 29 : Tâche d'effort physique (A) et d'effort mental (B).</b> .....	181
<b>Figure 30 : Différents niveaux de difficulté en fonction du type de paire dans la tâche d'effort mental.</b> .....	194
<b>Figure 31 : Temps de réaction en fonction des enjeux monétaires et du numéro d'essai dans la tâche d'effort physique et dans la tâche d'effort mental.</b> .....	197
<b>Figure 32 : Performance en fonction du numéro d'essai au sein de chaque session.</b> .....	198

## Liste des figures

<b>Figure 33 : Performance en fonction des enjeux monétaires et du numéro d'essai dans la tâche d'effort physique et dans la tâche d'effort mental.</b> .....	200
<b>Figure 34 : Résultats neuraux.</b> .....	202
<b>Figure 35 : Décomposition de l'activité du striatum, du dmPFC et du dlPFC au moment de l'apparition des enjeux en fonction des variables impactant la performance.</b> .....	205
<b>Figure 36 : Tâche d'apprentissage.</b> .....	219
<b>Figure 37 : Le comportement dans la tâche d'apprentissage.</b> .....	225
<b>Figure 38 : Les corrélats cérébraux de Val, Conf et DT dans la tâche d'apprentissage.</b> .....	227
<b>Figure 39 : Lien entre Val, Conf, DT et l'activité du vmPFC et du dmPFC.</b> .....	229
<b>Figure 40 : Erreur de Prédiction au moment du feedback et activité du vmPFC.</b> .....	230

## **Chapitre 3: Discussion générale**

<b>Figure 41 : Le dmPFC reflète les ressources allouées et pas le temps de réaction.</b> .....	248
--	-----

## **Appendices**

<b>Figure 42 : Les différentes séquences d'acquisition en IRMf donnent des résultats équivalents pour l'activité du vmPFC en corrélation avec la valeur.</b> .....	332
<b>Figure 43 : Corrélation entre le vmPFC et Val en utilisant les valeurs cachées issues du modèle computationnel.</b> .....	337
<b>Figure 44 : Position du regard au cours du temps dans la tâche de choix 1D-2O.</b> .....	345
<b>Figure 45 : Position du regard au cours du temps dans la tâche de choix 1O-2D.</b> .....	346
<b>Figure 46 : Position du regard au cours du temps dans la tâche d'apprentissage.</b> .....	348