



Décision et action

Alizée Lopez-Persem, Mehdi Khamassi

► To cite this version:

Alizée Lopez-Persem, Mehdi Khamassi. Décision et action. Khamassi, M. (Ed.) Neurosciences cognitives, 2021. hal-03411266

HAL Id: hal-03411266

<https://hal.science/hal-03411266>

Submitted on 2 Nov 2021

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

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

Décision et action

Alizée Lopez-Persem¹, Mehdi Khamassi²

¹Institut du Cerveau (ICM), INSERM, Sorbonne Université, Paris

²Institut des Systèmes Intelligents et de Robotique, CNRS, Sorbonne Université, Paris

Ceci est le preprint du chapitre de livre : Lopez-Persem, A. & Khamassi, M. (2021). *Décision et action*. Dans Khamassi, M. (Ed.) *Neurosciences cognitives*. Éditions De Boeck Supérieur.

Table des matières

Introduction	1
1. Les différents types de décisions	2
a. Décisions économiques	2
b. Décisions perceptives	3
2. Décisions : de la valeur à l'action.....	4
a. Le concept de valeur subjective	4
b. Valeurs et choix.....	10
c. Expliquer les choix à partir des valeurs subjectives	13
3. Bases neurales de la prise de décision.....	16
a. Représentation des valeurs de récompense dans le cerveau.....	17
b. Modulation des valeurs et contrôle	23
c. Choix	25
4. Les troubles de la prise de décision	27
5. À retenir	31
6. Pour aller plus loin	32
7. Bibliographie	33

Introduction

La prise de décision est un processus cognitif qui permet de sélectionner une option ou une action parmi un ensemble d'alternatives. Les choix, qui peuvent se traduire par le comportement du sujet, résultent des processus de décision. Ces processus peuvent s'appliquer à des situations très concrètes (choisir un plat dans un menu de restaurant, choisir une direction lorsque l'on marche), mais aussi à des situations plus abstraites (comme choisir parmi des croyances possibles). En neurosciences cognitives, l'étude de la prise de décision est principalement axée sur deux domaines : les décisions perceptives, et les décisions basées sur les récompenses (ou préférences). Comme nous allons le voir, ces dernières sont souvent qualifiées de « décisions économiques » car les économistes ont grandement contribué à modéliser ces décisions comme un

processus de comparaison des valeurs subjectives des options. Nous verrons dans ce chapitre les distinctions qui sont faites en psychologie cognitive sur les différentes façons de prendre des décisions, et ce que cela implique en termes de comportement. Nous verrons également quelles sont les bases neurales de la prise de décision, et évoquerons certains troubles (ou pathologies) de la prise de décision. Le **chapitre 10** de cet ouvrage présentera ensuite les méthodes pour concevoir des modèles computationnels de prise de décision.

1. Les différents types de décisions

Tout d'abord, il est important de bien définir les différents types de décisions auxquels on s'intéresse, car ceux-ci n'impliquent pas forcément les mêmes régions cérébrales et ne se traduisent pas de la même façon au niveau du comportement. Lorsque l'on pense aux types de décisions que l'être humain peut être amené à effectuer au cours de son existence, on pense naturellement au fait de décider quel type de nourriture manger, avec quels individus devenir ami ou se mettre en couple, à quelle activité s'adonner au cours de la journée, quelles zones de l'environnement explorer pour essayer d'y trouver de nouvelles sources de nourriture, ou encore décider si un danger imminent (comme croiser un lion) requière qu'on parte en courant ou au contraire qu'on s'immobilise sans un bruit.

Pourtant, il existe d'autres types de décisions que notre cerveau doit prendre fréquemment, sans nécessairement que nous en ayons conscience. Par exemple, si en nous baladant en forêt nous apercevons entre les feuilles des formes qui ressemblent à deux points au-dessus d'un trait, nous sommes dans une situation *d'ambiguïté perceptive* sur la présence ou non d'un visage, et donc potentiellement d'un danger (s'il s'agit bien d'un lion)¹. Avant de décider s'il faut partir en courant ou non, il nous faut donc d'abord décider de ce que nous voyons. On parle de *décision perceptive*.

Nous allons donc d'abord bien définir les différents types de décision considérés en psychologie cognitive, car elles n'impliquent pas les mêmes processus au niveau cérébral. Puis nous nous focaliserons dans le reste du chapitre sur les décisions que peut prendre un individu en vue de faire une action.

a. Décisions économiques

Comment décide-t-on entre manger une pizza ou un curry ? Quel processus (ou suite de processus) nous permet de décider d'aller nous promener en forêt plutôt qu'au bord

¹ Le phénomène consistant à spontanément voir des visages même quand il n'y en a pas s'appelle la paréidolie. D'un point de vue évolutif, on pense que ceci a permis aux individus de réagir plus vite dans le cas d'un danger réel (e.g., prédateur, ennemi), et donc d'avoir plus de chances de survivre.

d'un lac, ou encore de continuer en master plutôt que d'aller travailler en entreprise tout de suite après la licence ?

On considère dans tous ces cas qu'il y a besoin de comparer un certain nombre d'*attributs* des différentes options et ensuite de combiner ces attributs pour en déduire quelle est la meilleure option. Par exemple, un sujet peut préférer le goût de la pizza mais considérer qu'un curry est meilleur pour la santé. Il peut également considérer que la pizza coûte moins cher, mais qu'elle est plus longue à préparer (s'il doit la cuisiner tout seul), ou encore que même s'il préfère la pizza en général, celle proposée par le restaurant d'à côté n'est pas aussi bonne que leur curry. Afin de prendre une décision, il faut donc pouvoir combiner ces différents attributs et que le cerveau trouve une « monnaie commune » pour en déduire la *valeur subjective* totale de l'option A et de l'option B, afin de choisir la meilleure (pour l'individu).

Les travaux en économie sur la manière dont il convient de prendre des décisions basées sur la valeur (le plus souvent afin de maximiser un certain gain monétaire) ont nourri depuis une vingtaine d'années les recherches sur la prise de décision en psychologie cognitive et en neurosciences. Comme nous allons le voir dans ce chapitre, il s'avère que certaines aires cérébrales comme le cortex orbitofrontal ou le striatum ont des niveaux d'activité qui correspondent aux (« encodent » les) valeurs subjectives d'options parmi lesquelles un sujet doit choisir. On parle donc de *décisions économiques* pour caractériser toutes les décisions qui reposent sur la comparaison de valeurs, ceci même lorsque les gains en jeu ne sont pas monétaires mais par exemple de type énergétique (e.g., nourriture) ou encore social (e.g., l'acquiescement de nos pairs).

b. Décisions perceptives

Les processus de prise de décision perceptive concernent principalement les signaux sensoriels ambigus que nous recevons du monde extérieur. Comme par exemple les couleurs, les odeurs, les sons, etc. Ces signaux sont traités dans un premier temps par les cortex sensoriels, qui ont chacun leur propre modalité (visuel pour la vision, auditif pour les sons, olfactif pour les odeurs, etc. ; voir le **chapitre 2** de cet ouvrage). Ils sont ensuite traités par les cortex associatifs, comme le cortex pariétal, qui va intégrer les différentes informations sensorielles (Mesulam, 1998). Les processus de prise de décision perceptive dérivent des choix à partir de ces signaux, de manière à approuver une certaine hypothèse sur l'état estimé comme réel du monde (P. Glimcher, 2014). Des exemples typiques de prise de décision perceptive se produisent dans des situations d'ambiguïté perceptive (Qu'est-ce que je sens ? Que vois-je ?). Un exemple classique est celui du cube de Necker, dans lequel on doit décider de la perspective avec laquelle on perçoit le cube (**Figure 1**). Un autre exemple très connu, et que l'on peut trouver facilement en vidéo sur Internet, représente une danseuse qui tourne sur elle-même : notre système visuel doit alors décider s'il interprète la rotation de la danseuse comme se produisant dans le sens horaire ou antihoraire, car l'image en mouvement est compatible avec ces deux interprétations. De plus, avec un peu d'entraînement et de

concentration, notre système visuel peut réviser cette décision à de multiples reprises en l'espace de quelques secondes.

Dans les situations écologiques, il semble que la plupart des décisions perceptives soient implicitement et automatiquement traitées dans des systèmes perceptifs plutôt que de passer par des processus délibératifs explicites. C'est par exemple le cas dans la situation décrite en introduction, où un individu en forêt doit rapidement (i.e., avant même l'accès à la conscience, voir le **chapitre 7**) décider s'il aperçoit un visage, potentiellement indicateur de la présence d'un prédateur ou d'un ennemi, de façon à pouvoir rapidement détalé.

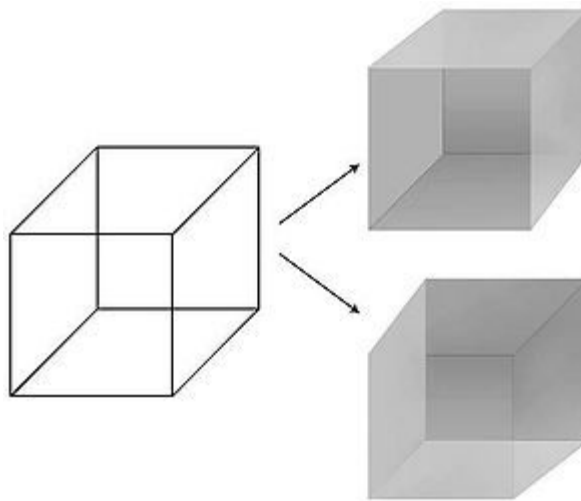


Figure 1. Illusion dite du « cube de Necker » illustrant le besoin de prendre une décision perceptuelle lors d'une situation ambiguë. La face située en haut à droite du cube de gauche peut être interprétée comme étant (1) soit la face du dessous d'un cube en 3D (image du haut), (2) soit la face du dessus (image du bas). En se concentrant sur le cube de gauche, on peut alterner entre ces deux percepts. Mais on ne peut pas les voir tous les deux en même temps car ce sont deux interprétations incompatibles du monde qui nous entoure. Le cerveau doit donc « décider » à un instant donné quelle configuration lui « convient », dans le sens où elle serait compatible avec son interprétation du monde, et donc quelle configuration il perçoit. Cette décision peut être influencée par le contexte : si je dessine une table sous le cube, selon la perspective avec laquelle j'ai représenté la table, il n'y aura plus qu'un seul des deux cubes en 3D qui sera compatible avec le reste de l'image. On ne parvient alors plus à alterner entre les deux percepts du cube.

2. Décisions : de la valeur à l'action

Concentrons-nous maintenant sur les décisions dites « économiques » car elles impliquent un processus d'évaluation de chaque option avant d'en choisir une.

a. Le concept de valeur subjective

Philosophie

Le concept de *valeur subjective* a été appréhendé pendant plus de 2000 ans et défini par trois domaines disciplinaires principaux : la philosophie avec Aristote, l'économie avec Bernoulli et Kahneman entre autres, et la psychologie de l'apprentissage par renforcement avec Thorndike, Pavlov, Rescorla et Wagner.

En philosophie classique, les valeurs subjectives ont été étudiées au travers du prisme de la notion de plaisir. La question centrale était : « Quelle est la nature du plaisir ? ». Le plaisir doit être distingué du bonheur, qui est considéré comme un bien-être stable et durable (Aristote, *Éthique à Nicomaque*). Le plaisir peut être vu comme une sensation d'un instant, mais aussi comme une expérience, comme un sentiment que nous cherchons à atteindre, ou même en tant que « force motrice » (Sidgwick, 1907). Pour dire les choses de façon courte et simple, on peut considérer que lorsque les philosophes discutent de ce que nous appelons la *valeur subjective*, ils parlent d'une quantité de plaisir, avec l'idée que le plaisir est précieux, et qu'on peut décider d'effectuer des actions pour l'atteindre. Chercher à obtenir du plaisir, ou chercher à éviter la douleur, devient alors un désir (que nous appellerions un « but »). John Stuart Mill a proposé que le désir d'un objet ne soit causé que par le plaisir, ou du moins par l'idée d'éprouver du plaisir (Mill, 1901). Cette dernière proposition est en accord avec les théories psychologiques hédonistes, qui supposent que le plaisir et la douleur sont les seules choses importantes qui nous poussent à agir.

Bien sûr, il faut considérer que certaines choses nous procurent du plaisir alors qu'elles ne sont pas spécialement bonnes pour nous, ou que leur excès est mauvais pour nous. La valeur subjective d'une option peut donc aussi refléter une évaluation délibérée des conséquences à long-terme de l'exécution de l'action. Nous y reviendrons. Néanmoins, les théories psychologiques considèrent en première approximation que les sensations de plaisir et de douleur sont le résultat d'un processus évolutif qui leur permet de refléter à peu près leur caractère bon ou mauvais pour nous.

L'aspect unitaire du plaisir est également central dans la tentative pour l'expliquer. En d'autres termes, savoir si le plaisir est un genre unique de sentiment avec différentes causes a fait l'objet de nombreux débats, puisque son lien avec la conscience, l'attention et la motivation sont complexes et qu'il peut être divisé en plusieurs types, comme aimer quelque chose ou être conscient d'aimer quelque chose.

La notion de désir est également sujet à controverse. Aristote y distingue deux composantes : l'appétit et la volition. La première étant originaire du corps et peut être perçue comme un besoin primaire tandis que la seconde est plus rationnelle et impliquée dans l'engagement dans une action orientée vers un désir. Spinoza considère aussi que l'appétit et la volition reflètent différents degrés de délibération derrière le désir, mais tous deux peuvent émaner du corps comme de l'esprit, ces deux

derniers étant deux facettes d'une seule et même substance². La nature du plaisir et du désir sont des sujets toujours débattus en philosophie. Cependant, la question de la « quantité », de l'« intensité » ou de la « force » du plaisir et du désir est rarement évaluée dans ce domaine alors qu'elle l'est précisément en neurosciences cognitives.

Nous définirons donc ici la *valeur subjective* comme une représentation interne de la quantité de plaisir qu'un objet, une action ou un état pourrait procurer et qui est capable de guider le comportement à la fois dans sa direction (faire un choix) et dans son intensité (vigueur ou rapidité d'exécution de l'action). L'aspect « directionnel » a été largement étudié en économie, que nous décrirons brièvement dans la prochaine section.

Économie

L'étude de la prise de décision a été dominée très longtemps par l'approche économique et le premier postulat de ce domaine est que l'humain devrait être rationnel, et que ses décisions devraient ainsi servir à optimiser une notion d'*utilité*³.

Cependant, il existe de nombreuses violations de la rationalité (que nous verrons plus loin dans ce chapitre) et Daniel Bernoulli (1700-1782) a été le premier à décrire formellement pourquoi l'utilité peut être définie comme une valeur subjective - i.e., comme une valeur spécifique à chaque individu et / ou contexte -, et non comme une valeur absolue. En effet, (Bernoulli, 1954) distingue la valeur objective d'une marchandise - comme son prix - et la valeur subjective de celui-ci - son utilité -, avec l'idée que le prix est égal pour tous et que l'utilité dépend de la personne et des circonstances. Pour lui, « il ne fait aucun doute qu'un gain de mille ducats est plus important pour un pauvre que pour un homme riche, bien que les deux gagneraient le même montant ». À partir de ce constat, il a proposé la première théorie pour expliquer le concept d'utilité marginale décroissante (même s'il n'a pas introduit le terme lui-même) en formalisant que « l'utilité résultant d'une petite augmentation de richesse sera inversement proportionnelle à la quantité de biens précédemment possédés ». Dit autrement, l'utilité est marginale parce qu'elle est relative à la quantité de biens déjà possédés, et elle est décroissante car plus l'individu aura de richesses, moins un billet de 500€ aura de valeur pour lui. Ce formalisme introduit le concept de subjectivité dans la valeur, ou utilité, qui est utilisée lors d'un choix. En 1879, William

² Spinoza rejette la dissociation entre l'âme et le corps, contrairement à la vision dualiste rendue classique par Descartes.

³ En Économie, l'utilité représente une espérance mathématique de gain. Elle se calcule en multipliant la probabilité d'obtenir un certain gain monétaire (e.g., probabilité de 0.8, qui représente 80% de chances d'obtenir ce gain) par la valeur de ce gain (e.g., 1000€ ; ce qui donne une utilité de 800). Voir le **chapitre 11** pour la façon de mettre ce calcul en équation.

Stanley Jevons développe la théorie de l'utilité, et la voit comme la mesure quantitative du plaisir : « expression numérique des quantités du sentiment » (Jevons, 1879).

Cependant, une difficulté centrale dans l'étude de la prise de décision est que l'utilité ne peut pas être mesurée directement. En effet, elle doit être déduite du comportement, par exemple par la quantité d'effort qu'un individu est prêt à fournir, ou par sa propension à payer un certain montant pour obtenir quelque chose. Ce type de mesure permet d'accéder à une représentation cardinale de l'utilité (donc une représentation quantitative). L'utilité peut également être mesurée grâce à des choix entre plusieurs options, permettant de déduire une représentation ordinale de l'utilité d'une option (i.e., quel est mon ordre de préférence parmi plusieurs options ?). La plupart des théories économiques sur l'utilité ont été développées en mesurant l'utilité par des expériences où les individus sont confrontés à des choix à faire entre plusieurs gains monétaires possibles ayant différentes probabilités (ce qu'on appelle des tâches de *loteries*, voir le **chapitre 10** pour des exemples).

Apprentissage par renforcement

Aux XIX^{ème} et XX^{ème} siècles, alors que les philosophes et les économistes se concentraient sur le plaisir et sa mesure au travers des choix, certains scientifiques et psychologues étudiaient le comportement pour tenter de décrire les facteurs qui le motivent et ceux qui pourraient expliquer sa répétition. Ils ont rapidement convergé vers le concept de *récompense*, i.e. un objet associé à une valeur positive.

Le concept de récompense est au cœur de l'étude physiologique des valeurs. Alors que la récompense est un objet ou une situation, le plaisir associé peut être vu comme la valeur subjective définie par les économistes et les philosophes. Pour les psychologues, la récompense est ce qui nous permet d'apprendre. Thorndike (1874-1949) a décrit la théorie de l'apprentissage comme l'association entre « un certain acte avec une certaine situation et le plaisir qui en résulte » (Thorndike, 1898). À l'aide de boîtes et d'animaux (chats, chiens et poussins), il a montré que le temps mis par les animaux pour s'échapper des boîtes diminuait avec le nombre d'essais jusqu'à atteindre un plateau et un temps de réponse court. Cette forme d'apprentissage a été appelée *apprentissage par essais et erreurs* (ou *apprentissage par renforcement*) et est *instrumental*, ce qui signifie qu'il s'agit d'une association entre une action et un résultat. À partir de cette expérience, Thorndike a défini la **loi de l'effet** (1905), qui stipule que les réponses comportementales qui produisent un effet positif dans une situation particulière deviennent plus susceptibles de se reproduire (elles acquièrent donc une valeur subjective plus grande), tandis que les réponses qui produisent un effet négatif sont moins susceptibles de se répéter (leur valeur subjective diminue). Avec cet exemple, on voit donc que **la valeur subjective sous-tendant nos décisions n'est pas nécessairement une valeur fixe**, mais peut être augmentée ou diminuée au fur et à

mesure de notre expérience d'interaction avec l'environnement, et donc être le fruit d'un apprentissage.

En parallèle, Pavlov (1849-1936) a mené son expérience bien connue sur des chiens et formalisé le concept de *conditionnement classique* (ou conditionnement pavlovien) (Pavlov, 1927). Pavlov a présenté un stimulus (e.g., le son d'une cloche) puis a donné de la nourriture à un chien. Après plusieurs répétitions, le chien a commencé à saliver dès la présentation du son de cloche, donc en réponse au stimulus, alors qu'aux premiers essais, le chien commençait à saliver uniquement en réponse à la nourriture. Pavlov a appelé ce phénomène le « conditionnement ». Le stimulus conditionnel (le son de la cloche) a été associé à la nourriture (réponse inconditionnée), ce qui signifie que le stimulus conditionné a acquis une valeur prédictive de la nourriture. Alors que les philosophes et les économistes semblaient considérer la question du plaisir *anticipé* et du plaisir *vécu* comme un concept unique, les physiologistes ont étudié le lien entre eux et ont montré comment l'un pouvait influencer l'autre par l'apprentissage.

À la suite des travaux de Pavlov, Rescorla et Wagner ont développé le modèle Rescorla-Wagner au début des années 1970 pour décrire formellement les circonstances dans lesquelles le conditionnement pavlovien se produit (Rescorla, Wagner, & others, 1972). Ils ont introduit l'idée que l'apprentissage vient de la différence entre ce qui est attendu et ce qui est réellement vécu, donc d'une forme d'*erreur de prédiction* (le **chapitre 10** du présent ouvrage décrit comment calculer cette erreur, et comment elle évolue au cours de l'apprentissage). Le concept central est que la quantité de surprise qu'un organisme éprouve face à un stimulus inconditionné (US)⁴ est supposée dépendre de la valeur associative totale que l'individu s'attend à observer dans l'environnement à partir de la somme de tous les signaux présents pendant cet événement. Par exemple, si un stimulus A prédit l'apparition de deux pommes, et un stimulus B l'apparition de trois pommes, alors la valeur associative des deux stimuli A+B sera de cinq pommes, et l'individu éprouvera une grande surprise si l'US ne consiste qu'en une seule pomme.

Grâce à cette notion de valeur associative, des travaux plus récents en psychologie ont pu également montrer que plus un individu s'attend à recevoir une grande quantité de récompense, plus il mettra de vigueur dans son action pour l'obtenir. Ainsi, les valeurs subjectives peuvent « énergiser » le comportement et l'orienter vers des indices ou des objets dans l'environnement. Comme nous allons le voir plus loin, elles sont également essentielles pour l'apprentissage.

Mesure des valeurs subjectives en neurosciences cognitives

⁴ Un stimulus inconditionné, ou inconditionnel, par exemple une récompense, se distingue d'un stimulus conditionné, ce dernier ayant été associé à la récompense par conditionnement.

Quatre grands types de mesures des valeurs subjectives existent en neurosciences cognitives.

Le premier type est ce que nous appelons une tâche de notation (ou d'évaluation) (Maël Lebreton, Jorge, Michel, Thirion, & Pessiglione, 2009). Ce type de tâche utilise les déclarations subjectives des individus. Un élément ou une situation est présenté(e) à un participant et il lui est demandé de noter (c'est-à-dire d'évaluer) à l'aide d'une échelle à quel point il lui serait agréable d'obtenir l'élément. Ce type de mesure est particulièrement utile pour connaître la valeur subjective de différentes options pour un individu donné, et ensuite vérifier que ses préférences respectent les différences de valeur entre les options lors d'une tâche de choix parmi ces options.

Les deuxième et troisième types de mesures utilisent la notion de coût. La « propension à payer » (« Willingness to pay » - (Varian, 1992; Plassmann, O'Doherty, & Rangel, 2007)) est mesurée dans une tâche d'enchère (Becker - DeGroot - Marschak - BDM, (Becker, DeGroot, & Marschak, 1964) qui demande aux participants d'indiquer sur une échelle combien ils seraient prêts à payer pour obtenir un objet présenté. Ce coût peut prendre la forme d'un effort pour obtenir une récompense, comme dans les « tâches d'effort » (par exemple, presser avec force pendant plusieurs secondes une pince afin d'obtenir un gain monétaire). Ce type de tâches se décline sous différentes formes, comme par exemple demander aux participants de produire un effort pour déterminer la probabilité de recevoir une récompense (M. Pessiglione et al., 2007; Cléry-Melin et al., 2011), ou leur demander s'ils sont prêts à produire une quantité X d'effort pour obtenir une récompense Y (Bonnelle et al., 2015). Ces mesures peuvent être considérées comme indirectes car on suppose que les participants minimiseront leurs coûts pour obtenir une récompense, c'est-à-dire qu'ils essaieront de produire la quantité d'effort suffisante pour obtenir la récompense tout en évitant les efforts supplémentaires qui seraient superflus. La valeur subjective des objets peut donc être déduite du comportement observé (i.e., combien d'effort les sujets effectuent en pratique pour obtenir la récompense).

Le quatrième type de mesure est une des plus répandues : les tâches de choix binaires, qui imposent aux participants de faire des choix entre deux options. Contrairement aux autres, cette mesure n'est pas cardinale, mais ordinale : la valeur de chaque option peut être déduite de son rang, ce qui signifie que si A est choisi contre B et B contre C, on peut supposer que A a une valeur supérieure à B, et B une valeur supérieure à C. Mais la valeur dépendra par construction de l'échantillon d'objets proposés. Cependant, mesurer les temps de réponses de ce type de choix peut indiquer si les deux valeurs sont proches ou non (un choix entre des items avec des valeurs similaires sera plus difficile et prendra donc plus de temps). De plus, la notion de distance entre deux items peut également être évaluée en examinant la proportion de choix pour une option donnée (fréquence de choix de A par rapport à B lorsque le choix est répété). Ce type de mesure est différent des autres tâches car elle concerne la valeur de décision

(différence entre les valeurs anticipées) tandis que les autres concernent la valeur anticipée ou les coûts.

Les mesures cardinales des valeurs subjectives sont étroitement liées à la gamme des échelles proposées aux sujets. De plus, l'utilisation des échelles de notation peut différer selon les sujets. Il est donc important de mentionner ici que même si nous nous référons au mot « cardinal » pour ces mesures, il s'agit toujours de valeurs subjectives rapportées relativement à une échelle donnée.

b. Valeurs et choix

Les étapes d'un choix

Une théorie unificatrice de la prise de décision basée sur la valeur a été proposée par Rangel, Camerer et Montague (Rangel, Camerer, & Montague, 2008). Elle a l'avantage d'unifier les choix étudiés par les économistes (que nous qualifierons de « comportements dirigés vers un but ») et les processus évalués par la théorie de l'apprentissage par renforcement. Cinq étapes principales y sont décrites :

- **Représentation**, y compris la représentation de l'environnement (états externes), du soi (états internes) et des options disponibles ;
- **Évaluation**, qui correspond à l'attribution d'une valeur subjective à chaque action en jeu, compte tenu des états internes et externes, qui peuvent aussi avoir une valeur. Notons qu'il existe plusieurs types de valeurs, comme la valeur d'action ou la valeur de récompense.
- **Sélection de l'action** qui correspond à la décision elle-même. Elle est basée sur l'étape précédente : les valeurs calculées sont comparées et la meilleure option est sélectionnée, ce qui entraîne la sélection et l'exécution d'une action. Plusieurs systèmes d'évaluation peuvent entrer en conflit à ce stade (nous y reviendrons).
- **Évaluation des conséquences de l'action**. C'est l'étape où le résultat de la décision est évalué et (parfois) expérimenté.
- **Apprentissage**, qui peut avoir un impact sur les trois premières étapes, puisqu'il permet de mettre à jour 1) la représentation des actions et des états (internes et externes) ; 2) la valeur de l'action grâce à **l'erreur de prédiction** (la valeur de la récompense attendue et la valeur de la récompense obtenue sont comparées ; voir **chapitre 10**) ; et enfin 3) les processus de sélection des actions.

Plusieurs systèmes d'évaluation

Les processus d'évaluation peuvent avoir lieu dans des types de comportements distincts. Nous décrivons ici trois systèmes principaux : le système pavlovien, le système habituel et le système dirigé vers les buts (Rangel et al., 2008).

Le système pavlovien

Le système dit pavlovien attribue des valeurs aux signaux environnementaux, qui peuvent être prédéterminés ou induits par des expositions répétitives (Pavlov, 1927). Cependant, il s'agit d'un petit ensemble de comportements qui sont généralement déclenchés par des stimuli appétitifs non conditionnés (Clark, Hollon, & Phillips, 2012). L'exemple du son de cloche décrit plus haut, qui par exposition répétée suivie de récompense acquière une valeur prédictive de la récompense, est typiquement un conditionnement pavlovien. Dans le **chapitre 10**, un modèle computationnel de ce processus vous sera présenté avec du code pour que vous puissiez le re-simuler vous-mêmes. L'important à retenir ici est que ce système n'est pas instrumental, contrairement aux deux autres systèmes. C'est-à-dire que l'apprentissage ne porte pas sur des actions mais sur des stimuli externes naturellement prédictifs d'une récompense, même lorsque l'agent reste passif. Du coup, dans des situations où nous faisons des choix entre des options à partir des valeurs instrumentales calculées en fonction de nos expériences passées par les deux autres systèmes (habituels et dirigés vers des buts, voir ci-dessous), il faut garder en tête que des petits bonus de valeurs peuvent être ajoutés aux stimuli liés aux options par le système pavlovien, et ainsi venir biaiser en partie nos décisions. C'est ce qu'on appelle le phénomène de Pavlovian-to-Instrumental Transfer (PIT), et qui est un des phénomènes psychologiques qui contribuent à biaiser certains de nos choix loin de la rationalité (Holland, 2004)(Holland, 2004).

Le système habituel

Contrairement au système pavlovien, le système habituel est instrumental : son apprentissage porte sur des actions. Il peut être appliqué à tout type de comportements dès lors qu'ils sont appris par des entraînements répétés. Ainsi, ce système va attribuer une valeur à une option par l'association d'un stimulus à une récompense (ou résultat d'action) par l'expérience d'essais et d'erreurs. Ce système repose fondamentalement sur le lien entre un stimulus et une action débouchant sur une récompense et est directement impliqué dans l'apprentissage (Balleine & Dickinson, 1998). Le point important à retenir sur ce système est que cet apprentissage dit « procédural » est implicite et automatique. On n'a pas besoin d'en avoir conscience pour qu'il ait lieu. Par conséquent, dans de nombreuses situations qui nous semblent familières, par fainéantise ou par souci d'économie de temps ou pour d'autres raisons que nous discuterons après, nos décisions peuvent émaner de notre système habituel, plus rapide car il évite les longues phases de délibération nécessaires au système dirigé vers des buts. Toutefois, le fait de prendre des décisions rapides nous rend davantage sujets de ce que Kahneman appelle nos « biais cognitifs », et qui peuvent rendre nos décisions apparemment irrationnelles du fait qu'on n'a pas pris le temps de la délibération (Kahneman, 2011).

Le système dirigé vers des buts

Contrairement au système habituel, le système dirigé vers des buts attribue des valeurs aux actions et crée des associations entre les actions et les résultats escomptés de celles-ci, ce qu'on appelle la *contingence* entre l'action et sa conséquence (Balleine & Dickinson, 1998). Ainsi, tout comportement reposant sur ce système serait normalement dirigé vers une récompense ou un « objectif » spécifique, i.e. un but. La différence fondamentale avec le système habituel est que dans ce dernier, l'action est déclenchée par le stimulus ou le contexte sans représentation des conséquences possibles de l'action. C'est simplement l'action qui a acquis la plus forte valeur subjective dans le contexte en cours qui est sélectionnée. À l'inverse, le système dirigé vers des buts déduit la valeur subjective des actions de l'estimation mentale de leurs conséquences futures. Ce processus d'estimation (ou *inférence*) requière une phase de délibération potentiellement longue s'il existe de nombreuses conséquences possibles et si ces conséquences ne sont pas immédiates mais à long-terme. Par exemple, lorsque l'on joue aux échecs, on peut soit jouer rapidement à partir d'une impression (ou *intuition*) de ce qui semble le meilleur coup dans une situation donnée (système habituel pour quelqu'un qui joue fréquemment aux échecs), ou bien on peut calculer plusieurs coups à l'avance, en estimant à chaque fois quelle va être la nouvelle situation et quel coup l'adversaire va probablement faire en réponse (système dirigé vers des buts).

Un tel système est donc lent à prendre une décision (à délibérer), comparé au système habituel qui est très réactif. Néanmoins, son avantage est qu'il permet de réagir de manière flexible à un changement de l'environnement ou de l'état interne (tel que la satiété). En effet, si une action ne donne plus la même conséquence dans l'environnement, ou si au contraire elle donne toujours de la nourriture alors que le sujet n'a plus faim, alors il suffit de mettre à jour la valeur subjective de l'action en fonction de l'estimation de ses conséquences futures (en termes de récompense) pour se rendre compte qu'elle n'est plus intéressante à exécuter. À l'inverse, le système habituel nécessite de nombreuses répétitions suivies d'une absence de récompense pour baisser la valeur subjective d'une action, symétriquement aux répétitions récompensées qui ont été nécessaires à son apprentissage initial (comme le montrent les simulations de modèles computationnels présentées au **chapitre 10**).

Nous verrons plus loin dans ce chapitre que l'existence de ces différents systèmes de décision a des conséquences importantes, puisqu'un même sujet ne prendra pas toujours la même décision dans le même contexte, en fonction du système sur lequel sa décision repose. La psychologie cognitive a donc mis en place une série de mesures (comme le temps de décision) pour pouvoir identifier sur quel système reposent les décisions d'un sujet à un moment donné. Les neurosciences peuvent également apporter leur contribution en montrant quelles aires du cerveau s'activent à un moment donné, certaines aires ou réseaux cérébraux étant davantage impliqués dans un système que dans les autres.

c. Expliquer les choix à partir des valeurs subjectives

Les valeurs subjectives, ou valeurs de récompense, sont aujourd'hui considérées comme nécessaires pour prendre une décision dirigée vers un but. En effet, si un individu doit faire un choix entre un item A et un item B, celui-ci doit estimer les valeurs de A et B, chaque valeur étant affectée par différents attributs comme l'immédiateté de la récompense ou son impact sur la santé. Ensuite, les valeurs doivent être comparées afin de pouvoir sélectionner l'item qui a la valeur la plus élevée (règle « *argmax* » : choix du maximum). Cependant, les choix observés violent souvent le principe du choix de la valeur subjective la plus élevée. Ainsi, la manière dont une décision est prise est une problématique qui a suscité un intérêt croissant au cours des dernières décennies, et plusieurs types de modèles ont été développés pour proposer des mécanismes cognitifs et cérébraux sous-jacents à la prise de décision. Deux types de modèles ont été largement utilisés dans le domaine. Le premier type correspond à des « modèles statiques », qui proposent des mécanismes pour expliquer les choix. Le second type concerne des modèles qui expliquent à la fois la décision et le temps nécessaire à la prise de la décision : il s'agit de « modèles dynamiques ». Ces deux types de modèles ont leurs origines dans la psychophysique et ils ont à l'origine été développés pour expliquer les décisions perceptuelles. Mais ils ont été récemment étendus à la prise de décision basée sur des valeurs.

Nous allons ici succinctement décrire les différences de mécanismes entre ces différents types de modèles, de façon à pouvoir illustrer comment relier ces mécanismes à des activités cérébrales mesurées en imagerie. Pour aller plus loin, le **chapitre 10** propose une introduction à la modélisation computationnelle, avec une explication de quelques-unes des équations des modèles discutés ici.

Modèles statiques de prise de décision

Modèles descriptifs

Comme nous l'avons vu, les choix ne peuvent souvent pas être expliqués ni modélisés par une simple règle déterministe « *argmax* », puisque tous les facteurs tels que le contexte ne peuvent être pris en compte dans le calcul de la valeur subjective. Une solution consiste à utiliser des méthodes probabilistes pour les expliquer en ajustant une fonction sigmoïdale telle que la fonction *softmax*, déjà utilisée dans les théories d'apprentissage (Luce, 1959) (voir **Figure 2**). Ces méthodes permettent de définir des fonctions psychométriques qui capturent le fait que le nombre d'incohérences dans les choix augmente lorsque les valeurs sont proches (Snodgrass, Townsend, & Ashby, 1985) : il est plus difficile de choisir entre deux objets de même valeur (ou de valeur proche) qu'entre deux objets ayant des valeurs très différentes. Les versions les plus simples de ces modèles ont deux paramètres : une constante, qui indique s'il y a un biais dans les choix (pour l'une ou l'autre des options) et une *température inverse* généralement appelée β , qui capture le bruit dans les choix. Cette fonction n'apporte

aucune indication sur les mécanismes cognitifs ou cérébraux sous-jacents aux décisions, et c'est pourquoi des modèles plus mécanistiques (i.e., explicatifs en termes de mécanismes) ont été développés.

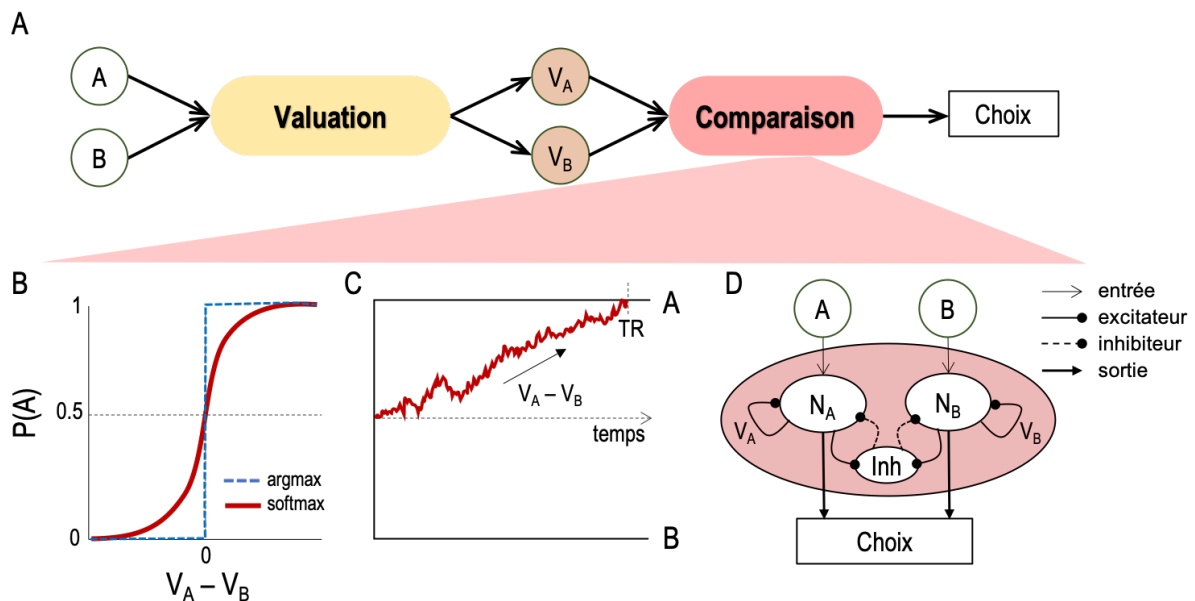


Figure 2. Représentation graphique des étapes d'une décision basée sur les valeurs et modèles de choix. A. Deux options A et B ont une valeur V_A et V_B respectivement. Ces deux valeurs sont comparées afin de choisir une des deux options. Les choix peuvent se modéliser par des modèles descriptifs (B) tels que le softmax, dans lequel la probabilité de choisir une option est déterminée par la différence de valeur entre les deux options, avec du bruit lorsque les valeurs sont proches ; par des modèles dynamiques (C) comme les modèles d'accumulation, dans lesquels de l'évidence s'accumule en fonction du temps et proportionnellement à la différence de valeurs entre les deux options. Un choix est fait avec un temps de réponse TR lorsqu'une borne (A ou B) est franchie ; ou par des modèles biologiques (D) dans lesquels deux ensembles de neurones excitateurs N_A et N_B reçoivent en entrée chacune des deux options, et dont l'activité auto-excitatrice est proportionnelle à la valeur de chaque option. Un ensemble de neurones inhibiteurs intermédiaire (Inh) reçoit de chaque ensemble des afférences excitatrices et inhibe à son tour chaque autre ensemble. Le choix est fait lorsqu'un des deux ensembles N_A ou N_B est complètement inhibé. Figure adaptée de (Ratcliff, 1978) (C), (P. W. Glimcher & Fehr, 2013) (D).

Théorie de la détection du signal (TDS)

Le but initial de la théorie de la détection du signal était d'expliquer les décisions sensorielles avec incertitude (Peterson, Birdsall, & Fox, 1954; Tanner Jr & Swets, 1954). Elle suppose qu'il y a quatre types de réponses possibles lorsque l'on doit détecter la présence d'un stimulus : détection correcte (répondre « oui » lorsque le stimulus est présent), erreur d'omission (répondre « non » lorsque le stimulus est présent), fausse alarme (répondre « oui » quand le stimulus est absent), et rejet correct (répondre

« non » lorsque le stimulus est absent). À partir de la proportion de ces types de réponses, on peut calculer la sensibilité d' (ou discriminabilité) au stimulus, qui peut être vue comme un pur paramètre perceptif, et un critère β , parfois appelé « biais de réponse », qui peut indiquer si le sujet est « conservateur » (tendance à dire « non » : pour éviter les fausses alarmes, avec le risque de manquer une certaine détection) ou « libéral » (tendance à dire « oui » : pour éviter les omissions, avec le risque d'augmenter les fausses alarmes) (Figure 3).

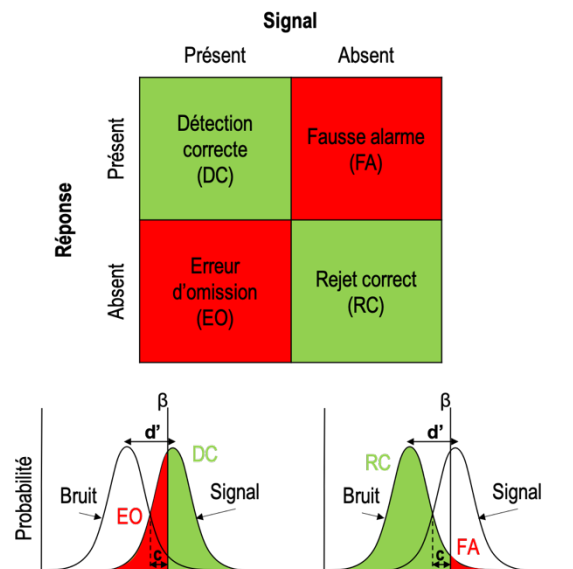


Figure 3. Théorie de la détection du signal. Quatre types de réponses sont possibles pour la détection d'un stimulus selon la présence ou absence réelle du signal ou selon la réponse (signal perçu ou non perçu) : détection correcte, fausse alarme, erreur d'omission ou rejet correct. La théorie de la détection du signal représente la probabilité de chaque réponse à l'aide de deux paramètres principaux : la sensibilité d' et le critère β . En bas à droite : exemple où la sensibilité moyenne est un critère positif ($c > 0$) : Si le signal est présent (gauche), il y aura beaucoup d'erreurs d'omission ; si le signal est absent (droite), il y aura peu de fausses alarmes. Adaptée de (Anderson, 2015).

Ce modèle a été largement utilisé dans le domaine de la psychophysique. Cependant, si on considère l'idée que la valeur subjective est une distribution interne des probabilités, alors quand on doit comparer deux valeurs, plus elles sont proches et moins elles seront discriminables (d' est petit, donc le pourcentage d'erreurs augmente). Dans ces cas, les omissions et les fausses alarmes ne sont qu'un type d'erreur : ne pas choisir l'option avec la valeur la plus élevée. Le biais de décision s'exprimerait (ou non) sur l'une ou l'autre option, selon le contexte ou les a priori. Nous pouvons considérer la théorie de la détection du signal (TDS) comme la base mathématique de la prise de décision perceptuelle. Cependant, comme elle a été développée pour expliquer les décisions liées à la détection de stimuli bruités, elle a ses limites. De plus, elle ne prend pas en compte la dynamique du processus de

décision. En effet, il est systématiquement observé que plus une décision est difficile (perceptuelle ou non), plus elle prendra de temps. La TDS a été étendue à des formes dynamiques compatibles avec la prise de décision basée sur la valeur.

Modèles dynamiques de prise de décision

La TDS a été étendue aux modèles dynamiques grâce à l'idée qu'à chaque point micro-temporel du processus de décision général, si le stimulus est présent, une information sera ajoutée à la précédente, et ainsi de suite. Des modèles d'échantillonnage séquentiel ont été développés sur cette idée d'accumulation d'évidence en faveur d'une ou l'autre des options disponibles.

Les modèles d'échantillonnage séquentiel supposent que l'information sur une option s'accumule au fil du temps jusqu'à ce qu'un seuil particulier soit atteint (**Figure 2C**). Le temps auquel ce seuil est atteint correspond au temps de réaction (RT) et le critère de décision atteint définit le choix qui est fait (s'il y a plusieurs bornes, comme nous le verrons plus loin). L'accumulation commence à un point de départ S et augmente avec un taux de dérive μ , correspondant à la quantité d'évidence disponible. Fondamentalement, ce processus est stochastique, ce qui rend l'accumulation bruitée, donc le processus d'accumulation variera selon les répétitions. Ensuite, ce sont généralement les probabilités de choix et les distributions de RT qui sont déduites de ce type de modèles dynamiques. Un temps de non-décision (Tnd) peut être introduit avant le début de l'accumulation, ce qui peut capturer le temps nécessaire par des processus non liés à la décision tels que l'identification du stimulus ou l'accès à des informations stockées en mémoire. Enfin, le rôle de ces paramètres peut être vu comme suit : une augmentation de la difficulté diminuerait le taux de dérive μ ; une augmentation de la pression temporelle (ou compromis vitesse-précision) diminuerait le critère de décision A (et B) ; un a priori sur l'évidence modifierait la valeur du point de départ S (comme les croyances) ; et enfin, un stimulus difficile à catégoriser augmenterait le temps de non-décision.

Plusieurs versions de modèles d'échantillonnage séquentiel ont été développées en neurosciences cognitives, le principal étant le *Drift Diffusion Model*. Le principe de base de ce modèle est que les choix découlent d'un processus d'accumulation déterminé par la différence entre les valeurs d'options A et B , avec un taux $\mu = \Delta V$ ($\Delta V = V_A - V_B$ est la différence entre les valeurs), (Ratcliff, 1978; Ratcliff & McKoon, 2008). Ce modèle ne peut être mis en œuvre qu'avec deux options alternatives (A et B) et a été largement appliqué pour étudier les choix lors de prise de décision perceptuelle et mnésique (Kinchla & Smyzer, 1967; Link & Heath, 1975; Pearson, Hayden, Raghavachari, & Platt, 2009).

3. Bases neurales de la prise de décision

En neurosciences cognitives, valeur et récompense sont deux concepts étroitement liés. Les premières études des bases neurales de la prise de décision étudiaient principalement la réponse du cerveau à l'exposition à des récompenses.

a. Représentation des valeurs de récompense dans le cerveau

Cortex orbitofrontal (OFC)

Les premières études mentionnant un lien entre récompense et cerveau sont des études d'auto-stimulation (Olds & Milner, 1954) et des études de lésions (Butter, McDonald, & Snyder, 1969; Mora, Avrith, Phillips, & Rolls, 1979; Rolls, Burton, & Mora, 1980) chez les rongeurs et les singes. Les études d'auto-stimulation consistent à implanter des électrodes dans le cerveau de rongeurs ou de singes, et à l'aide d'un petit levier, les rongeurs peuvent activer électriquement l'électrode et stimuler une région de leur cerveau eux-mêmes. Or les premières études d'auto-stimulation utilisaient des électrodes qui venaient activer électriquement un réseau de fibres provenant du tronc cérébral, le medial forebrain bundle. Ce dernier constitue un canal important de transmission d'un neuromodulateur associé au plaisir, la dopamine, vers le cortex préfrontal et les ganglions de la base. Les neurobiologistes considèrent donc que ces fibres participent à induire la sensation de plaisir liée à l'obtention de la récompense.

Des études plus récentes ont révélé qu'une sous-partie bien particulière du cortex préfrontal semblait être impliquée dans l'évaluation des récompenses : le cortex orbitofrontal (OFC). Par exemple, le taux d'auto-stimulation de l'OFC chez les singes augmentait davantage lorsque le singe avait faim que lorsqu'il était rassasié (Mora et al., 1979). Cela illustre l'idée d'un codage de valeur subjective par rapport à un codage de valeur objective : la nourriture a une valeur subjective plus forte lorsque nous avons faim que lorsque nous sommes rassasiés. D'autre part, les études sur les lésions de l'OFC ont révélé des comportements de persévération (Iversen & Mishkin, 1970) et une augmentation des comportements d'approche des produits non alimentaires chez les singes (Butter et al., 1969). En d'autres termes, lorsque les singes étaient face à des objets sans aucune valeur particulière pour eux, ils le portaient à la bouche, comme si cet objet avait un intérêt alimentaire. Ces résultats ont été initialement interprétés comme une absence d'inhibition induite par les lésions OFC. Cependant, les études suivantes montreront que c'est l'association entre le stimulus et la récompense qui est altérée et qui conduit à une réponse à des stimuli et des items non récompensant (Jones & Mishkin, 1972). En d'autres termes, si la lésion de l'OFC brouille l'information sur l'association entre le stimulus et la récompense, le système dirigé vers des buts ne peut plus estimer quel stimulus a des chances d'être suivi de récompense, laissant place à un choix au hasard parmi les stimuli présentés. Plus tard, les premiers enregistrements électrophysiologiques de neurones dans l'OFC ont montré que les neurones répondaient à la suppression de nourriture (Mora et al., 1979) et étaient supposés coder la disponibilité de la récompense (Rosenkilde, Bauer, & Fuster, 1981), puisque leur

taux de décharge augmentait avec la présence d'une récompense obtenue grâce à une réponse correcte dans une tâche de mémoire à court terme. Au regard de ces résultats, l'OFC semble donc une structure cérébrale importante pour l'attribution d'une valeur à un stimulus ou un objet, alimentaire ou non.

Dopamine

Si l'OFC permet de mémoriser des associations entre stimulus et récompense, de quelle partie du cerveau lui provient l'information sur la valeur de récompense lui permettant de renforcer la valeur de ces associations ? Comme nous l'avons vu brièvement, un réseau de fibres provenant du tronc cérébral, le *medial forebrain bundle*, transmet de la dopamine jusque dans le cortex orbitofrontal. Or il s'avère que la dopamine ne signale pas simplement la présence de récompense, mais aussi sa quantité (Roesch, Calu, & Schoenbaum, 2007), sa probabilité d'occurrence (Fiorillo, Tobler, & Schultz, 2003), et la pertinence de sa mémorisation (W Schultz, Dayan, & Montague, 1997). En effet, lorsqu'une association entre un stimulus et une récompense est bien connue par l'animal, cela veut dire qu'elle est déjà bien mémorisée et qu'il n'y a pas besoin de la renforcer davantage au niveau neural. Il est donc pertinent d'envoyer des signaux de renforcement au cortex préfrontal uniquement lorsqu'une nouvelle récompense est obtenue. En d'autres termes, le niveau de surprise lié à une récompense inattendue est important à prendre en compte dans tout signal de renforcement.

Or justement, les neurones dopaminergiques – ceux qui émettent de la dopamine – du mésencéphale enregistrés dans l'aire tegmentale ventrale (VTA) ou dans la Substantia Nigra pars compacta (SNc) répondent spécifiquement selon un principe d'*erreur de prédiction de la récompense* (récompense attendue - récompense expérimentée (Wolfram Schultz, 1998) ; dont le modèle computationnel sera présenté au **chapitre 10**). L'expérience la plus connue concernant cette propriété est l'étude réalisée par Wolfram Schultz et son équipe (W Schultz et al., 1997). Ils ont enregistré des neurones dopaminergiques alors que des singes recevaient une récompense prédite ou non par un stimulus (tâche de conditionnement pavlovien). Ils ont montré que la réception de récompenses, lorsqu'elle n'était pas prévue, déclenchait une augmentation du taux de décharge des neurones dopaminergiques, ce qui correspondrait à une erreur de prédiction positive (la situation est meilleure que prévue). Ensuite, ils ont montré que lorsque la récompense pouvait être prédite grâce à la présentation du stimulus (résultat du conditionnement), les neurones augmentaient leur fréquence de décharge en réponse au stimulus prédisant la récompense, mais plus à la réception de la récompense. L'interprétation est que cette récompense est dorénavant parfaitement prédite. Il n'y a donc plus d'erreur de prédiction. Enfin, lorsque le signal arrive mais que la récompense n'est pas délivrée (omission ; donc situation moins bien de prévue), une diminution de la fréquence de décharge a été observée au moment où la récompense était attendue. Ces trois résultats ont démontré que l'erreur de prédiction de la récompense est une notion clé pour comprendre l'information représentée par

l'activité des neurones dopaminergiques. Pour une illustration graphique de l'activité dopaminergique enregistrée pendant cette étude, voir la figure 6 du **chapitre 10**.

Striatum

Il est important de noter que les signaux d'erreur de prédiction de la récompense transmis par la dopamine ne sont (en première approximation) pas directement impliqués dans les processus décisionnels (mais voir (Humphries, Khamassi, & Gurney, 2012) pour un raffinement de cette vision). Ils sont plutôt considérés comme une façon d'apprendre les valeurs de récompense associées à différents stimuli de l'environnement, de façon à pouvoir plus tard décider vers quel stimulus s'orienter pour obtenir de la récompense. Or pour bien comprendre les processus d'apprentissage qui sous-tendent les décisions en vue d'obtenir de la récompense, il nous manque encore deux ingrédients importants : (1) les signaux de prédiction (ou anticipation) de la récompense ; et (2) les signaux de valeur de récompense associée non plus à des stimuli, mais à des actions. Une aire cérébrale joue un rôle très important dans l'encodage de ces deux types de signaux : le striatum.

Le striatum est la principale structure d'entrée d'un ensemble de noyaux sous-corticaux appelés les ganglions de la base. Ces derniers sont organisés selon des boucles anatomiques qui relient le cortex, les ganglions de la base, le thalamus, et projettent en retour vers le cortex (Alexander & Crutcher, 1990). Or les ganglions de la base jouent un rôle essentiel dans l'organisation du comportement, de façon à sélectionner les actions qui conduisent à des récompenses (Redgrave, Prescott, & Gurney, 1999). De plus, le striatum est la partie du cerveau qui reçoit la plus grande partie des projections dopaminergiques du tronc cérébral, davantage encore que n'en reçoit le cortex orbitofrontal (Haber, 2014).

La partie ventrale du striatum reçoit des projections dopaminergiques de l'aire tegmentale ventrale et projette en retour vers les neurones dopaminergiques, ce qui en fait une région d'intérêt particulier pour les processus d'apprentissage basé sur la récompense. Or les neurones enregistrés dans le striatum ventral chez le singe pendant des tâches de conditionnement pavlovien montrent une activité qui anticipe la quantité de récompense que l'animal s'attend à recevoir (Cromwell & Schultz, 2003). De plus, lorsque des rats naviguant dans des labyrinthes reçoivent une séquence de récompenses sous forme de gouttes d'eau successives qu'ils ont bien appris à prédire, les neurones du striatum ventral répondent en anticipation de chaque récompense dans la séquence (Khamassi, Mulder, Tabuchi, Douchamps, & Wiener, 2008). Le striatum ventral est donc considéré comme une structure clé pour calculer les prédictions de récompense qui seront envoyées aux neurones dopaminergiques et leur permettre ainsi de calculer la fameuse *erreur de prédiction de la récompense* décrite plus haut. Les études IRMf chez l'humain ont confirmé ce rôle en étudiant les corrélations entre le signal cérébral enregistré et l'erreur de prédiction de récompense pendant une

variété de tâches. Ces études ont identifié un engagement du striatum ventral dans une grande diversité de paradigmes expérimentaux, comme des tâches d'effort, d'apprentissage ou de choix (Abler, Walter, Erk, Kammerer, & Spitzer, 2006; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Knutson, Fong, Adams, Varner, & Hommer, 2001; Lin, Adolphs, & Rangel, 2011; McClure, Berns, & Montague, 2003; Mathias Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Schmidt, Lebreton, Cléry-Melin, Daunizeau, & Pessiglione, 2012).

Enfin, un animal faisant l'expérience d'un conditionnement instrumental ne doit plus seulement associé de la récompense à des stimuli particuliers de l'environnement (comme c'est le cas dans le conditionnement pavlovien), mais également aux actions qui ont joué un rôle instrumental dans l'obtention de la récompense (comme presser un levier pour obtenir une boulette de nourriture). Un grand nombre d'expériences chez les rongeurs et chez les primates non-humains ont montré que l'activité neurophysiologique des régions dorsolatérale et dorsomédiale du striatum encodent la valeur associée aux actions (Samejima, Ueda, Doya, & Kimura, 2005; Thorn, Atallah, Howe, & Graybiel, 2010). Chez l'humain, l'utilisation de l'IRMf a permis de confirmer la dissociation entre le rôle du striatum dans l'apprentissage des valeurs des stimuli prédicteurs de récompense pendant un conditionnement pavlovien, et le rôle du striatum dans l'apprentissage des valeurs des actions pendant un conditionnement instrumental (O'Doherty et al., 2004).

On voit donc à partir de ces études, principalement neurophysiologiques chez l'animal ou l'humain en comportement, montrent qu'un ensemble de régions cérébrales sont impliquées dans le traitement de l'information liée à la récompense. Toutefois, il n'est pas encore clair comment la communication entre ces régions est organisée au sein de réseaux cérébraux pour permettre l'apprentissage, le traitement et la comparaison de valeurs de récompense sous-tendant la prise de décision.

Le système cérébral d'évaluation : un réseau de régions cérébrales

Alors que les études électrophysiologiques sur la valeur subjective et ses processus neuronaux ont l'avantage d'apporter des résultats précis sur la dynamique et sur l'activité neuronale au niveau local, les études IRMf ont l'avantage de pouvoir étudier l'activité neurale de plusieurs régions cérébrales qui forment des réseaux puisqu'elles permettent d'étudier la réponse BOLD dans tout le cerveau (**Figure 4**).

Un grand nombre d'études et de méta-analyses ont montré que le signal BOLD enregistré dans le cortex préfrontal ventro-médian (vmPFC) et dans le striatum ventral est corrélé aux valeurs subjectives d'options présentées dans des choix, d'items à évaluer ou sur lesquels enchérir (Maël Lebreton et al., 2009; Plassmann et al., 2007), mais aussi avec la valeur subjective de la récompense lorsqu'elle est reçue (O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). Bartra et al., (Bartra, McGuire, & Kable, 2013) ont mené une méta-analyse qualitative confirmant que le signal enregistré dans

le vmPFC et le striatum ventral, mais aussi dans le cortex cingulaire postérieur (PCC) est corrélé positivement avec les valeurs subjectives.

Cependant, le vmPFC semble réagir plus fortement à l'étape d'évaluation par rapport au moment de la réception d'une récompense. L'activité striatale ventrale présente la même corrélation positive pour les valeurs d'items à valence positive mais c'est l'inverse pour les pénalités, suggérant une relation quadratique (« en U ») entre la valeur et la réponse BOLD, tant au moment de l'anticipation que de la réception de la récompense.

Notons ici que dans les études chez l'humain, le vmPFC est beaucoup plus souvent rapporté comme région répondant à la valeur que l'OFC, alors que dans les études chez le singe, c'est l'inverse. Une étude récente utilisant de l'électrophysiologie intracérébrale chez l'humain a montré que ces deux régions répondaient de la même manière à la valeur, ce qui suggère que cette différence inter-espèce est probablement liée à la différence de méthodes d'enregistrement de l'activité cérébrale (IRMf chez l'humain et électrophysiologie chez le singe) (Lopez-Persem et al., 2020). L'OFC est une région voisine du vmPFC et selon les études, les labélisations peuvent varier, ainsi que leur distinction, parfois l'OFC comprend le vmPFC et parfois non. Par souci d'harmonisation, nous nous référons dans ce chapitre à l'OMPFC (orbito-medial prefrontal cortex) lorsque nous considérons l'OFC et le vmPFC ensemble (voir **Figure 4**).

En fonction des contextes et des dessins expérimentaux utilisés pour étudier le système cérébral d'évaluation, plusieurs partenaires du vmPFC et du striatum ventral ont été identifiés : le cortex cingulaire postérieur (PCC), l'hippocampe, le cortex préfrontal dorso-latéral (dlPFC) et le cortex cingulaire dorso-antérieur (dACC) pour les zones les plus communes impliquées dans le processus d'évaluation (**Figure 4**).

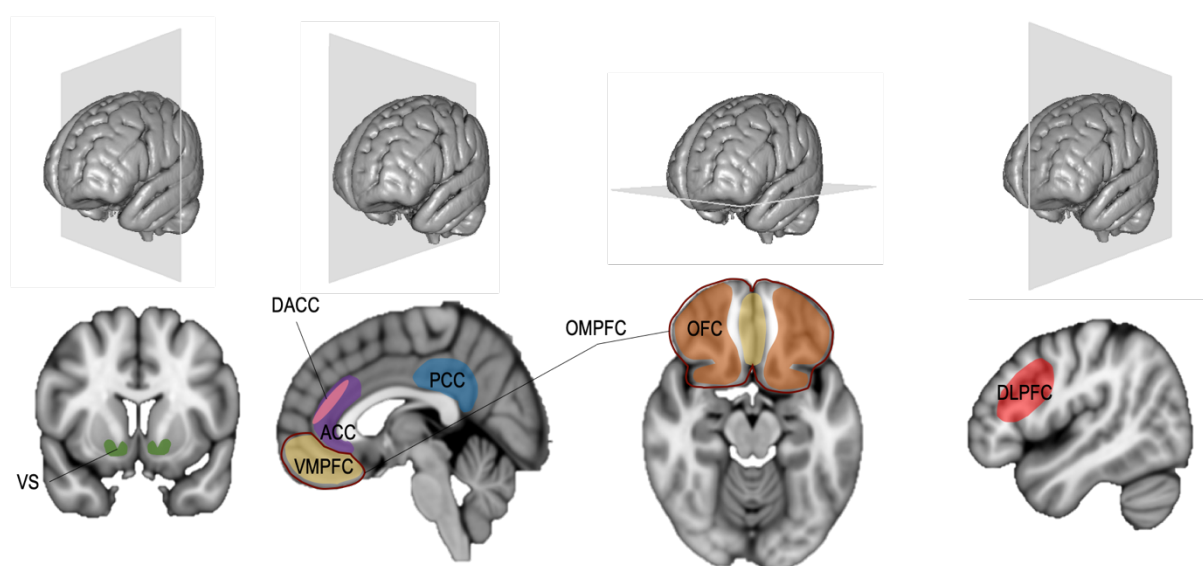


Figure 4. Régions cérébrales impliquées dans l'évaluation et la prise de décision. Première ligne : indication des plans de coupe pour représenter les régions cérébrales dans la deuxième ligne. VS : striatum ventral, VMPFC : cortex préfrontal ventro-médian, ACC : cortex cingulaire antérieur, dACC : partie dorsale de l'ACC, PCC : cortex cingulaire postérieur, OFC : cortex orbitofrontal, formant avec le VMPFC l'OMPFC : cortex préfrontal orbito-médian, DLPFC : cortex préfrontal dorso-latéral.

La zone cérébrale la plus souvent identifiée avec le vmPFC et le striatum ventral comme répondant à la valeur subjective et aux choix est le PCC (Bartra et al., 2013). Le PCC est une région centrale du réseau par défaut (DMN), ce qui signifie qu'il est systématiquement activé lorsque les sujets doivent se reposer dans le scanner. Il a été proposé que son activité est liée à la récupération de souvenirs autobiographiques, mais aussi à des fonctions attentionnelles (Leech & Sharp, 2013). Il est également connu pour être anatomiquement connecté avec le vmPFC via le faisceau de cingulum (Greicius, Supekar, Menon, & Dougherty, 2009). Ces résultats nous permettent d'émettre plusieurs hypothèses non exclusives sur l'implication de cette région dans le système cérébral d'évaluation : (1) son rôle pourrait être de récupérer des informations sur des items pour estimer leurs valeurs subjectives, et en déduire sur quels items préférés focaliser l'attention ; (2) une autre hypothèse serait que le PCC traite les signaux sensoriels pour reconnaître les objets, et que la connectivité entre le PCC et le vmPFC permettrait d'estimer la valeur subjective de ces objets (Polanía, Moisa, Opitz, Grueschow, & Ruff, 2015).

L'hippocampe est une structure cérébrale qui contribue au système cérébral d'évaluation. En effet, il a été démontré que l'hippocampe joue un rôle critique dans l'évaluation de récompenses imaginées (Barron, Dolan, & Behrens, 2013; MaÅ«l Lebreton et al., 2013; Peters & Büchel, 2010). Classiquement, l'hippocampe est connu pour son rôle dans le codage et la récupération de souvenirs épisodiques (Huijgen et Samson, 2015) (voir aussi le **chapitre 4**), mais il a été suggéré que l'hippocampe peut biaiser les décisions en créant des associations entre des éléments et des événements se produisant ensemble, puis en réactivant la représentation neuronale d'un événement A lorsqu'un événement B est présenté car il y était auparavant associé (Wimmer & Shohamy, 2012). Il a également été supposé que l'hippocampe biaise les décisions vers des éléments mémorisés en augmentant la connectivité entre l'hippocampe et le striatum ventral (Gluth, Sommer, Rieskamp, & Büchel, 2015). De manière plus générale, on suppose que le vmPFC et l'hippocampe interagissent lors de l'élaboration de valeurs qui nécessitent une récupération (donc une inférence) ou une délibération par anticipation des conséquences (par exemple par une simulation épisodique). Il en ressort que le vmPFC et l'hippocampe semblent des structures qui participent au système décisionnel dirigés vers des buts.

Propriétés de la représentation de la valeur subjective dans le système cérébral d'évaluation

Comme nous l'avons vu, il existe un grand nombre d'études soutenant l'idée que le vmPFC, l'OFC, le striatum ventral et le PCC, parfois avec l'hippocampe, forment le système d'évaluation du cerveau. Ce réseau a plusieurs propriétés intéressantes en termes de représentation de la valeur subjective (Lopez-Persem et al., 2020). En effet, il est générique, c'est-à-dire qu'il représente la valeur des items indépendamment de la catégorie à laquelle ils appartiennent : nourriture, tableaux, visages, argent, etc. ; il semble également être automatique : il répond à la valeur subjective même si les participants ne sont pas engagés dans un processus d'évaluation ou de décision basée sur la valeur, comme par exemple lorsqu'ils doivent estimer l'âge d'un visage ou d'un tableau. Enfin, son activité spontanée est sensible au contexte externe ou interne et influence la valeur subjective qui sera rapportée par un individu.

b. Modulation des valeurs et contrôle

Dans les modèles de prise de décision basé sur la valeur, les valeurs subjectives correspondent au résultat d'un apprentissage par renforcement (par essai-erreur, voir **chapitre 10**) sur lequel les choix sont ensuite basés (le plus souvent, l'individu choisira l'option ayant la plus forte valeur subjective). Pour refléter tous les aspects subjectifs liés aux options, ces valeurs doivent donc intégrer tous les attributs et fonctionnalités des options disponibles. Par exemple, la valeur attendue d'un comportement donné, dont l'avantage potentiel est très élevé mais très incertain (par exemple, l'achat d'un billet de loterie) réduit généralement ledit avantage par l'incertitude. Un raisonnement similaire s'applique, par exemple, au coût ou au retard : une option qui annonce une forte récompense, mais à un coût élevé, ou dont la consommation est considérablement décalée dans le temps verrait sa valeur actualisée par ledit coût et le délai d'obtention. La façon dont la probabilité, les coûts et les délais diminuent la valeur subjective ont fait l'objet de nombreuses recherches en économie, en psychologie et en neurosciences, mais est au-delà de la portée de ce chapitre. Pourtant, de manière critique, les signaux d'évaluation du vmPFC semblent être sensibles aux probabilités, aux coûts et aux délais, d'une manière qui reflète les préférences et les choix comportementaux. De plus, les signaux de valeurs observés dans le vmPFC semblent combiner des signaux intéroceptifs (internes) et perceptifs (externes) (Montague & King-Casas, 2007). Par exemple, le signal de valeur d'une option alimentaire dans le vmPFC est typiquement diminué en fonction de la satiété. Ces observations soutiennent l'idée que l'activité du vmPFC encode des signaux de valeur flexibles qui peuvent guider de manière adaptative les décisions dans diverses situations écologiques. D'un point de vue fonctionnel, cela signifie également que le vmPFC reçoit et intègre plusieurs informations provenant des zones sous-corticales et corticales qui contribuent au signal de valeur.

Toutefois, le comportement des individus, qu'il s'agisse d'humains ou d'autres mammifères, ne reflète pas toujours des valeurs subjectives qui intègrent bien tous les attributs possibles. D'une façon générale, on constate que lorsque les individus

prennent leurs décisions à partir d'un processus de délibération cherchant à estimer la meilleure façon d'atteindre un but souhaité (comportement dirigé vers un but), un réseau cérébral impliquant le cortex préfrontal contribue à intégrer de façon « rationnelle » tous les attributs d'une option pour en estimer la valeur subjective. À l'inverse, lorsque les individus prennent des décisions rapides, souvent dans des contextes familiers où ils ont l'habitude de répéter les mêmes actions, le processus de prise de décision court-circuite le cortex préfrontal pour donner la main au système habituel décrit au début de ce chapitre (Khamassi & Humphries, 2012).

Par exemple, dans la prise de décision basée sur la valeur, nous devons parfois utiliser notre cortex préfrontal pour exercer un contrôle et ainsi choisir l'option qui nous convient le mieux (par exemple, une alimentation saine) et non l'option que nous préférierions vraiment (barre de chocolat). La manière dont les valeurs des options sont modifiées par le contrôle que nous devons exercer a été étudiée par Hare, Camerer et Rangel (Hare, Camerer, & Rangel, 2009). Ils ont montré que les personnes qui intégraient des caractéristiques saines dans leur évaluation des aliments avaient leur activité vmPFC qui était corrélée à l'utilité globale (goût et santé) à l'inverse des autres personnes. Ils ont également observé que l'activité du dlPFC augmente lorsque les gens exercent une certaine maîtrise de soi dans une décision entre un aliment savoureux mais non sain et un aliment sain mais moins savoureux. De plus, ils ont constaté que le vmPFC était fonctionnellement connecté au dlPFC via une troisième zone proche du dlPFC (Gyrus frontal inférieur / BA46). Ces résultats illustrent le processus par lequel le dlPFC, une région impliquée dans la planification et le raisonnement (Koechlin, Ody, & Kouneiher, 2003), vient moduler l'activité du vmPFC pour réguler les processus d'évaluation par délibération (manger trop de chocolat est mauvais pour la santé, je décide donc de prendre un dessert plus sain cette fois-ci).

On sait qu'une compétition similaire entre décisions habituelles et décisions dirigées vers des buts est en jeu lors des comportements de navigation (Khamassi & Humphries, 2012). Par exemple, lorsque l'on visite une ville pour la première fois, on regarde souvent une carte et on essaie de mémoriser une séquence de déplacements que l'on a à effectuer. Plus tard, en se déplaçant dans la ville, si l'on constate qu'à un moment on ne se trouve pas à l'endroit escompté (e.g., on pensait trouver un musée après un dernier virage, mais on s'est trompé de rue), on est capable de re-visualiser mentalement une partie de la carte (par exemple en fermant les yeux), même imparfaitement, de réfléchir à là où on aurait dû tourner. Il s'agit d'un processus de délibération mentale dans lequel on assigne des valeurs subjectives à différentes options de déplacement en fonction de la longueur estimée de chaque chemin. On préfère en effet emprunter le chemin le plus court, ce qui fait qu'il aura une valeur subjective plus élevée. De façon frappante, ce processus de délibération n'a plus lieu lorsqu'on se déplace de manière habituelle dans une ville qui nous est familière. Par exemple, lorsque l'on va de chez soi à son université ou à son lieu de travail, on ne réfléchit plus à la séquence de déplacements qu'on s'apprête à faire. On a beau être

encore endormi, ou discuter avec quelqu'un, le système habituel déroule automatiquement la séquence de mouvements que nous avons l'habitude de faire. Il n'y a donc plus de processus de comparaison de valeurs subjectives intégrées par la délibération, mais une exécution d'une chaîne d'actions qui ont été associées ensemble par notre exposition répétitive au même déplacement.

D'une façon générale, des chercheurs comme Daniel Kahneman ont souligné que dans de nombreux domaines où nous devons prendre des décisions (navigation, décisions économiques, résolution de problèmes, etc.), nous n'utilisons le plus souvent pas notre système dirigé vers des buts, dont les calculs nécessitent un effort cognitif, mais plutôt notre système habituel qui repose sur des heuristiques et des approximations mentales (Kahneman, 2011). Or, comme nous l'avons vu, ceci nous rend plus sujet à influence par les stimuli de l'environnement, et peut nous rendre plus difficile d'échapper à certaines de nos « mauvaises » habitudes, comme le fait de grignoter entre les repas. Il est donc intéressant de réfléchir sur soi-même, et sur ce qu'on le sait à l'heure actuelle de la compétition entre systèmes d'évaluation et de prise de décision dans le cerveau, de façon à essayer de prendre moins souvent des décisions trop rapides qu'on pourrait regretter après, et plus souvent des décisions reposant sur un processus délibératif qui rende nos actions plus cohérentes avec nos propres valeurs.

c. Choix

Pour étudier les mécanismes de choix proprement dits, et non d'évaluation, de nombreuses études se sont intéressées aux choix perceptuels, qui permettent d'avoir des protocoles expérimentaux très contrôlés et des stimuli qui ont des caractéristiques objectives (comme la direction d'un mouvement, le contraste d'une image, etc.). Ces études ont beaucoup utilisé les modèles de primates non-humain, comme le macaque, sur des tâches classiques du type « *Random Dot Motion* », dans laquelle les singes doivent déterminer la direction d'un ensemble de points se déplaçant plus ou moins aléatoirement avec un degré de cohérence variable (**Figure 5**). Ils doivent indiquer leur réponse à l'aide d'une saccade oculaire dans la direction de leur choix et sont récompensés lorsque leur réponse est correcte. L'activité neurale enregistrée dans les régions du système visuo-saccadique du cerveau, telles que l'aire intrapariétale latérale (LIP ou le champ oculomoteur frontal (Frontal Eye Field FEF), augmente jusqu'à ce que la réponse soit effectuée (Roitman & Shadlen, 2002). Ce type de signal est compatible avec les modèles d'accumulation décrits précédemment, et il a été observé dans de nombreuses études (Gold & Shadlen, 2007). De manière intéressante, des signaux cérébraux compatibles avec des signaux d'accumulation (comme ceux illustrés à la **Figure 2C**) ont été observés dans d'autres régions, comme dans le réseau moteur pour des tâches où l'action à effectuer était motrice. Ces signaux reflètent une accumulation progressive d'informations dans le processus de délibération de façon à pouvoir déterminer si le nuage de points bouge principalement dans une direction ou dans l'autre. Enfin, les neurones du LIP ont aussi montré une activité de type

accumulation en réponse à différentes probabilités d'obtenir des récompenses (Platt & Glimcher, 1999). Dans l'ensemble, ces résultats suggèrent que les régions qui effectuent les choix, c'est-à-dire sélectionnent les options choisies via des actions, procèdent en accumulant de l'information sur une action spécifique, similairement à ce qui est proposé dans les modèles d'accumulation (**Figure 2**) ; lorsqu'un seuil critique est atteint, ces régions effectuent l'action pour mettre en œuvre le choix. En accord avec cette interprétation, il a été montré que le temps de décision augmente avec le degré d'ambiguïté de l'information (Camerer & Weber, 1992), comme si l'accumulation prenait davantage de temps à passer le seuil critique, parce qu'il y a moins « d'évidence » dans l'information sensorielle.

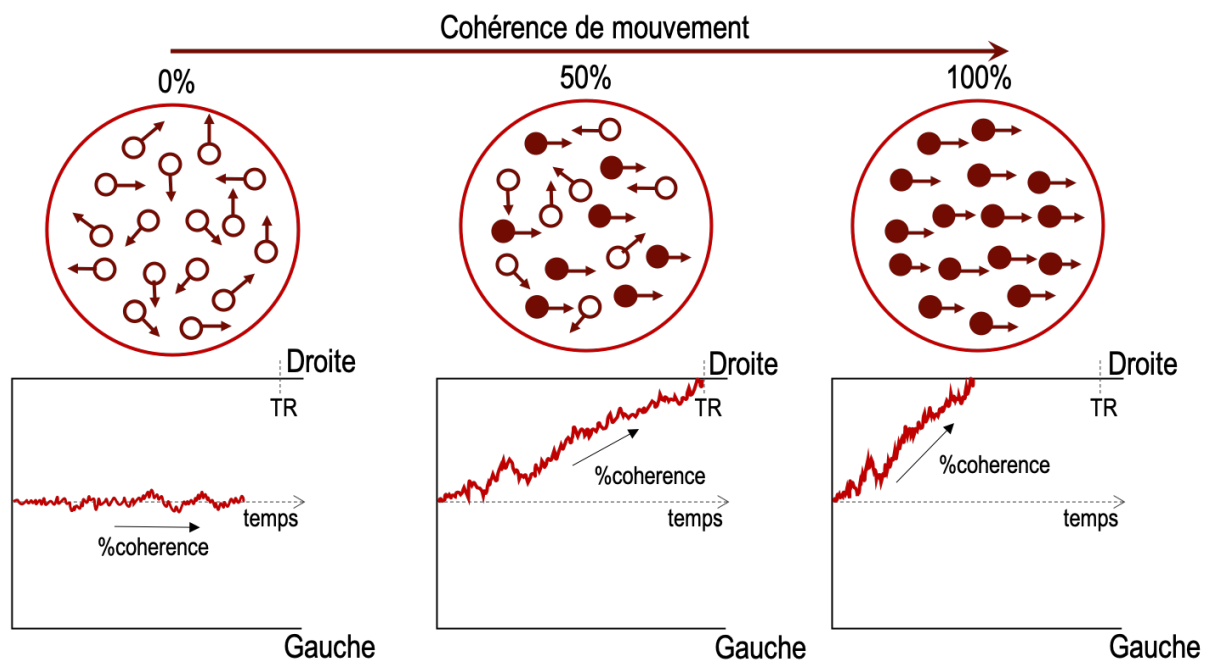


Figure 5. Exemple de stimuli classiquement utilisés dans les paradigmes de décision perceptuelle et processus d'accumulation d'évidence associés. Des points se déplacent complètement au hasard (cohérence de mouvement de 0%), partiellement dans la même direction (50% de cohérence) ou tous dans la même direction (100%). Lorsqu'il n'y a pas de cohérence de mouvement, l'accumulation vers la direction droite ou gauche se fait de manière bruitée, et la décision sera prise au hasard. Lorsque la cohérence est à 50% vers la droite, l'accumulation se fait en faveur de la direction « droite » mais le temps de réponse est lent, contrairement à lorsque la cohérence est à 100% et que l'accumulation se fait rapidement.

Chez l'humain, les techniques de neuroimagerie fonctionnelle comme l'IRMf, la MEG ou l'EEG sont limitées à cause de leur résolution temporelle ou spatiale pour étudier ce type de dynamique neuronale. Cependant, l'utilisation de modèles computationnels couplés à la neuroimagerie est de plus en plus répandue pour pallier ces limitations et identifier les mécanismes neuraux sous-jacents aux choix. Cette approche a notamment permis d'identifier des corrélations entre l'activité du vmPFC

et des variables de décisions, comme par exemple la différence entre la valeur de l'option choisie et de l'option non choisie dans un choix binaire (Boorman, Behrens, Woolrich, & Rushworth, 2009; Kable & Glimcher, 2007). Ce type de signal est attendu et prédit par les modèles d'accumulation. Cependant, l'interprétation doit être prudente car ce type de signal peut être observé à la fois dans des régions qui implémenteraient réellement le choix et dans des régions qui calculeraient la différence entre la valeur de la récompense qui va être obtenue grâce au choix (valeur attendue) et celle qui sera perdue à cause de celui-ci (notion de regret).

Conformément aux résultats chez les primates non humains, les activations cérébrales qui différencient les choix faciles des choix difficiles ont été identifiées dans les parties plus dorsales du cortex préfrontal (dACC, dlPFC), ce qui suggère que ces régions pourraient jouer un rôle essentiel dans la mise en œuvre de la comparaison. Plusieurs études tirant parti de la modélisation computationnelle ont montré que l'activité du dACC et du dlPFC pendant les choix reflètent des caractéristiques importantes telles que les signaux d'accumulation - qui sont cohérentes avec les modèles d'accumulation et leurs implémentations biologiquement plausibles (Hunt, Dolan, & Behrens, 2014; Hutcherson, Rangel, & Tuschke, 2020).

Dans l'ensemble, ces études sur les choix chez les primates humains et non humains sont cohérentes avec les résultats obtenus sur la prise de décision perceptive, où l'activité liée au choix a été identifiée dans des régions codant les entrées de décisions sous la forme de statistiques sommaires (perceptives dans les aires visuelles et LIP, information sur la valeur dans le vmPFC et OFC), mais semblent dominer dans les parties dorsales et latérales du PFC, et finalement converger vers les régions effectrices mettant en œuvre l'action résultant du choix (aires motrices).

4. Les troubles de la prise de décision

Élucider les mécanismes de prise de décision peut non seulement nous aider à mieux comprendre le fonctionnement du cerveau en condition normale, mais également nous permettre de mieux appréhender certaines pathologies neuro-psychiatriques dans lesquelles la prise de décision est altérée. Nous résumons ici les traits principaux liés à la prise de décision de certaines de ces pathologies.

Comme nous l'avons vu dans ce chapitre, des choix qui diffèrent de ceux qui seraient observés chez un agent considéré comme purement « rationnel » (au sens où ces choix ne lui permettraient pas de maximiser son *utilité* ; voir la partie 2.a Économie) ne correspondent pas forcément à un trouble de la décision, surtout s'ils sont observés dans la population générale (saine). Il existe de nombreux biais (irrationnels) dans les choix effectués par les individus sains.

Par troubles de la prise de décision, il faut entendre des troubles qui affectent les choix des individus. Ces choix semblent à la fois particulièrement inadaptés et

significativement différents des choix observés dans les mêmes situations dans la population générale. Il est important de mentionner ici que cette seconde condition est nécessaire pour qu'il y ait trouble de la prise de décision proprement dit.

Les troubles de la prise de décision ne sont pas systématiquement évalués dans les diagnostics cliniques. Ils sont en général catégorisés comme des troubles de la motivation, qui sont plus généraux et qui font partie des fonctions évaluées lors de diagnostics neuro-psychiatriques. Par exemple, l'apathie – qui correspond à une diminution pathologique de la motivation – peut être conceptualisée comme un trouble de la prise de décision, où la valeur de toutes les options est similaire ou bruitée de telle sorte qu'aucune action n'est entreprise (Pessiglione et al., 2018). De même, à l'autre extrémité du spectre, la dépendance, la manie ou l'impulsivité pourraient être causées par une évaluation aberrante de certains types de comportement.

Sous ce prisme motivationnel, des troubles de la prise de décision – parfois non spécifiques – sont observés dans de nombreuses pathologies. Par conséquent, il y a un effort récent pour évaluer plus spécifiquement les aspects des processus décisionnels chez les patients atteints de lésions cérébrales, de maladies neuro-dégénératives ou de troubles psychiatriques (Ernst & Paulus, 2005).

Troubles de la prise de décision causés par des lésions cérébrales

Le cas le plus connu de lésion cérébrale préfrontale est celui de Phineas Gage. En 1848, alors que M. Gage est en train de travailler sur la construction d'une ligne de chemin de fer aux États-Unis, une explosion se produit et une barre de fer lui transperce le crâne, provoquant une lésion dans la partie antérieure de son cortex préfrontal. Suite à cet accident, sa personnalité change et il devient plus instable dans ses décisions, et d'humeur changeante, entre autres.

En général, les lésions focales sont plutôt causées par des accidents vasculaires cérébraux ou des tumeurs. Dans le domaine de la prise de décision, de nombreuses études se sont concentrées sur les conséquences des lésions dans le cortex préfrontal (Fellows, 2019), et notamment dans le vmPFC et l'OFC. Ces lésions semblent affecter de manière assez spécifique la prise de décision basée sur la valeur : alors que d'autres aspects de la cognition sont apparemment relativement préservés, les patients ont du mal à apprendre la valeur des options dans les tâches d'apprentissage, ils présentent des préférences plus irrationnelles et incohérentes entre des paires d'objets par rapport à la population générale, et ils ont aussi des difficultés dans les tâches de prise de décision impliquant des aspects sociaux, moraux ou une prise de risque (au sens économique d'incertitude sur les gains ; voir le chapitre 10).

Quelques études ont étudié la prise de décision chez des patients présentant des lésions dorsales du cortex préfrontal (plus rares que les lésions ventrales) et ont décrit des déficits comportementaux dans diverses tâches de prise de décision, en particulier celles nécessitant du contrôle cognitif. Les décisions qui requièrent d'utiliser sa mémoire de travail, de planifier, de déplacer le focus de son attention, ou d'inhiber

une réponse semblent plus susceptibles d'être perturbées par des lésions dans le cortex préfrontal dorsal que dans le cortex orbito-frontal.

Troubles de la prise de décision dans les maladies neuro-dégénératives

Les maladies neuro-dégénératives sont caractérisées par une perte neuronale progressive, se propageant à partir de régions cérébrales spécifiques.

Maladie de Parkinson

La maladie de Parkinson est caractérisée par des pertes neurales focales des neurones dopaminergiques du mésencéphale qui innervent les circuits cortico-striataux. Les patients présentent des troubles principalement liés au mouvement (tremblements et initiation de mouvements). Cependant, les patients atteints de la maladie de Parkinson présentent également des altérations dans l'apprentissage basé sur la récompense (Frank, Seeberger, & O'reilly, 2004). Par exemple, choisir un plat au restaurant qui ne s'avère pas très bon (récompense négative) et pourtant choisir à nouveau ce même plat lors d'une prochaine visite. Il a été également souvent rapporté des déficits spécifiques dans la prise de décision liée au risque ainsi que des décisions impulsives pathologiques, parfois directement liées aux traitements dopaminergiques que les patients reçoivent. Ce dernier résultat confirme d'une autre manière (que les études plus classiques en neuro-imagerie ou en électrophysiologie) que la dopamine joue un rôle essentiel dans l'apprentissage et la prise de décision.

Démence fronto-temporale et maladie d'Alzheimer

La démence fronto-temporale (FTD) et la maladie d'Alzheimer (MA) sont deux maladies neuro-dégénératives courantes qui se caractérisent par des pertes neurales provenant des lobes frontaux et temporaux (FTD) ainsi que des structures sous-corticales comme l'hippocampe (MA). Elles provoquent des déficits dans diverses fonctions cognitives, telles que la mémoire, le langage et les réponses émotionnelles. Les pertes neurales dans le cortex préfrontal (ventral et dorsal) chez ces patients produisent des troubles de la prise de décision, notamment parce qu'ils ont tendance à faire des choix plus impulsifs, risqués et aléatoires, et qu'ils présentent des altérations de l'apprentissage basé sur les récompenses.

Troubles de la prise de décision dans les maladies psychiatriques

Des troubles de la motivation et de la prise de décision ont été observés dans tout le spectre des troubles psychiatriques, mais le tableau d'ensemble est loin d'être clair. En effet, il existe une grande variété de troubles psychiatriques, dont la plupart sont encore mal compris en termes d'origines neurobiologiques.

Troubles du comportement et de la prise de décision

Les troubles qui se manifestent principalement par des addictions ou des compulsions sont ceux qui affectent le plus la prise de décision. Les dépendances – à une substance (par exemple, l'alcool) ou à une activité (par exemple, les jeux de hasard ou de paris) –

se caractérisent par une consommation excessive de drogues ou un comportement inadapté malgré les conséquences négatives, probablement causées par un dérèglement fronto-striatal (potentiellement dopaminergique). Les troubles obsessionnels et compulsifs sont caractérisés par l'exécution d'actions répétitives et fonctionnelles (c'est-à-dire des compulsions, comme la vérification répétée de certaines choses), principalement effectuées pour soulager la détresse induite par des pensées intrusives (c'est-à-dire des obsessions). Ces deux types de troubles semblent liés à des perturbations de la prise de décision basée sur les valeurs, telles qu'une impulsivité accrue ou une recherche de risque dans les choix économiques, ainsi qu'à des déficits dans l'apprentissage basé sur la récompense, la maîtrise de soi ou les comportements dirigés vers des buts (Limbrick-Oldfield, van Holst, & Clark, 2013; Monterosso, Piray, & Luo, 2012).

Troubles des émotions et de l'humeur

Les troubles anxieux et la dépression ont en commun un aspect négatif des émotions et de la perception de soi, ainsi qu'une augmentation de l'indécision. Les patients atteints de troubles anxieux présentent des comportements d'évitement des menaces ainsi que de l'aversion au risque.

La dépression quant à elle est également associée à l'apathie, à la surestimation des efforts physiques à effectuer et à des déficits dans la recherche de récompense.

De nombreuses études récentes ont pour but de donner une description computationnelle précise des déficits décisionnels dans les troubles de l'émotion et de l'humeur, proposant notamment des déséquilibres spécifiques dans l'évaluation des récompenses par rapport aux coûts (Mathias Pessiglione, Le Bouc, & Vinckier, 2018), dans l'estimation des valeurs liées aux probabilités ou aux a priori (Paulus & Yu, 2012), ou dans l'évaluation des aspects positifs versus négatifs des options disponibles (Hartley & Phelps, 2012).

Troubles du développement

Il existe de nombreux troubles du développement, mais deux en particulier semblent impacter les processus impliqués dans la prise de décision : la schizophrénie et l'autisme.

Ces pathologies sont caractérisées par des symptômes très complexes : par exemple, la schizophrénie est associée à la fois à des symptômes dits « positifs » (comme des hallucinations et des délires) et à des symptômes dits « négatifs » (comme une perte de plaisir et un retrait social, et diverses formes de troubles cognitifs).

Il a été observé dans ces troubles des altérations de la prise de décision. Par exemple, les patients atteints de schizophrénie ont du mal à ajuster leurs réponses après avoir perdu une récompense, alors que la population générale s'adapte à ce type de retour négatif. Ce défaut d'ajustement a été associé à des activités anormales dans l'OFC (Waltz et al., 2010).

Aujourd'hui, on ne sait toujours pas comment les symptômes classiques de ces troubles peuvent être liés à des aspects spécifiques du processus décisionnel.

Néanmoins, des études récentes en psychiatrie computationnelle ont proposé que ces troubles complexes pourraient être expliqués par des processus de traitement de récompense et de motivation aberrants (Strauss, Waltz, & Gold, 2014), ou par des perturbations des processus d'apprentissage (Sterzer et al., 2018).

Complémentarité de la recherche clinique et fondamentale sur la prise de décision

Dans l'ensemble, les troubles de la prise de décision observés dans un large éventail de conditions cliniques constituent une autre façon de confirmer l'implication critique de l'OMPFC et du PFC dorsal dans les processus de prise de décision, en parallèle des études fondamentales présentées dans ce chapitre. En d'autres termes, comprendre comment le cerveau fonctionne en situation pathologique apporte aussi son lot d'informations pour nous aider à comprendre comment le cerveau fonctionne en situation dite « normale ».

De même, les études cliniques confirment que le système dopaminergique du mésencéphale, qui cible principalement le cortex préfrontal, apparaît comme central dans les mécanismes cérébraux de la prise de décision. Alors que les études de recherche fondamentale sont principalement basées sur des résultats de corrélation, les déficits observés chez les patients fournissent des explications causales de l'implication des régions cérébrales ou des systèmes de neurotransmetteurs. Les études fondamentales permettent de tester facilement des nouveaux paradigmes expérimentaux, et de développer des modèles computationnels de fonctionnement sain, et les études cliniques permettent quant à elles d'identifier les régions nécessaires au fonctionnement sain d'un processus cognitif. En effet, si on se rend compte qu'un mécanisme donné de prise de décision est systématiquement altéré à chaque fois que des patients ont une lésion ou une inactivation provisoire d'une région cérébrale particulière, alors on peut considérer que le lien entre le mécanisme et la région va au-delà de la simple corrélation.

La complémentarité entre études fondamentales et cliniques est capitale. De plus en plus d'études adoptent ainsi des approches neuro-computationnelles dans l'idée de développer des outils de diagnostic pour mieux comprendre le très large éventail de troubles neurologiques et psychiatriques qui affectent la prise de décision. (Wang & Krystal, 2014)(Wang et Krystal, 2014). Certains de ces modèles computationnels sont présentés dans le **chapitre 10**.

5. À retenir

En résumé, nous avons vu dans ce chapitre que des mécanismes cérébraux de prise de décision sont étudiés dans deux grandes familles de décisions : les décisions perceptuelles (décider de ce que l'on perçoit lorsque les stimuli sont ambigus), et les

décisions économiques (décider entre plusieurs options ou plusieurs actions à effectuer après avoir comparé leurs valeurs subjectives).

Nous avons également vu que différents systèmes correspondant à des réseaux cérébraux pouvaient contribuer à ces décisions : (1) le système pavlovien, qui biaise le comportement en donnant une tendance à être attiré par les stimuli prédicteurs de récompense, et à être repoussé par les stimuli prédicteurs de conséquences négatives ; (2) le système dirigé vers des buts, qui implique tout un processus de délibération pour estimer les conséquences possibles des actions avant de décider que faire ; (3) le système habituel, qui développe progressivement des automatismes comportementaux, des routines, à force d'être exposés de manière répétée aux mêmes décisions dans des mêmes contextes. Ce dernier système nous pousse le plus souvent à prendre des décisions rapides sans anticiper les conséquences de nos actions, et ainsi nous rend davantage influençables par les stimuli de l'environnement, par le contexte, ou encore par nos émotions.

Nous avons ensuite vu que de nombreux enregistrements de l'activité cérébrale par des méthodes de neurophysiologie chez l'animal en comportement ont permis de mieux comprendre le fonctionnement et l'organisation de ces systèmes de prise de décision dans le cerveau. Les études d'imagerie par résonance magnétique fonctionnelle (IRMf) ont plus récemment permis de confirmer chez l'humain de nombreux résultats d'abord établis chez l'animal. Elles ont également permis d'apporter d'autres connaissances sur la communication entre aires cérébrales. Elles n'apportent toutefois pas le degré de précision spatio-temporelle des enregistrements neurophysiologiques intra-cérébraux.

Enfin, nous avons brièvement présenté un certain nombre de pathologies et de troubles neurodéveloppementaux qui impliquent des troubles de la prise de décision. L'étude de ces cas pathologiques permet à la fois une dimension clinique, pour contribuer à trouver de meilleurs traitements pour les patients, et permet également de contribuer à la compréhension du fonctionnement des mécanismes de prise de décision en mettant en lumière leurs propriétés lorsque des déficits précis sont en jeu.

6. Pour aller plus loin

BORAUD, Thomas. Matière à décision. CNRS Éditions, 2015.

Daw, N. D., & O'Doherty, J. P. (2014). Multiple systems for value learning. In *Neuroeconomics* (pp. 393-410). Academic Press.

GLIMCHER, Paul W. Decisions, uncertainty, and the brain: The science of neuroeconomics. MIT press, 2003.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Khamassi, M., & Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, OCTOBER 2012.

KHAMASSI, Mehdi et PACHERIE, Elisabeth. L'action. Dans Collins, T., Andler, D. and Tallon-Baudry, C. (Eds.), *La cognition : du neurone à la société*, Paris, France : Gallimard. 2018.

Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., ... Pessiglione, M. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nature Neuroscience*, 23(5), 664-675.

Pessiglione, M. (2021). *Les vacances de Momo Sapiens. Notre cerveau, entre raison et déraison*. Éditions Odile Jacob.

7. Bibliographie

Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens.

NeuroImage, 31(2), 790-795. <https://doi.org/10.1016/j.neuroimage.2006.01.001>

Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits : Neural substrates of parallel processing. *Trends in Neuroscience*, 13(7), 266-271.

Anderson, N. D. (2015). Teaching signal detection theory with pseudoscience.

Educational Psychology, 6, 762. <https://doi.org/10.3389/fpsyg.2015.00762>

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action :

Contingency and incentive learning and their cortical substrates.

Neuropharmacology, 37(4-5), 407-419. [https://doi.org/10.1016/S0028-3908\(98\)00033-1](https://doi.org/10.1016/S0028-3908(98)00033-1)

- Barron, H. C., Dolan, R. J., & Behrens, T. E. J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nature Neuroscience*, 16(10), 1492-1498. <https://doi.org/10.1038/nn.3515>
- Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system : A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, 76, 412-427. <https://doi.org/10.1016/j.neuroimage.2013.02.063>
- Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3), 226-232.
- Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, 23-36.
- Bonnelle, V., Veromann, K.-R., Burnett Heyes, S., Lo Sterzo, E., Manohar, S., & Husain, M. (2015). Characterization of reward and effort mechanisms in apathy. *Journal of Physiology, Paris*, 109(1-3), 16-26. <https://doi.org/10.1016/j.jphysparis.2014.04.002>
- Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. *Neuron*, 62(5), 733-743. <https://doi.org/10.1016/j.neuron.2009.05.014>
- Butter, C. M., McDonald, J. A., & Snyder, D. R. (1969). Orality, Preference Behavior, and Reinforcement Value of Nonfood Object in Monkeys with Orbital Frontal

Lesions. *Science*, 164(3885), 1306-1307.

<https://doi.org/10.1126/science.164.3885.1306>

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences : Uncertainty and ambiguity. *Journal of risk and uncertainty*, 5(4), 325-370.

Clark, J. J., Hollon, N. G., & Phillips, P. E. (2012). Pavlovian valuation systems in learning and decision making. *Current Opinion in Neurobiology*, 22(6), 1054-1061.
<https://doi.org/10.1016/j.conb.2012.06.004>

Cléry-Melin, M.-L., Schmidt, L., Lafargue, G., Baup, N., Fossati, P., & Pessiglione, M. (2011). Why Don' t You Try Harder? An Investigation of Effort Production in Major Depression. *PLoS ONE*, 6(8), e23178.
<https://doi.org/10.1371/journal.pone.0023178>

Cromwell, H. C., & Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *Journal of neurophysiology*, 89(5), 2823-2838.

Ernst, M., & Paulus, M. P. (2005). Neurobiology of decision making : A selective review from a neurocognitive and clinical perspective. *Biological psychiatry*, 58(8), 597-604.

Fellows, L. K. (2019). The functions of the frontal lobes : Evidence from patients with focal brain damage. In *Handbook of Clinical Neurology* (Vol. 163, p. 19-34). Elsevier. <https://doi.org/10.1016/B978-0-12-804281-6.00002-1>

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(March), 1898-1902.

- Frank, M. J., Seeberger, L. C., & O' reilly, R. C. (2004). By carrot or by stick : Cognitive reinforcement learning in parkinsonism. *Science*, 306(5703), 1940-1943.
- Glimcher, P. (2014). Understanding the Hows and Whys of Decision-Making : From Expected Utility to Divisive Normalization. *Cold Spring Harbor Symposia on Quantitative Biology*, 79, 169-176. <https://doi.org/10.1101/sqb.2014.79.024778>
- Glimcher, P. W., & Fehr, E. (2013). *Neuroeconomics : Decision making and the brain*. Academic Press. Consulté à l' adresse <https://books.google.fr/books?hl=fr&lr=&id=qToTAAAAQBAJ&oi=fnd&pg=P1&dq=neuroeconomics&ots=t7yqvE22ng&sig=R21-ZrkrPuJN14RKXFIrxPH7Sgs>
- Gluth, S., Sommer, T., Rieskamp, J., & Büchel, C. (2015). Effective Connectivity between Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices from Memory. *Neuron*, 86(4), 1078-1090. <https://doi.org/10.1016/j.neuron.2015.04.023>
- Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, 30(1), 535-574. <https://doi.org/10.1146/annurev.neuro.29.051605.113038>
- Greicius, M. D., Supekar, K., Menon, V., & Dougherty, R. F. (2009). Resting-State Functional Connectivity Reflects Structural Connectivity in the Default Mode Network. *Cerebral Cortex*, 19(1), 72-78. <https://doi.org/10.1093/cercor/bhn059>
- Haber, S. N. (2014). The place of dopamine in the cortico-basal ganglia circuit. *Neuroscience*, 282, 248-257. <https://doi.org/10.1016/j.neuroscience.2014.10.008>

- Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. *Science*, 324(5927), 646-648. <https://doi.org/10.1126/science.1168450>
- Hare, T. A., O' Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the Role of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and Prediction Errors. *Journal of Neuroscience*, 28(22), 5623-5630. <https://doi.org/10.1523/JNEUROSCI.1309-08.2008>
- Hartley, C. A., & Phelps, E. A. (2012). Anxiety and decision-making. *Biological psychiatry*, 72(2), 113-118.
- Holland, P. C. (2004). Relations Between Pavlovian-Instrumental Transfer and Reinforcer Devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, 30(2), 104-117. <https://doi.org/10.1037/0097-7403.30.2.104>
- Humphries, M. D., Khamassi, M., & Gurney, K. (2012). Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Frontiers in Neuroscience*, (FEB). <https://doi.org/10.3389/fnins.2012.00009>
- Hunt, L. T., Dolan, R. J., & Behrens, T. E. J. (2014). Hierarchical competitions subserving multi-attribute choice. *Nature Neuroscience*. <https://doi.org/10.1038/nn.3836>
- Hutcherson, C. A., Rangel, A., & Tusche, A. (2020). *Evidence accumulation, not “self-control,” explains why the dlPFC activates during normative choice* [Preprint]. *Neuroscience*. <https://doi.org/10.1101/2020.10.06.328476>

- Iversen, S. D., & Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research*, 11(4), 376-386.
- Jevons, W. S. (1879). *Political economy* (Vol. 2).
- Jones, B., & Mishkin, M. (1972). Limbic lesions and the problem of stimulus-Reinforcement associations. *Experimental neurology*, 36(2), 362-377.
- Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10(12), 1625-1633.
<https://doi.org/10.1038/nn2007>
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Khamassi, M., & Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, (OCTOBER 2012).
<https://doi.org/10.3389/fnbeh.2012.00079>
- Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., & Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal neurons of behaving rats. *European Journal of Neuroscience*, 28(9). <https://doi.org/10.1111/j.1460-9568.2008.06480.x>
- Kinchla, R. A., & Smyzer, F. (1967). A diffusion model of perceptual memory. *Perception & Psychophysics*, 2(6), 219-229. <https://doi.org/10.3758/BF03212471>

- Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, 12(17), 3683-3687.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, 302(5648), 1181-1185.
<https://doi.org/10.1126/science.1088545>
- Lebreton, Maël, Bertoux, M., Boutet, C., Lehericy, S., Dubois, B., Fossati, P., & Pessiglione, M. (2013). A Critical Role for the Hippocampus in the Valuation of Imagined Outcomes. *PLoS Biology*, 11(10), e1001684.
<https://doi.org/10.1371/journal.pbio.1001684>
- Lebreton, Maël, Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. *Neuron*, 64(3), 431-439.
<https://doi.org/10.1016/j.neuron.2009.09.040>
- Leech, R., & Sharp, D. J. (2013). The role of the posterior cingulate cortex in cognition and disease. *Brain*, 137(1), 12-32. <https://doi.org/10.1093/brain/awt162>
- Limbrick-Oldfield, E. H., van Holst, R. J., & Clark, L. (2013). Fronto-striatal dysregulation in drug addiction and pathological gambling: Consistent inconsistencies? *NeuroImage: Clinical*, 2, 385-393.
- Lin, A., Adolphs, R., & Rangel, A. (2011). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, 7(3), 274-281. <https://doi.org/10.1093/scan/nsr006>

- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40(1), 77-105. <https://doi.org/10.1007/BF02291481>
- Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., ... Pessiglione, M. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nature Neuroscience*, 23(5), 664-675. <https://doi.org/10.1038/s41593-020-0615-9>
- Luce, R. D. (1959). On the possible psychophysical laws. *Psychological review*, 66(2), 81.
- McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, 38(2), 339-346.
- Mesulam, M. M. (1998). From sensation to cognition. *Brain: A Journal of Neurology*, 121 (Pt 6), 1013-1052. <https://doi.org/10.1093/brain/121.6.1013>
- Mill, J. S. (1901). *Utilitarianism*. Longmans, Green and Company.
- Montague, P. R., & King-Casas, B. (2007). Efficient statistics, common currencies and the problem of reward-harvesting. *Trends in cognitive sciences*, 11(12), 514-519.
- Monterosso, J., Piray, P., & Luo, S. (2012). Neuroeconomics and the study of addiction. *Biological psychiatry*, 72(2), 107-112.
- Mora, F., Avrith, D. B., Phillips, A. G., & Rolls, E. T. (1979). Effects of satiety on self-stimulation of the orbitofrontal cortex in the rhesus monkey. *Neuroscience Letters*, 13(2), 141-145. [https://doi.org/10.1016/0304-3940\(79\)90031-4](https://doi.org/10.1016/0304-3940(79)90031-4)
- O' Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science*, 304(5669), 452-454. <https://doi.org/10.1126/science.1094285>

- O' Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature neuroscience*, 4(1), 95-102.
- Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, 47(6), 419-427. <https://doi.org/10.1037/h0058775>
- Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making : Affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, 16(9), 476-483. <https://doi.org/10.1016/j.tics.2012.07.009>
- Pavlov, I. P. (1927). Conditioned reflexes. *An Investigation of the physiological activity of the cerebral cortex*.
- Pearson, J. M., Hayden, B. Y., Raghavachari, S., & Platt, M. L. (2009). Neurons in Posterior Cingulate Cortex Signal Exploratory Decisions in a Dynamic Multioption Choice Task. *Current Biology*, 19(18), 1532-1537. <https://doi.org/10.1016/j.cub.2009.07.048>
- Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the Brain Translates Money into Force : A Neuroimaging Study of Subliminal Motivation. *Science*, 316(5826), 904-906. <https://doi.org/10.1126/science.1140459>
- Pessiglione, Mathias, Le Bouc, R., & Vinckier, F. (2018). When decisions talk : Computational phenotyping of motivation disorders. *Current Opinion in Behavioral Sciences*, 22, 50-58. <https://doi.org/10.1016/j.cobeha.2017.12.014>

- Pessiglione, Mathias, Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106), 1042-1045. <https://doi.org/10.1038/nature05051>
- Peters, J., & Büchel, C. (2010). Neural representations of subjective reward value. *Behavioural Brain Research*, 213(2), 135-141. <https://doi.org/10.1016/j.bbr.2010.04.031>
- Peterson, W., Birdsall, T., & Fox, W. (1954). The theory of signal detectability. *Transactions of the IRE professional group on information theory*, 4(4), 171-212.
- Plassmann, H., O' Doherty, J., & Rangel, A. (2007). Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions. *Journal of Neuroscience*, 27(37), 9984-9988. <https://doi.org/10.1523/JNEUROSCI.2131-07.2007>
- Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400(6741), 233-238.
- Polanía, R., Moisa, M., Opitz, A., Grueschow, M., & Ruff, C. C. (2015). The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nature Communications*, 6, 8090. <https://doi.org/10.1038/ncomms9090>
- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7), 545-556. <https://doi.org/10.1038/nrn2357>
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59-108. <https://doi.org/10.1037/0033-295X.85.2.59>

- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model : Theory and data for two-choice decision tasks. *Neural computation*, 20(4), 873-922.
- Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia : A vertebrate solution to the selection problem? *Neuroscience*, 89(4), 1009-1023.
[https://doi.org/10.1016/S0306-4522\(98\)00319-4](https://doi.org/10.1016/S0306-4522(98)00319-4)
- Rescorla, R. A., Wagner, A. R., & others. (1972). A theory of Pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2, 64-99.
- Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience*, 10(12), 1615-1624. <https://doi.org/10.1038/nn2013>
- Roitman, J. D., & Shadlen, M. N. (2002). Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. *The Journal of Neuroscience*, 22(21), 9475-9489.
- Rolls, E. T., Burton, M. J., & Mora, F. (1980). Neurophysiological analysis of brain-stimulation reward in the monkey. *Brain research*, 194(2). Consulté à l' adresse <http://www.sciencedirect.com/science/article/pii/0006899380912160>
- Rosenkilde, C. E., Bauer, R. H., & Fuster, J. M. (1981). Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain research*, 209(2), 375-394.
- Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Neuroscience : Representation of action-specific reward values in the striatum. *Science*, 310(5752), 1337-1340.
<https://doi.org/10.1126/science.1115270>

- Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural Mechanisms Underlying Motivation of Mental Versus Physical Effort. *PLoS Biology*, 10(2), e1001266. <https://doi.org/10.1371/journal.pbio.1001266>
- Schultz, W, Dayan, P., & Montague, P. R. (1997). A neural substrate for prediction and reward. *Science*, 275, 1593-1599.
- Schultz, Wolfram. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80(1), 1-27. <https://doi.org/10.1152/jn.1998.80.1.1>
- Sidgwick, H. (1907). *The methods of ethics*. Hackett Publishing. Consulté à l' adresse https://books.google.fr/books?hl=fr&lr=&id=YYYsGTuH8HYC&oi=fnd&pg=PR1&dq=method+of+ethics&ots=_e39EboP9M&sig=aYlFVM-NI__8cXm8MYi6qyVyek
- Snodgrass, J. G., Townsend, J. T., & Ashby, F. G. (1985). Stochastic Modeling of Elementary Psychological Processes. *The American Journal of Psychology*, 98(3), 480. <https://doi.org/10.2307/1422636>
- Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., ... Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological psychiatry*, 84(9), 634-643.
- Strauss, G. P., Waltz, J. A., & Gold, J. M. (2014). A review of reward processing and motivational impairment in schizophrenia. *Schizophrenia bulletin*, 40(Suppl_2), S107-S116.
- Tanner Jr, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological review*, 61(6), 401.

- Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. *Neuron*, 66(5), 781-795. <https://doi.org/10.1016/j.neuron.2010.04.036>
- Thorndike, E. L. (1898). Animal intelligence : An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, 2(4), i.
- Varian, H. R. (1992). *Microeconomic analysis*. Norton & Company. Consulté à l'adresse <http://thuvien.due.udn.vn:8080/dspace/handle/TVDHKT/19371>
- Waltz, J. A., Schweitzer, J. B., Ross, T. J., Kurup, P. K., Salmeron, B. J., Rose, E. J., ... Stein, E. A. (2010). Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology*, 35(12), 2427-2439.
- Wang, X.-J., & Krystal, J. H. (2014). Computational psychiatry. *Neuron*, 84(3), 638-654.
- Wimmer, G. E., & Shohamy, D. (2012). Preference by Association : How Memory Mechanisms in the Hippocampus Bias Decisions. *Science*, 338(6104), 270-273. <https://doi.org/10.1126/science.1223252>