

L'estimation des valeurs positives le réseau des valeurs

Dans le Trésor de la Langue Française², 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

² <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.

« 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^{er} 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^{ème} siècle ap. JC), tout comme les *Pensées pour moi-même* de **Marc-Aurèle** (II^{ème} siècle ap. JC), *La Vie heureuse* de Sénèque (I^{er} siècle ap. JC), les apophtegmes des **Sept Sages** de la Grèce antique (VII^{ème}-VI^{ème} 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^{ème}-V^{ème} 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^{ème} siècle av. JC) ou **Eudoxe de Cnide** (IV^{ème} 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^{ème} 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^{ème} 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^{ème} 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 ^{XX}^{ème} 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^{ème} et le début du XX^{ème} 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^{ème} 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 2013³. 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.⁴. 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*

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

⁴ https://www.lemonde.fr/planete/article/2016/10/01/le-sulfureux-parcours-du-telephone-portable-des-mines-aux-filieres-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 classifier 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 mathématiciens⁵. C'est à la suite de ces échanges épistolaires en effet que Pascal proposera la formule de l'**espérance** :

$$E = p * x$$

⁵ <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.* »⁶. 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 :

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

⁶ <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 α 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).

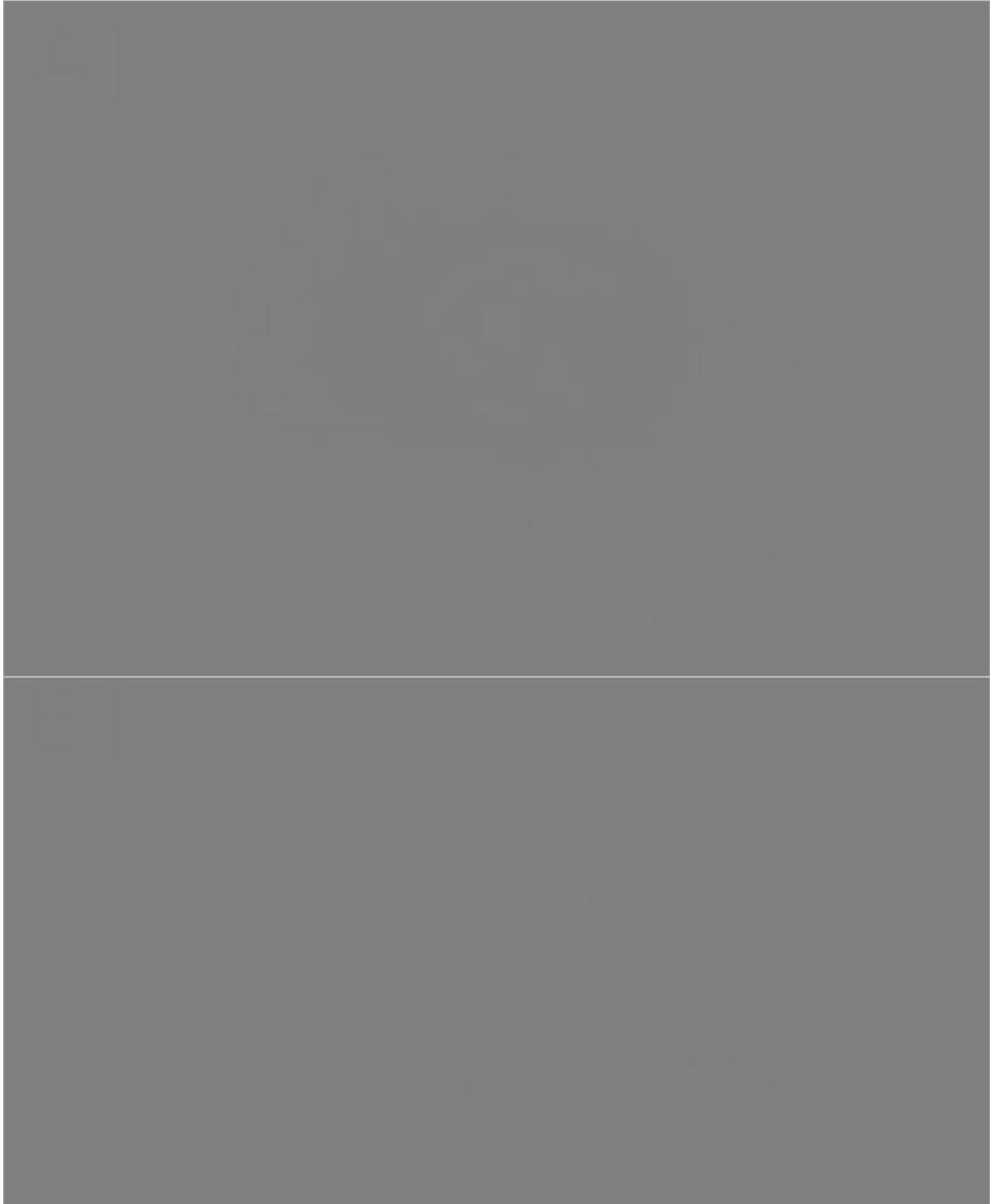


Figure 3 : Zones effectrices des neurones dopaminergiques. A] Schéma simplifié des zones innervées par des neurones dopaminergiques dans le cerveau humain. Figure issue de (Glimcher, 2014). B] Schéma détaillé du circuit reliant l'aire tegmentale ventrale au noyau accumbens dans un cerveau de rongeur. 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'homme⁷. 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,

⁷ <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)

Jusqu'à-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^{ème} 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^{ème} 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 ;