HAL open science

# Initiation à la modélisation computationnelle

Anne Collins, Mehdi Khamassi

**To cite this version:**

Anne Collins, Mehdi Khamassi. Initiation à la modélisation computationnelle. Khamassi, M. (Ed.) Neurosciences cognitives, 2021. hal-03411274

**HAL Id: hal-03411274**

**https://hal.science/hal-03411274**

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.

# INITIATION À LA MODÉLISATION COMPUTATIONNELLE

A PREPRINT

**Anne Collins**
Department of Psychology
University of California, Berkeley
Berkeley, CA 394712
annecollins@berkeley.edu

**Mehdi Khamassi**
Sorbonne Université, CNRS
Institut des Systèmes Intelligents et de Robotique
F-75005 Paris, France
mehdi.khamassi@sorbonne-universite.fr

February 3, 2021

## 1 Introduction

***Qu'est-ce que la modélisation computationnelle ?***

La modélisation computationnelle, parfois aussi appelée modélisation mathématique, consiste à décrire un phénomène par une ou plusieurs équations, selon la complexité du phénomène. Elle est de plus en plus en vogue en psychologie expérimentale et en neurosciences cognitives, car elle apporte un outil complémentaire à l'expérimentation.

***A quoi ça sert ?***

La modélisation apporte plusieurs bénéfices importants. Tout d'abord, elle permet de décrire les phénomènes observés dans la nature de façon synthétique. En d'autres termes, il s'agit de rendre compte de grandes quantités de données expérimentales avec seulement quelques équations qui offrent une *explication* au phénomène observé en termes de mécanismes simples.

Ensuite, elle permet de bénéficier de la précision du langage mathématique. Ainsi, au lieu de se contenter de mettre en mots une description du phénomène qui risque d'être comprise de façon ambiguë selon la connotation des mots, ou même selon le sens des mots qui peut varier d'une discipline à l'autre, une équation mathématique ne souffre pas de mésinterprétation. La modélisation peut donc apporter de la clarté, particulièrement dans le dialogue interdisciplinaire.

Troisièmement, on observe actuellement un nombre croissant d'études où les chercheurs analysent les données d'une expérience de façon dite 'model-based', c'est-à-dire en confrontant directement les données aux variables du modèle. Par exemple, dans des tâches de prise de décision comme celles décrites au **chapitre 6** on peut vouloir étudier si le temps de décision des sujets augmente après un essai d'erreur (comme si les sujets étaient moins sûrs d'eux et prenaient plus de temps pour réfléchir). Une méthode d'analyse classique consistera à simplement dissocier les essais qui suivent une erreur des essais qui suivent une bonne réponse, et d'utiliser un test statistique pour déterminer si les temps de décision sont significativement différents entre ces deux groupes d'essais. Alternativement (ou plutôt, de façon complémentaire), si l'on dispose d'un modèle de prise de décision, on peut alors regarder si l'évolution d'essai en essai du temps de décision des sujets corrèle avec une variable qui représente le temps de délibération dans notre modèle. Nous y reviendrons plus tard en décrivant concrètement les méthodes de liage d'un modèle aux données (on parle en anglais de *'model fitting'*), puis l'utilisation des variables du modèle comme *régresseurs* des données expérimentales.

Enfin, et peut-être de façon encore plus importante, la modélisation permet d'être prédictif. En effet, si on réussit à modéliser un phénomène sous forme d'une équation mathématique, alors cette équation nous permet de prédire ce qu'on devrait observer dans telle ou telle nouvelle situation expérimentale, même si l'expérience n'a encore jamais été faite ! En effet, si on reste sur notre exemple des temps de décision et qu'on dispose d'un modèle qui dit grossièrement que le temps de délibération augmente de la même quantité après une erreur et qu'il diminue après une bonne réponse, alors on peut imaginer une expérience dans laquelle on contraint artificiellement le sujet à faire une bonne réponse, puis une erreur, puis une bonne réponse, puis une erreur, et ainsi de suite. Notre modèle prédit dans ce cas qu'au bout d'un nombre pair d'essais, le temps de décision des sujets devrait être identique à leur temps de décision initial. On peut donc mettre en place le protocole expérimental et faire passer des sujets pour voir si la prédiction du modèle est juste ou non.

Un modèle computationnel peut donc donner des idées de nouvelles expériences à réaliser pour tester la validité du modèle. Le modèle peut parfois être invalidé par les résultats de la nouvelle expérience. On aura alors appris que le modèle est faux et qu'il faut en concevoir un nouveau. Dans le meilleur des cas, si des chercheurs différents ont proposé deux modèles différents, le modèle A et le modèle B, qui font des prédictions contradictoires, alors on peut réaliser une expérience conçue spécifiquement pour déterminer lequel des deux modèles est faux, et lequel reste toujours valide. La modélisation peut ainsi guider l'expérimentation, et c'est un de ses bénéfices sur lequel nous reviendrons tout au long du chapitre avec des exemples.

***Qu'est que le mot 'computationnel' veut dire ?***

Le mot *'computation'* en anglais signifie calcul. Un modèle computationnel est donc un modèle qui fait des calculs à partir d'informations données en entrée, et qui donne en sortie un résultat numérique.

Dans le champ d'étude de la cognition, cette vision des choses s'est grandement répandue à partir de ce qu'on appelle *'la révolution cognitive'* des années 60, qui a donné naissance aux sciences cognitives. Au cours de ce changement de perspective, des psychologues comme Jerry Fodor et des linguistes comme Noam Chomsky (mais pas seulement eux) ont proposé de considérer le cerveau comme un système de traitement de l'information, et la pensée comme un ensemble d'opérations, donc de calculs, effectués sur l'information.

Au début des années 80, le neurobiologiste David Marr a fondé le champ des *neurosciences computationnelles* en cherchant à élaborer une méthode qui permette de modéliser, et donc de mieux comprendre, les traitements d'information effectués par le système visuel dans le cerveau (Marr, 1982). Il a pour cela proposé de distinguer 3 niveaux nécessaires à l'élaboration de toute théorie computationnelle, le but dans son cas ayant été de proposer une théorie du fonctionnement du système visuel. Ces trois niveaux sont :

1. Le niveau computationnel, ou fonctionnel,
2. Le niveau algorithmique,
3. Le niveau implémentationnel.

Le niveau computationnel vise à décrire quelle est la fonction du système étudié. Par exemple, dans le cas de la vision, la théorie de Marr est que le système visuel a pour fonction de reconstruire (ou simplement construire, mais c'est une question philosophique, voir le **chapitre 2**) une représentation des éléments physiques qui nous entourent à partir des informations reçues par les photorécepteurs sur la rétine. Le niveau algorithmique vise à décrire les représentations et les algorithmes, donc les suites de calculs, que le système opère afin de réaliser sa fonction. Le niveau implémentationnel décrit comment ces algorithmes sont mis en oeuvre dans un substrat naturel, donc en l'occurrence dans le substrat biologique qu'est le cerveau.

Par exemple, si la fonction du système visuel est de reconnaître à un moment donné un objet perçu comme étant distinct de l'environnement dans lequel il est situé, alors un modèle computationnel décrira au premier niveau l'objectif de la série de calculs qu'il effectue comme étant de délimiter les contours de l'objet pour l'isoler du reste. Au niveau algorithmique, le modèle décrira la suite d'opérations qui permettent de détecter des bords à partir de mesures de contraste dans la lumière reçue par la rétine. Et le niveau implémentationnel (si le modèle en contient un) proposera une façon dont ces calculs sont répartis et communiquent entre différentes aires du cerveau, de par leur connectivité mutuelle. Dans ce chapitre, nous nous concentrerons sur les modèles de prise de décision, et les lecteurs qui veulent comprendre la fonction visuelle au-delà de cette description très schématique sont invités à lire le **chapitre 3**.

***Quelle différence y a-t-il entre un modèle descriptif et un modèle prédictif ?***

On fait très souvent la distinction entre les modèles descriptifs et les modèles prédictifs. Les premiers visent à rendre compte de données expérimentales sans proposer de fonction, au sens de raison d'être, de ces données. On se contentera alors de décrire mathématiquement les relations qu'il peut y avoir entre telles et telles mesures enregistrées par l'expérience. C'est ce qu'on fait lorsqu'on effectue une régression logistique des données expérimentales à partir d'une équation, ce qui permet déjà d'apporter le premier bénéfice de synthéticité décrit plus haut.

À l'inverse, les modèles prédictifs vont chercher à décrire quelle est la fonction sous-jacente aux données observées, de façon à pouvoir expliquer pourquoi on s'attend à observer ces données dans telle situation et pas dans telle autre. On cherche donc à proposer une explication mécanistique du phénomène.

Pour comprendre cette distinction, prenons un exemple dans lequel on cherche à étudier l'influence de la température ambiante sur le temps quotidien que passent les bébés rats à se blottir les uns contre les autres. On commencera par faire une expérience dans laquelle on notera une liste d'observations : il fait 3°C, le rat 1 a passé 65 minutes à se blottir contre les autres ; il fait 9°C, le rat 1 a passé 5 minutes à se blottir contre les autres ; il fait 12°C, le rat 2 a passé 30 minutes à se blottir contre les autres, etc. A partir de cette série d'observations, on pourrait se rendre compte qu'il y a

une relation de proportionnalité : plus il fait froid, plus les bébés rats passent de temps blottis les uns contre les autres, et plus il fait chaud moins c'est le cas. Bien sûr, on se rendrait compte qu'il y a une certaine variabilité inter-individuelle, de sorte qu'à une température donnée, certains individus aiment plus se blottir que d'autres.

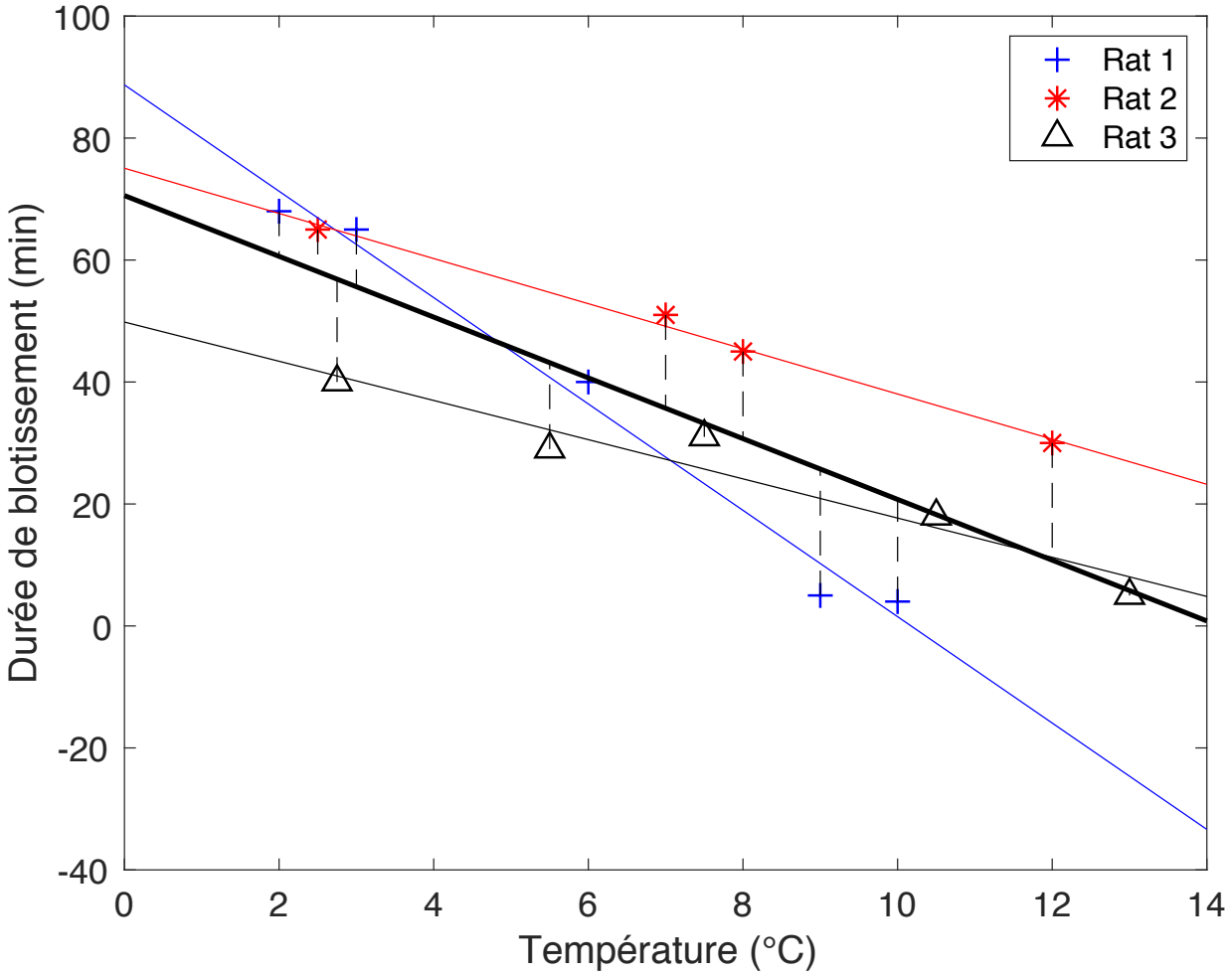


Figure 1: Durée de blotissement fictif de bébés rats en fonction de la température ambiante. À titre illustratif, on imagine ici que différents rats ont différentes tendances à diminuer leur temps de blotissement en fonction des variations de température. Ces tendances pourraient être *fittées* par des droites (fonctions affines) spécifiques à chaque individu (traits fins), ou par une seule droite pour l'ensemble des données (trait épais). Dans ce dernier cas, l'erreur de fit est illustrée en pointillé comme la distance de chaque point à la droite. Le code source pour construire cette figure est disponible à l'adresse `https://github.com/MehdiKhamassi/basicModelingTutorial`.

Si l'on dessine un graphique où chaque point représente une mesure, avec la température sur l'axe des abscisses et la durée de blotissement sur l'axe des ordonnées, on obtiendra un nuage de points étiré vers le bas droit du graphique, qui correspondra à peu près à une droite linéaire qui représenterait cette relation de proportionalité (**Figure 1**). Dans un premier temps, on pourrait donc être tenté de faire passer une droite par ce nuage de points, donc de *'fitter'* les données avec une équation linéaire ou affine. La pente de cette équation permettant à la droite de passer le plus au milieu des points possible sera le premier paramètre libre de notre modèle. L'ordonnée à l'origine sera notre deuxième paramètre. On peut aussi parfois vouloir ajouter un terme à l'équation du modèle qui représente un certain bruit statistique de faible amplitude. Ceci permettrait notamment de rendre compte des erreurs de mesure expérimentale, ou de rendre compte superficiellement des différences inter-individuelles observées. Néanmoins, comme on peut le voir sur la figure, ce modèle descriptif simple *fitté* à l'ensemble des données sans tenir compte de la variabilité inter-individuelle n'est pas satisfaisant. En effet, si on fitte des modèles affines distincts sur les données de chaque rat individuellement, on peut constater que les paramètres du modèle ne sont pas les mêmes entre les rats : les rats 2 et 3 diminuent à peu près d'autant leur durée de blotissement quand la température augmente, mais le rat 2 passe en moyenne plus de temps à se

blottir que le rat 3. Le rat 1, lui, passe principalement du temps à se blottir quand il fait froid, et quasiment pas dès qu'il fait plus de 7-8°C. Ce dernier cas suggère qu'un modèle non-linéaire (par exemple, une sigmoïde) serait peut-être plus approprié, car cela permettrait notamment d'éviter des durées de blotissement négatives au-delà de 10°C.

De façon plus centrale pour la question traitée, ce modèle descriptif reliant la température à la durée de blotissement n'a qu'un pouvoir prédictif limité. Certes, si l'expérimentateur n'a jamais fait de mesure à une température de 25°C, le modèle permet néanmoins de prédire quelle durée moyenne de blotissement on s'attend à observer dans notre population de bébés rats, ou même chez chaque rat pris individuellement. Néanmoins, si on changeait plus drastiquement les conditions de l'expérience, on ne serait plus capable de faire des prédictions avec ce modèle descriptif. Par exemple, si au préalable on mettait les bébés rats dans une pièce chauffée à une température $x_0$, avant de les remettre dans la zone expérimentale où il fait une température $x$ et où l'on souhaite faire notre mesure expérimentale, alors notre modèle n'aurait aucune manière de prédire l'influence de $x_0$ sur la durée de blotissement. Si l'on refaisait plusieurs fois l'expérience avec différents $x_0$, on se rendrait sûrement compte que la durée de blotissement moyenne n'est pas la même, de sorte que notre nuage de points montrera une plus grande variance qu'on ne sait pas expliquer (à ce stade de la modélisation).

Au contraire, un modèle qui chercherait à donner une explication du phénomène observé en proposant une explication des mécanismes sous-jacents pourrait nous permettre d'avoir une plus grande capacité prédictive. Par exemple, on pourrait proposer un modèle qui considère qu'à chaque fois que la température du corps passe en-dessous d'un certain seuil $s_i$ ($i \in \{1, 2, .., n\}$, avec $n$ le nombre de bébés rats dans la population), le cerveau déclenche un comportement de blotissement qui a pour effet de faire remonter la température. Ce seuil varie d'un individu à l'autre : $s_i \neq s_j$ si $i \neq j$. Ainsi notre modèle devra contenir une équation pour décrire la probabilité de déclencher un comportement de blotissement en fonction de la température $t_i$ actuelle du corps de l'individu $i$ et de son seuil $s_i$. Le modèle devra également contenir une deuxième équation pour décrire l'influence du comportement de blotissement sur la température du corps : plus précisément, comment celle-ci varie au cours du temps en fonction de la surface du corps en contact avec d'autres corps et de la température de ces corps, et comment l'absence de contact fait diminuer la température du corps.

On peut voir qu'avec ce deuxième modèle (qui dit que si on néglige tous les aspects sociaux et émotionnels, la fonction du comportement de blotissement est d'augmenter la température) on a un pouvoir prédictif plus grand qu'avec le premier modèle. En effet, le fait de placer un bébé rat dans une première pièce chauffée à une température $x_0$ avant de le replacer dans la zone expérimentale pourra influencer la température initiale du corps du bébé rat au début de la mesure, selon les équations du modèle. Et en fonction du seuil $s_i$ de tolérance au froid que l'on aurait trouvé en *'fittant'* le modèle sur les données de ce bébé rat lors de la première expérience, le modèle nous permettra de prédire combien de temps cet individu va passer à se blottir contre ses congénères.

De la même manière, quel que soit le protocole expérimental que l'on pourrait imaginer (par exemple de placer le bébé rat au milieu de petits robots rat dont la température est plus basse que celle d'un rat biologique), du moment qu'il est question de température et de comportements de blotissement, notre modèle peut prédire quelque chose. Avec des robots, il prédirait par exemple que le bébé rat passera davantage de temps à se blottir car ce comportement produit ici un moins bon réchauffement pour une même durée.

***J'ai dessiné un schéma avec des boîtes et des flèches. Est-ce un modèle ?***

Avant de passer à une étape de formalisation mathématique, il est souvent commode de commencer par dessiner un schéma avec des boîtes et des flèches, pour décrire les échanges d'information que l'on conçoit entre différents composants du modèle. C'est un type de représentation que le lecteur a pu rencontrer dans plusieurs chapitres de cet ouvrage dès lors qu'il était question d'échange d'information entre aires cérébrales, par exemple.

Toutefois, si ce type de schéma n'est pas accompagné d'équations, on ne parlera pas de modèle computationnel mais plutôt de *modèle conceptuel*. En effet, un tel schéma ne dit rien du caractère quantitatif des opérations effectuées. Si par exemple on mettait en entrée de la première boîte une information qu'on suppose égale à 1, on ne peut pas dire quelle valeur de sortie on va retrouver au niveau de la dernière boîte. On ne sait pas si les flèches représentent une simple transmission d'information, ou une transmission avec amplification de l'information, et dans ce dernier cas on ne sait pas de quel facteur l'information doit être amplifiée. On ne peut donc pas écrire un programme qui *implémente* ce modèle ni le simuler pour déterminer quel est le résultat du calcul.

Des schémas de type boîtes-et-flèches peuvent néanmoins parfois être assortis d'informations numériques et de conventions qui font que l'on peut déduire comment traduire ces schémas en suite d'opérations numériques précises. Par exemple, un schéma de type *réseau de neurones* peut représenter les noeuds comme étant des neurones artificiels connectés entre eux, tout en indiquant sur chaque flèche le facteur de connectivité (on parle de *poids synaptique*) et en indiquant au niveau de chaque neurone quelle fonction de transfert $f(.)$ (sigmoïde, fonction à seuil ou fonction linéaire) est appliquée sur l'information (**Figure 2**). Dans ce cas, on sait par convention que la sortie d'un neurone donné du schéma va être égale à la somme des informations d'entrée (envoyées par d'autres neurones) pondérées

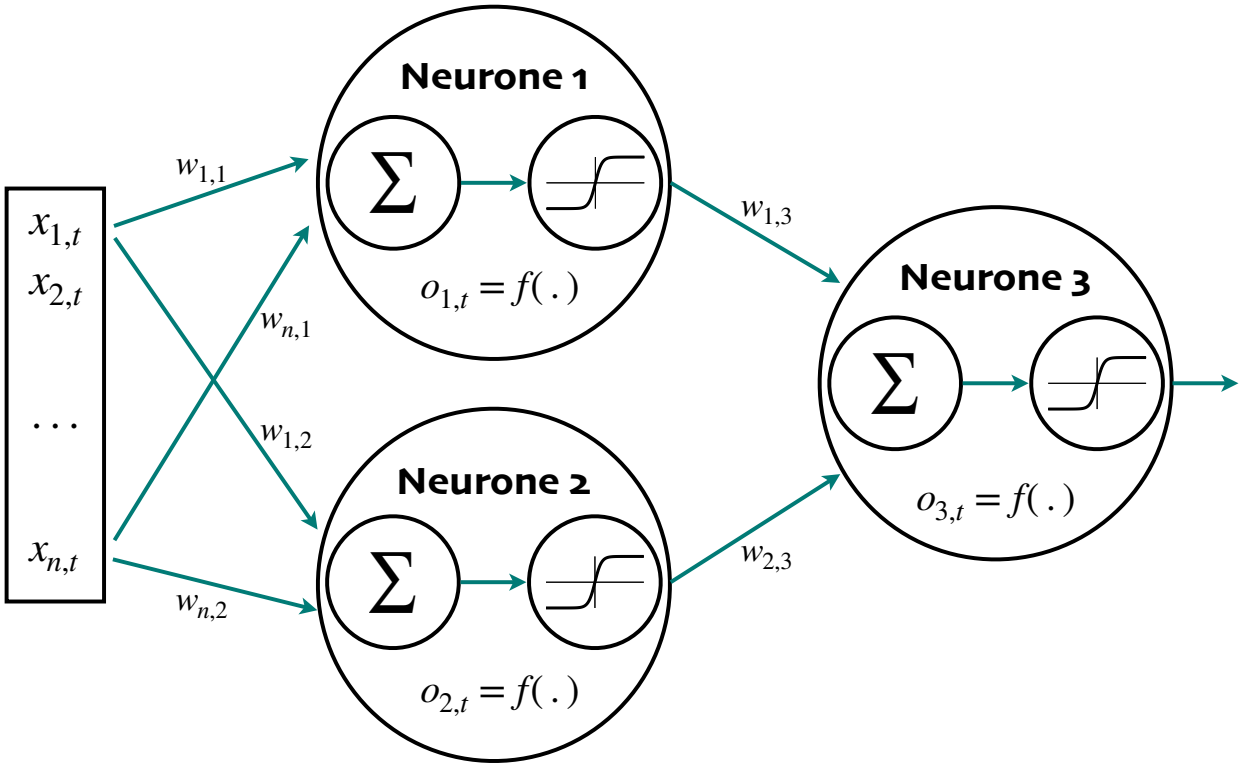


Figure 2: Schéma d'un réseau de neurones. Tous les neurones sont ici dotés de la même fonction de transfert $f(.)$ : une sigmoïde. Le vecteur d'entrée $x$ est pondéré par les poids synaptiques $w$ spécifiques à chaque neurone. La sortie $o_{i,t}$ à un instant $t$ du neurone $i$ est alors calculée comme étant $o_{i,t} = f(\sum_{k=1}^{n} w_{ki}x_{k,t})$. Pour le neurone 3, qui est le neurone de sortie du réseau, la même opération est effectuée non plus sur le vecteur d'entrée $x$ mais sur les sorties des neurones de la couche précédente : $o_{3,t} = f(w_{1,3}o_{1,t} + w_{2,3}o_{2,t})$.

chacune par le poids synaptique de la flèche par laquelle l'information est passée. On écrirait donc la sortie $o_{i,t}$ (o pour output) à un instant $t$ du neurone $i$ comme étant $o_{i,t} = f(w_{1i}x_{1,t} + w_{2i}x_{2,t} + .. + w_{ni}x_{n,t})$ qui s'écrit aussi $o_{i,t} = f(\sum_{k=1}^{n} w_{ki}x_{k,t})$, où $x_{k,t}$ est l'information émise par le neurone $k$ au temps $t$ (cette information est considérée ici comme une entrée du neurone $i$), et où $w_{ki}$ est le poids synaptique de la connexion qui relie le neurone $k$ au neurone $i$. On voit bien ici que si notre schéma contient un ensemble de neurones et qu'on connait les poids synaptiques qui les relient entre eux et les fonctions de transfert, alors à partir d'une information initiale introduite à l'entrée d'un premier neurone, on peut calculer comment l'information va circuler dans le réseau et quelle va être la sortie numérique émise par le dernier neurone. Ce type de modèle est à la base de nombreux travaux en neurosciences computationnelles.

Pour bien reconnaître un modèle computationnel, il faudra donc vérifier qu'il est bien constitué (1) d'un ensemble d'équations qui décrivent le processus supposé sous-jacent au phénomène observé, (2) d'un ensemble de paramètres de ces équations (dans l'exemple de la température et du blotissement, on avait une fonction linéaire dont les paramètres étaient la pente et l'ordonnée à l'origine ; dans l'exemple du réseau de neurones, les paramètres sont les poids synaptiques), (3) et d'un ensemble de variables qui représentent l'état du modèle à un instant $t$ donné, ces variables étant ce qu'on peut lire comme étant le résultat des opérations (la probabilitié d'aller se blottir est une variable du premier modèle, qui va changer au cours du temps en fonction de la température du corps et des paramètres ; la sortie d'un neurone est une des variables du deuxième modèle).

***Est-ce qu'un modèle peut être considéré comme juste s'il est validé par une expérience ?***

Malheureusement, non. Nous avons vu précédemment qu'un des grands intérêts des modèles computationnels est d'être prédictifs : on ne se contente pas de décrire ni synthétiser des données expérimentales existantes, mais on peut aussi faire des prédictions que l'on va ensuite pouvoir valider ou infirmer par une nouvelle expérience. Et nous avons vu que quand les résultats de l'expérience sont contradictoires avec la prédiction, alors on peut déduire que le modèle est faux. Mais l'inverse est-il vrai ? Que peut-on conclure si les résultats de l'expérience vont dans le même sens que les prédictions du modèle ?

Et bien on ne peut rien en conclure. Ou plutôt, on peut seulement en conclure que le modèle reste valide jusqu'à ce qu'une nouvelle expérience prouve le contraire. En effet, même si on peut dire qu'une prédiction d'un modèle est vérifiée, on ne peut jamais dire qu'un modèle est vrai, tout simplement parce qu'un modèle n'est pas la réalité qu'il est sensé décrire.

Pour comprendre cette subtilité, imaginons que l'on veuille concevoir le modèle d'un vélo. On aura beau le dessiner avec précision, indiquer son poids, son volume, ses caractéristiques mécaniques, il y aura toujours une information qui manquera. Par exemple, avec ce modèle on ne pourra pas représenter la fraîcheur du cadre du vélo, qu'un humain peut ressentir en le saisissant. Si l'on veut à chaque fois incorporer de nouvelles informations au modèle pour qu'il se rapproche de la réalité, à la fin on aboutira à un autre vélo (quasi) identique. Or justement, un vélo est le plus souvent fabriqué de façon industrielle, de sorte que de nombreux vélos (quasi) identiques ont été produits à la chaîne. C'est pour cela que c'est un exemple intéressant à considérer ici avant de se poser la question pour des choses biologiques. Du fait de ce caractère industriel, chaque vélo sorti de cette même chaîne de production sera un bon modèle des autres vélos qui vont en sortir. Un bon modèle, oui, mais pas un modèle vrai. Car il y aura toujours une micro imperfection qui fera une petite différence entre deux vélos, de sorte que le meilleur modèle d'un vélo donné, c'est le vélo lui-même.

On comprend bien que si un modèle ne peut être vrai pour un objet produit industriellement, ce sera a fortiori encore moins le cas pour un objet biologique. En d'autres termes, le meilleur modèle d'un cerveau est ce cerveau lui-même. Mais ça ne nous avance guère car on ne peut pas en déduire une meilleure compréhension du fonctionnement de ce cerveau. Donc même si un modèle du cerveau n'est jamais vrai, son intérêt résiderait dans le fait qu'il pourrait nous aider à mieux comprendre le fonctionnement du cerveau. Et pour qu'un modèle puisse nous aider à comprendre, il faut qu'il soit clair et simple. Clair, donc précis, ce qui est justement l'avantage que nous fournit le langage mathématique. Simple, donc simplifié par rapport à la réalité. Mais ce n'est pas grave du moment que cette simplification nous est utile pour comprendre la réalité.

Cette explication peut être résumée par une phrase du statisticien George Box : "Tous les modèles sont faux, mais certains sont utiles".

***Est-ce qu'une équation très simple suffit pour constituer un modèle ?***

Oui et c'est même souhaitable. Comme nous l'avons expliqué au paragraphe précédent, le but d'un modèle computationnel est d'être utile à notre compréhension du phénomène étudié. Et pour qu'il soit utile, un modèle doit être clair (donc précis) et simple.

Cet objectif de simplicité des modèles, aussi appelé *principe de parcimonie*, peut être bien résumé par la formule philosophique qu'on appelle *le rasoir d'Ockham*. L'idée est qu'il faut éliminer (avec le rasoir) toutes les explications improbables qui ajoutent de la complexité au modèle sans augmenter les capacités descriptives et prédictives du modèle. En d'autres termes, entre deux explications possibles, la meilleure est toujours la plus simple. Cela ne veut pas dire que le modèle le plus simple ne pourra pas être invalidé plus tard par de nouvelles données expérimentales. Cela ne veut pas dire non plus que le modèle plus compliqué, qui n'expliquait ni mieux ni moins bien les données précédentes, ne pourrait pas, lui, demeurer valide suite à sa confrontation à ces nouvelles données. Cela veut juste dire qu'à un instant donné, en considérant un ensemble de données qui sont aussi bien modélisées par un modèle A et un modèle B, celui des deux modèles qui a le plus de complexité dans ses mécanismes aura statistiquement moins de chances d'être vrai. Nous y reviendrons plus en détails dans la section méthode de ce chapitre.

Donc il est justement souhaitable de concevoir des modèles les plus simples possibles pour contribuer à la recherche scientifique. Si l'on revient sur l'exemple du temps de délibération dans les tâches de prise de décision, on peut justement concevoir des modèles simples qui peuvent faire des liens avec les données expérimentales qui ont été présentées au **chapitre 6**. Par exemple, considérons une tâche dans laquelle un individu doit choisir entre deux objets pour lesquels il a exprimé au préalable sa préférence lorsqu'il était exposé à chaque objet isolément (on connaît donc la *valeur subjective* de chacun de ces objets pour cet individu). Alors un modèle simple du temps de décision du sujet stipulera que plus les valeurs subjectives des deux objets sont proches, plus le sujet sera incertain, et donc plus il mettra de temps à choisir. Au contraire, si un des deux objets a une valeur très forte pour ce sujet, tandis que l'autre objet n'en a pas, alors le sujet n'hésitera pas ou peu avant de choisir. On peut donc modéliser ce processus décisionnel en considérant que le temps de délibération sera inversement proportionnel à la différence entre les valeurs. On peut imaginer que ce modèle a un paramètre multiplicatif qui pourra varier d'un sujet à l'autre et ainsi représenter le fait qu'à différence égale de valeurs, certains sujets hésiteront davantage que d'autres avant de choisir. On se retrouve donc avec un modèle très simple qui constitue une abstraction rudimentaire des *drift-diffusion models* qui ont été discutés au **chapitre 6**. Et ce modèle a été utile pendant de nombreuses années pour décrire et prédire de nombreux résultats expérimentaux. Ce n'est que récemment que de nouveaux travaux ont montré que le temps de délibération était non seulement proportionnel à la différence de valeurs, mais aussi à l'ordre de grandeur de ces valeurs (Steverson, Chung, Zimmermann, Louie, & Glimcher, 2019). En d'autres termes, si un sujet attribue comme valeur subjective 5 à un objet

A, 10 à un objet B, 50 à un objet C, et 55 à un objet D, même si la différence de valeur entre A et B est la même qu'entre C et D, le temps de délibération du sujet ne sera pas le même. Il s'avère au contraire que les humains ont tendance à prendre plus de temps pour décider entre des options de forte valeur, comme s'il y avait un plus grand enjeu.

Ceci nous amène à un dernier exemple de modèle très simple qui nous permettra de conclure cette introduction : la loi de Weber-Fechner, qui a révolutionné la psychologie expérimentale au milieu du $XIX^{eme}$ siècle en donnant naissance au champ de la psychophysique.

C'est en étudiant la relation entre la force d'un stimulus et l'intensité perçue de ce stimulus que le médecin Ernst Weber, puis le psychologue Gustav Fechner, ont découvert qu'il y avait une relation quantitative régulière que l'on pouvait formaliser par une équation mathématique. De façon très proche de l'exemple précédent, ils se sont rendus compte que la différence ressentie subjectivement entre les forces perçues de deux stimuli n'était pas la même selon la magnitude de ces stimuli. Par exemple, entre un objet d'1kg et un objet de 2 kg, on ressent une très grande différence, alors qu'entre des objets de 10kg et 11kg, on ne ressent quasi aucune différence. Pour que le ressenti soit le même, il faut que la différence soit multipliée par la magnitude, donc que l'on compare plutôt 10kg et 20kg. En d'autres termes, la différence des intensités ressenties divisée par la magnitude est constante, ce qui s'écrit : $\frac{(i_1 - i_2)}{i} = k$, qui s'écrit aussi $\frac{\Delta i}{i} = k$, où $i$ est l'intensité de référence (qui peut se calculer ici comme la moyenne entre les intensités $i_1$ et $i_2$), et $k$ est une constante, donc un paramètre du modèle. On peut alors écrire l'équation donnant l'intensité $i$ perçue d'un stimulus comme étant : $i = k \cdot log(s)$, où $s$ est la magnitude du stimulus, sa grandeur. On se retrouve donc avec une relation qui n'est pas linéaire, mais logarithmique.

Pour confirmer les prédictions de cette loi, des expériences ont été élaborées dans lesquelles on recherchait le *seuil différentiel* des sujets, c'est-à-dire la plus petite différence d'intensité perçue, que l'on écrit $\Delta i$. Ces expériences ont été appliquées avec des stimuli de nature et de modalité différentes. C'est par cette démarche qu'on considère que Gustav Fechner a fondé le domaine de la psychophysique, avec pour objectif d'établir des relations quantitatives entre stimuli et cognition (notamment dans le cadre de la perception). C'est justement l'intérêt des modèles computationnels de formaliser ces relations quantitatives, apportant un pouvoir prédictif. Ils sont de plus en plus utilisés en psychologie et en neurosciences, ce qui a donné naissance à des domaines plus récents comme les neurosciences computationnelles.

## 2 Un exemple approfondi : la théorie des perspectives

Dans cette deuxième partie, nous allons approfondir autour d'un exemple lié à la prise de décision chez l'humain : la théorie des perspectives. Ceci nous permettra à la fois de faire des liens avec les résultats expérimentaux présentés dans le **chapitre 6** de cet ouvrage, et à la fois de mieux comprendre certains succès récents dans la modélisation des phénomènes d'apprentissage par renforcement, tels que décrits dans les parties suivantes de ce chapitre.

### 2.1 Prise de décision fondée sur les valeurs en économie

Comme le lecteur aura pu le constater tout au long des différents chapitres de cet ouvrage, il existe une variété de mécanismes de prise de décision dans notre cerveau, qui concernent différents domaines dans lesquels les humains peuvent avoir à décider : perception, attention, mouvement, action, navigation, mémorisation, raisonnement ou même langage. Dans la suite de ce chapitre, nous allons nous focaliser sur un sous-ensemble de décisions bien précises qu'on appelle *les décisions fondées sur les valeurs* (*value-based decision-making* en anglais). Cela concerne les situations où, pour décider parmi 2 ou N options (que ces options soient des options d'actions, comme choisir entre aller au musée ou au cinéma, ou des options d'objets, comme choisir parmi différents produits à manger), nous commençons par attribuer des valeurs subjectives à ces options, puis comparons ces valeurs, et le plus souvent nous choisissons l'option qui a la plus forte valeur subjective pour nous.

Cette vision des choses est historiquement issue du champ de la recherche en économie, et est ensuite venue nourrir les sciences cognitives pour donner naissance à un domaine récent de recherche qu'on appelle *la neuro-économie* (voir le **chapitre 2** pour une perspective historique). En effet, la recherche en économie s'intéresse à comprendre dans quelle mesure les humains sont des agents rationnels qui choisissent des options qui leur semblent les meilleures pour eux (ou pour la collectivité) afin d'en récolter un certain bénéfice (le plus souvent, un gain monétaire) dans un futur plus ou moins proche et avec une probabilité plus ou moins forte.

Mais il est très important de comprendre que cette formalisation des décisions à partir des valeurs ne se restreint pas au domaine des gains monétaires. Il a été depuis démontré que les mêmes propriétés comportementales et les mêmes mécanismes cérébraux sont en partie observés pour des décisions dont les buts peuvent être différents, comme par exemple l'obtention de nourriture, ou encore d'une récompense sociale. Cette généricité du système cérébral des valeurs se comprend bien si l'on conçoit que le cerveau doit pouvoir mettre en comparaison différents types d'options afin de les classer selon les préférences du sujet avant de décider. Si l'on imagine que la préférence pour une option A sur

une option B est représentée dans le cerveau par une activité électrique plus forte pour A que pour B, alors on peut considérer abstraitement que l'intensité de ces activités cérébrales représente la valeur subjective de chaque option. Ceci est illustré par divers résultats expérimentaux mentionnés au **chapitre 6**.

En économie expérimentale, on confronte souvent des sujets humains à ce qu'on appelle des *loteries* (*gambles* en anglais). Par exemple, on leur demande : "**loterie 1** : est-ce que vous préférez avoir 80% de chances de gagner 4000€, et 20% de chances de ne rien gagner **(option A)**, ou bien être sûr de gagner 3000€ **(option B)** ?" Avant de lire la suite, demandez-vous quelle option vous choisiriez si la possibilité de gain était réelle. Ça y est vous avez choisi ? Alors continuons : un rapide calcul nous dit que l'espérance mathématique de l'option A vaut 0,8 x 4000 + 0,2 x 0 = 3200€. Donc un agent rationnel devrait préférer l'option A. Et pourtant, la plupart des gens préfèrent l'option B. En effet, il semble souvent plus raisonnable d'empocher de façon sûre 3000€ que de prendre 20% de risques de repartir les mains vides. On appelle ce phénomène psychologique *l'aversion au risque*. Pourtant un agent purement rationnel devrait choisir l'option A car celle-ci a une espérance mathématique plus forte. Cela veut dire que si on est exposé de façon répétée à un choix comme la loterie 1 et qu'on prend l'option ayant la plus forte espérance, on gagnera davantage en moyenne. C'est donc le comportement que stipule la théorie économique classique.

Considérons maintenant un deuxième exemple : "**loterie 2** : est-ce que vous préférez avoir 80% de chances de perdre 4000€, et 20% de chances de ne rien perdre **(option C)**, ou bien être sûr de perdre 3000€ **(option D)** ?" Avant de lire la suite, demandez-vous quelle option vous choisiriez si la possibilité de gain était réelle. Bon maintenant c'est un peu biaisé, car vous connaissez l'astuce. Mais essayez de répondre en suivant votre intuition initiale. Ça y est vous avez choisi ? Alors on peut y aller : comme vous vous en doutez maintenant, on se doit de calculer l'espérance mathématique de l'option C. Celle-ci vaut -3200€. Donc d'un point de vue normatif, il vaut mieux choisir l'option D. Et pourtant, la plupart des gens préfèrent l'option C car ils préfèrent se donner ne serait-ce qu'une petite chance d'éviter une perte. On appelle ce phénomène *l'aversion à la perte*.

Vous voyez grâce à ces deux exemples de loteries qu'il y a une asymétrie dans la façon dont la psychologie humaine considère les gains et les pertes monétaires. Nous avons tendance à avoir une aversion pour le risque dans le domaine des gains, tandis que nous semblons rechercher le risque dans le domaine des pertes. D'une façon générale, il s'avère de plus que les humains ont tendance à surestimer la probabilité d'occurrence des événements rares (*i.e.,* les 20% de chances dans les options A et C), et à sous-estimer la probabilité des événements très probables (non illustré ici).

### 2.2 Un modèle computationnel

Nous allons maintenant voir comment on peut modéliser la tendance qu'ont les humains à choisir parmi des loteries.

On attribue souvent les premiers raisonnements probabilistes à Blaise Pascal, philosophe, mathématicien et physicien français du XVII$^e$ siècle (avec notamment son fameux pari). Néanmoins, le calcul de l'espérance mathématique que nous avons effectué juste au-dessus a été formalisé au milieu du XX$^e$ siècle par deux mathématiciens américains : John von Neumann et Oskar Morgenstern. Ceux-ci ont proposé la *théorie des utilités espérées* (*expected utility theory* en anglais) stipulant qu'un agent rationnel doit chercher à maximiser l'utilité $U(X)$ d'une certaine variable aléatoire $X$.

Le résultat d'une loterie est ici considéré comme une variable aléatoire. Mais contrairement à l'emploi du terme "aléatoire" dans le langage commun en français, cela ne veut pas dire que cette variable se comporte de manière complètement aléatoire. Cela veut simplement dire qu'elle a un caractère probabiliste et non déterministe. L'utilité se calcule alors selon l'équation :

$$U(X) = \mathbb{E}[X] = \sum_{i=1}^{n} p_i x_i \tag{1}$$

où $x_1$, $x_2$, ..., $x_n$ sont les différentes valeurs que peut prendre la variable aléatoire $X$, et $p_1$, $p_2$, ..., $p_n$ sont les probabilités associées à chacune de ces valeurs, avec $p_1 + p_2 + ... + p_n = 1$. $\mathbb{E}[X]$ s'appelle l'espérance mathématique, ou simplement l'espérance, de $X$.

Mais contrairement aux résultats des loteries mentionnées à la section précédente, ce modèle rationnel prédit une symétrie parfaite entre les gains et les pertes. De plus, il considère les probabilités absolues, sans les sur-estimer ni les sous-estimer. Ce modèle considérerait donc que la valeur subjective que les humains construisent face à ces loteries est égale à l'utilité théorique telle que donnée par l'équation 1 ci-dessus, donc une fonction parfaitement linéaire des gains et des pertes.

Pour rendre compte des phénomènes d'aversion au risque et d'aversion à la perte que les sujets humains montrent face à des loteries, les psychologues Daniel Kahneman et Amos Tversky proposèrent un modèle alternatif en 1979, qu'ils appelèrent la *théorie des perspectives* (*prospect theory* en anglais). L'idée est que l'agent considère la valeur subjective pour lui des valeurs possibles $x_i$ que peut prendre la variable aléatoire $X$, par le biais d'une fonction $v(x_i)$ qui "déforme" l'utilité théorique pour créer une asymétrie gain/perte. De même, l'agent ne considère pas les probabilités

absolues $p_i$ mais une certaine fonction $\pi(p_i)$ qui donne un poids gonflé aux probabilités faibles (les événements rares) et un poids réduit aux probabilités proches de 1 (les événements très probables). L'agent prend donc sa décision non plus sur la base de son calcul de l'utilité théorique, mais à partir de la valeur subjective calculée selon :

$$V(X) = \sum_{i=1}^{n} w(p_i)v(x_i) \tag{2}$$

La **Figure 3** illustre une forme que peut prend cette fonction, et en particulier l'asymétrie entre la manière dont la valeur subjective considère les gains et les pertes.

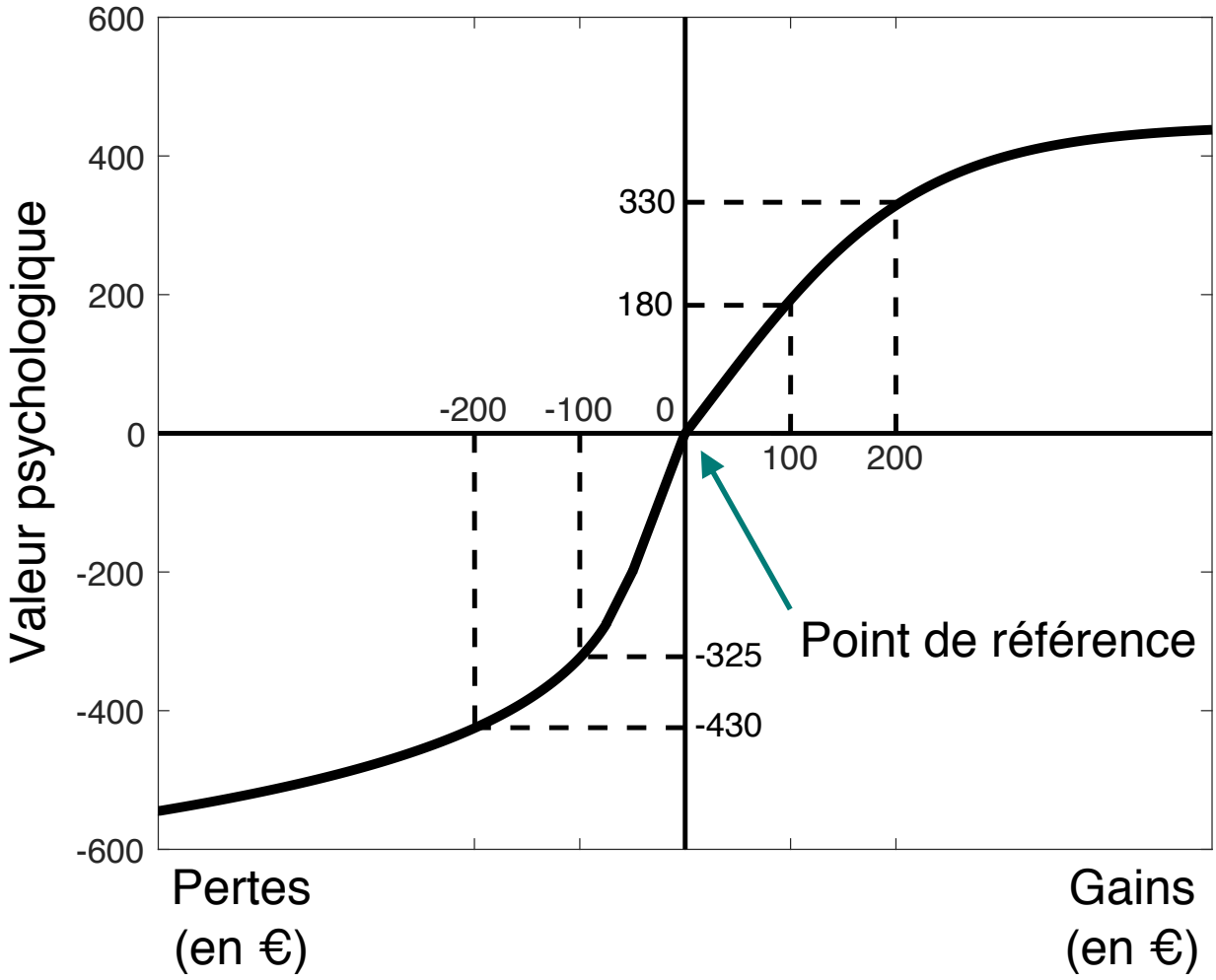


Figure 3: Illustration d'une fonction de valeur subjective qui résume l'asymétrie gains/pertes dans les choix économiques telle qu'elle a été formalisée dans la *Théorie des perspectives (prospect theory)* par (Kahneman & Tversky, 1979). Par exemple, une perte de 100€ sera ressentie plus fortement qu'un gain de 100€. Le code source pour construire cette figure est disponible à l'adresse `https://github.com/MehdiKhamassi/basicModelingTutorial`.

Dans ce type de formalisation, $V$ réfère à la valeur subjective que l'agent attribue à la variable aléatoire $X$, et est également souvent appelée *la fonction d'utilité* de l'agent. En d'autres termes, on considère que l'agent cherche à maximiser son estimation subjective de l'utilité pour lui-même de ses décisions. Comme nous allons le voir par la suite, ce type de formalisme est à la base des modèles de prise de décision. Que les agents soient modélisés comme prenant des décisions rationnelles ou non, on pourra représenter leurs valeurs subjectives estimées sous la forme d'une fonction d'utilité.

Dans les modèles d'apprentissage par renforcement présentés par la suite, l'utilité subjective $V$ d'une option peut être augmentée ou diminuée par apprentissage, en fonction des expériences passées avec cette option. Cette valeur

subjective construite au cours du temps de la vie d'un individu peut contribuer à expliquer pourquoi deux individus ne prendront pas les mêmes décisions face à la même offre d'options.

## 3 Un exemple de succès - l'apprentissage par renforcement

Un autre exemple de succès remarquable de la modélisation computationnelle en psychologie et en neurosciences ces dernières années concerne le domaine de *l'apprentissage par renforcement*. Les modèles d'apprentissage par renforcement décrivent comment l'exposition répétée à des stimuli prédicteurs de récompense conduit à un renforcement de leur valeur, de sorte que ces stimuli deviennent attractifs. De même, effectuer plusieurs fois une même action systématiquement suivie de récompense conduit au renforcement de sa valeur, de sorte que nous avons plus de chances d'effectuer cette action à nouveau dans le même contexte.

Ces idées trouvent leur origine il y a plus d'un siècle chez des behavioristes tels Pavlov et Skinner, qui avaient observé les effets du conditionnement lié à l'obtention de récompenses ou de punitions. Rescorla et Wagner ont ensuite proposé un des premiers modèles mathématiques expliquant plusieurs phénomènes de conditionnement. Ces théories ont depuis été largement étendues pour mieux décrire 1) l'apprentissage de manière plus générale, car il s'avère que les mécanismes d'apprentissage par renforcement se généralisent à de nombreuses autres situations d'apprentissage que le seul conditionnement ; 2) comment certaines régions du cerveau implémentent (*i.e.,* mettent en oeuvre) ces processus d'apprentissage.

### 3.1 Les origines - conditionnement pavlovien, Rescorla-Wagner.

La fameuse expérience de Pavlov montre un exemple d'apprentissage par renforcement appelé *le conditionnement Pavlovien*. Un chien est soumis plusieurs fois à l'expérience d'entendre une cloche sonner, suivie immédiatement par l'arrivée de nourriture. Petit à petit, l'animal se met à saliver quand il entend la cloche, montrant qu'il est capable de prévoir l'arrivée d'un événement récompensant suivant ce son : la nourriture. Ceci révèle donc l'apprentissage d'une association entre un stimulus (le son de cloche) et une récompense (la nourriture).

Comment pourrait-on modéliser cet apprentissage, caractérisé par une augmentation de la salivation en réponse a la cloche ?

L'idée originale est simple : on définit une variable $V$ qui décrit la force de l'association entre un stimulus neutre (US: *Unconditioned Stimulus*, ici la cloche), et un stimulus appétitif (CS: *Conditioned Stimulus*, la nourriture). À chaque fois que l'animal fait l'expérience d'un US après la présentation d'un CS, cette valeur d'association est augmentée : $V(t+1) = V(t) + increment$. Quel incrément fait sens ? Si l'incrément est une valeur fixe, la valeur d'association pourrait augmenter infiniment, ce qui n'est pas raisonnable (la quantité de salive produite par le chien atteint un plateau pendant l'apprentissage). Une option plus raisonnable consiste à définir une valeur maximale pour la valeur d'association, par exemple 1, dont on considère qu'elle représente ce plateau. Puis à chaque essai, on incrémente l'association en proportion de la distance à cette valeur maximale, $1 - V(t)$. En d'autres termes, à mesure que $V$ s'approchera de 1, on incrémentera $V$ de plus en plus faiblement, de façon à la faire *converger* petit à petit vers 1. Cela nous donne l'équation suivante :

$$V(t+1) = V(t) + \alpha \times (1 - V(t))$$

Dans cette équation, $\alpha$ est un paramètre qui prend une valeur entre 0 et 1, et indique à quelle vitesse on apprend : si $\alpha = 0$, on n'apprend pas du tout ; si $\alpha = 1$, on atteint la valeur maximale après une seule expérience de la cloche suivie de la nourriture. Le problème d'une vitesse rapide d'apprentissage est qu'à la moindre absence fortuite de récompense on désapprendra d'un coup la valeur $V$. Une valeur intermédiaire de $\alpha$ est donc plus probable. L'animal a ainsi besoin de plusieurs répétitions avant d'être confiant que le son de cloche est annonciateur de récompense.

Que se passe-t-il quand l'expérience change, et que le chien continue à être exposé à la cloche, mais sans que celle-ci ne soit suivie par de la nourriture ? Dans cette phase dite *d'extinction*, l'association est désapprise petit à petit par le chien, qui arrête progressivement de saliver quand il entend la cloche. On peut rendre compte de ce phénomène par une équation très similaire : cette fois, la valeur d'association diminue à chaque essai où l'US n'est pas présente, avec une valeur minimale fixée cette fois-ci à 0 : $V(t+1) = V(t) + \alpha \times (0 - V(t))$. Ainsi, il faudra plusieurs absences de récompense après la présentation du son de cloche pour que l'animal considère que celui-ci n'est plus un prédicteur fiable de récompense.

Il apparaîtra à un oeil avisé que ces deux équations ont la même structure. Seule la valeur entre parenthèse change : 1 quand il y a une récompense, 0 sinon. On peut donc ré-écrire ces deux équations ensemble, en une seule équation, de la manière suivante:

$$V(t+1) = V(t) + \alpha \times (r(t) - V(t))$$

Où $r(t)$ indique le *renforcement*, ou la valeur appétitive de ce qui a été observé après l'US : 1 pour la nourriture, 0 pour l'absence de nourriture. Cette équation simple, parfois appelée *delta-rule*, est au coeur des modèles d'apprentissage par renforcement. Elle décrit une variable importante : l'erreur de prédiction $\delta = r(t) - V(t)$ qui peut être interprétée comme la différence entre une récompense obtenue $r(t)$ et une récompense prédite $V(t)$. Une forte erreur de prédiction conduit à changer la valeur d'association estimée plus qu'une faible erreur de prédiction (**Figure 4**).

Le modèle de Rescorla-Wagner (RW), proposé dans les années 1970, étend ce modèle simple pour rendre compte d'autres phénomènes bien établis dans le domaine du conditionnement. Par exemple, le phénomène d'*overshadowing* montre que si deux stimuli de salience différente sont présentés simultanément (par exemple un flash de lumière et un son faible), le stimulus le plus saillant produit un conditionnement plus fort que le stimulus le moins saillant. Le modèle RW est capable de produire cet effet en faisant l'hypothèse que le taux d'apprentissage est lié à la salience.

Un autre phénomène intéressant est celui du *blocking*, qui montre que si l'animal a déjà appris que la cloche est suivie de nourriture, lui présenter simultanément la cloche et un nouveau stimulus aussi saillant (par exemple une lumière), suivis de l'US, ne conduira pas à un conditionnement du nouveau stimulus : on dit que son apprentissage est bloqué. Intuitivement, l'idée est que la cloche prédit déjà le renforcement. Il n'y a donc plus besoin d'apprendre (**Figure 4**). L'équation précédente ne permet pourtant pas de reproduire cet effet. Pour ce faire, le modèle RW doit tenir compte de la prédiction faite par l'ensemble des stimuli, et non plus seulement par chaque stimulus indépendamment. Au final, le modèle RW consiste en une équation qui a la même structure que précédemment : incrémenter la valeur d'association d'un stimulus en fonction de l'erreur de prédiction, où la prédiction est simplement la somme des valeurs d'association des stimuli présents :

$$V_{lumiere}(t+1) = V_{lumiere}(t) + \alpha\beta_{lumiere} \times (r(t) - [V_{lumiere}(t) + V_{cloche}(t)])$$

$$V_{cloche}(t+1) = V_{cloche}(t) + \alpha\beta_{cloche} \times (r(t) - [V_{lumiere}(t) + V_{cloche}(t)])$$

Ici, $\beta_{stimulus}$ est un paramètre qui indique la salience du stimulus.

Le modèle de Rescorla-Wagner capture beaucoup d'aspects du conditionnement et était une preuve du succès potentiel des méthodes computationnelles en psychologie. Cependant, il a aussi un certain nombre de limites : par exemple, ce modèle ne peut pas rendre compte d'un phénomène important, le *second-order conditioning*, où l'animal est exposé à une lumière suivie du son d'une cloche, et apprend que la lumière prédit de la nourriture, puisqu'elle prédit la cloche qui elle-même prédit la nourriture. Les modèles d'apprentissage par renforcement plus modernes, que nous présentons dans la section suivante, étendent ce modèle historique à des domaines plus vastes (dont l'apprentissage instrumental et la planification de l'action).

### 3.2 Apprentissage par renforcement - les bases

La théorie de l'apprentissage par renforcement (RL, pour *Reinforcement Learning* en anglais) (Sutton & Barto, 2018) propose un ensemble d'algorithmes qui résolvent des problèmes où un agent apprend un comportement en utilisant un signal de "renforcement" de l'environnement : soit une récompense, soit une punition. En particulier, ces algorithmes peuvent se trouver dans différents *états* $s$ (tels que la présentation des différents stimuli de Pavlov), répondre par différents comportements - les *actions* $a$ (telle picorer un levier pour les pigeons de Skinner), et obtenir un renforcement $r$ (le même $r(t)$ que nous avons vu dans les équations précédentes).

En tant que généralisation du modèle de RW, les modèles de RL n'estiment pas seulement l'*association* entre un état et l'US qui le suit juste après. Ils essaient au contraire d'estimer la *valeur future cumulée* d'un état, $V(s)$, qui correspond à l'ensemble des renforcements que je m'attend à recevoir dans le futur, pas de temps après pas de temps, à partir de l'état actuel. Dit autrement, les modèles de RL posent la question : si je suis dans un état $s$, quelles récompenses et punitions vais-je recevoir immédiatement, puis à l'essai suivant, puis à l'essai d'après, etc. ? Ceci permet au modèle d'apprendre à prédire la récompense, même quand elle arrive après un certain délai. Les récompenses des essais futurs sont alors pondérées moins fortement que les récompenses immédiates. En anglais on dit que les récompenses futures sont "discounted" (dévaluées) par un paramètre de dévaluation $\gamma$ : on considère la somme $R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots$, avec $0 \leq \gamma \leq 1$. Si $\gamma = 0$, seulement les récompenses immédiates sont prises en compte. Si $\gamma = 1$, une récompense dans dix jours a le même effet qu'une récompense immédiate. En général, une valeur intermédiaire est appropriée.

Pour estimer la valeur d'un état, les algorithmes de RL peuvent utiliser une simple extension du modèle RW. Comme dans RW, on incrémente la valeur estimée en proportion d'une erreur de prédiction, avec un taux d'apprentissage $\alpha$ contrôlant cette proportion :

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

Pour tenir compte de l'estimation non seulement de la récompense immédiate, mais aussi de la récompense future, l'algorithme TD-RL, dit de *différence temporelle* (*temporal-difference* reinforcement learning), compare la prédiction

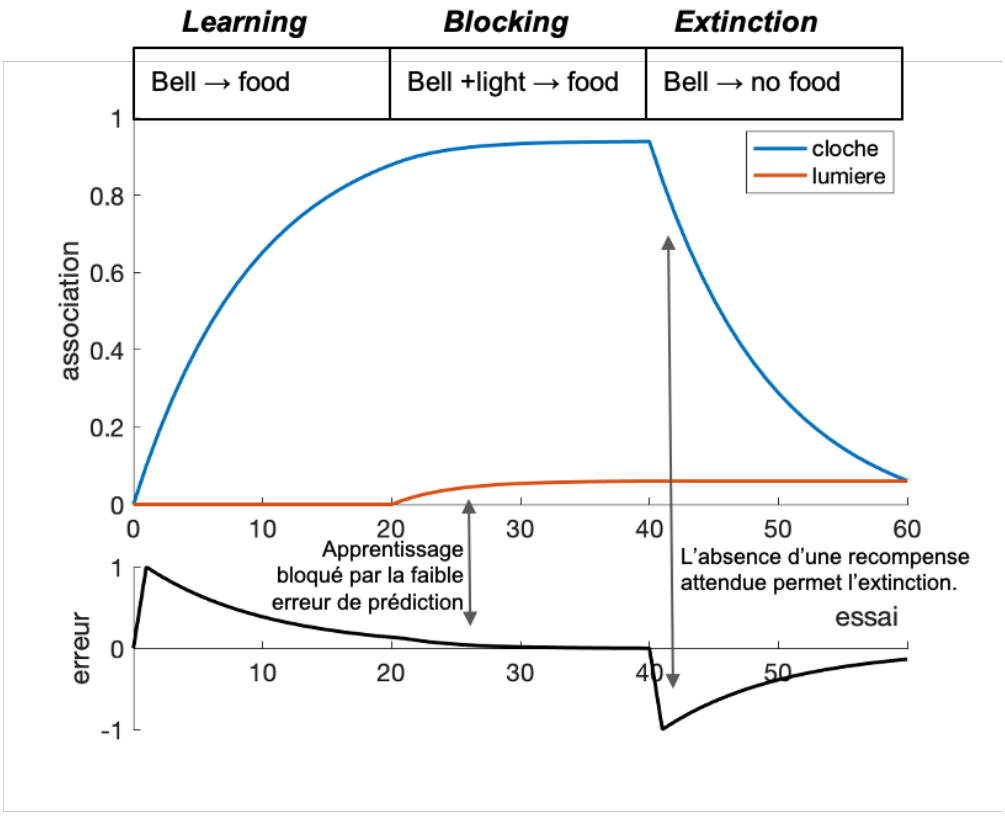


Figure 4: Simulation du modèle de Rescorla-Wagner. Pendant les 20 premiers essais, l'animal apprend à associer la cloche à la nourriture - la valeur d'association augmente. Pendant les 20 essais suivants, une lumière est présentée en même temps que la cloche sonne, et suivie de nourriture. Néanmoins, l'animal n'apprend pas que la lumière prédit la nourriture (sa valeur d'association reste faible), parce que l'erreur de prédiction est faible : c'est le phénomène de blocking. Pendant les 20 derniers essais, la cloche sonne, mais n'est pas suivie de lumière. L'animal désapprend son association. Le code source pour construire cette figure est disponible à l'adresse `https://github.com/MehdiKhamassi/basicModelingTutorial`.

$V(s_t)$ non seulement à la récompense obtenue $r_t$, mais aussi à la valeur de l'état suivant $V(s_{t+1})$. Ainsi, l'erreur de prédiction TD est :

$$\delta_t = (r_t + \gamma V(s_{t+1})) - V(s_t).$$

Cette extension permet de tenir compte du futur, et résout par exemple les problèmes de RW pour le phénomène de conditionnement de deuxième ordre (**Figure 5**).

Si l'agent a une bonne estimation de la valeur des états, il peut sélectionner une action qui l'amènera au meilleur des états suivants - une stratégie appelée *gloutonne* (*greedy* en anglais). Cependant, des théorèmes prouvent que l'algorithme ci-dessus n'est capable d'estimer correctement la valeur des états que si l'agent explore suffisamment : c'est-à-dire si l'agent choisit régulièrement une action qui a une valeur estimée non optimale. Cette exploration permet de s'assurer que tout l'espace des états est appris, et que l'agent ne s'est pas limité à un maximum local.

Le degré d'exploration est souvent modélisé de manière stochastique : la stratégie *greedy* est choisie avec une certaine (forte) probabilité $p$, et d'autres actions sont choisies le reste du temps (c'est-à-dire avec une probabilité $1-p$). Plusieurs formules sont possibles, toutes parametrisées par un certain degré de stochasticité. Mais nous ne détaillerons pas ces formules ici et suggérons quelques pistes de lecture en fin de chapitre pour aller plus loin.

### 3.3 Dopamine et troubles psychiatriques

Les modèles d'apprentissage par renforcement peuvent être vus comme des solutions aux niveaux *computationnel* et *algorithmique* de Marr (tels que présentés en introduction de ce chapitre). Au niveau computationnel, ils proposent que l'apprentissage a pour but d'estimer et d'optimiser les renforcements futurs cumulés. Au niveau algorithmique,

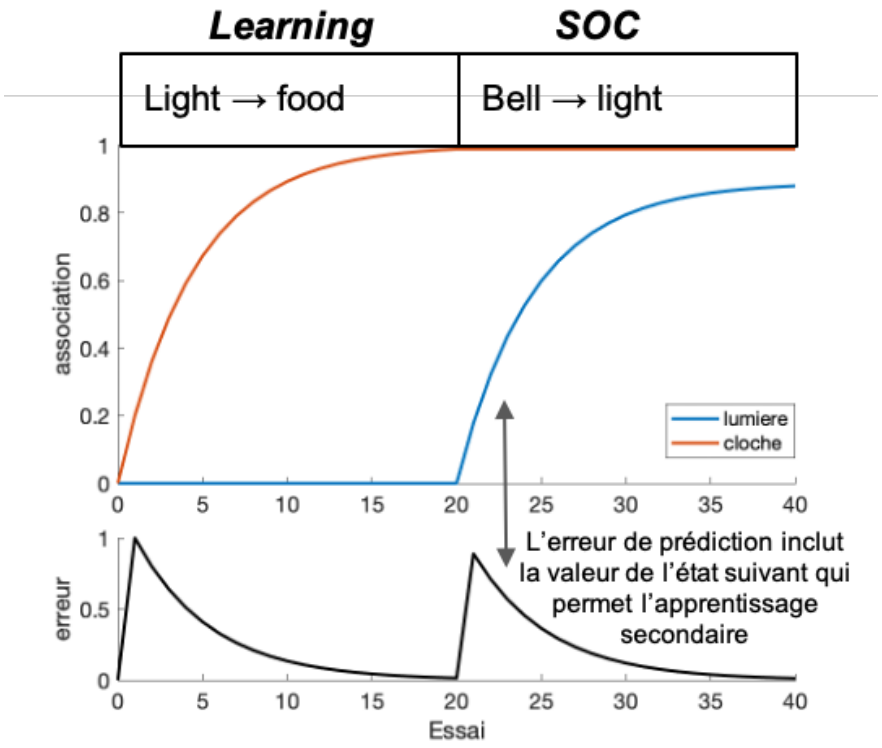


Figure 5: Simulation d'un modèle d'apprentissage par renforcement TD avec un conditionnement de deuxième ordre. Après 20 essais où l'agent apprend à associer la cloche avec une récompense, l'agent fait l'expérience d'une lumière suivie de la cloche. La valeur acquise de la cloche provoque une erreur TD, qui renforce la valeur de la lumière. Le code source pour construire cette figure est disponible à l'adresse `https://github.com/MehdiKhamassi/basicModelingTutorial`.

ils proposent une solution simple pour ce calcul, qui permet de très bien rendre compte de beaucoup de phénomènes psychologiques. Une des raisons du succès majeur des algorithmes RL est qu'ils sont aussi un bon modèle au troisième niveau de Marr : le niveau implémentationnel, c'est-à-dire ce qui se passe au niveau physiologique dans le cerveau pendant ce type d'apprentissages. Dans cette section, nous expliquons comment ce modèle représente la façon dont une partie du cerveau apprend.

Pour cela, il nous faut parler de la *dopamine*. Celle-ci est un neuromodulateur, c'est-à-dire une molécule biochimique neurotransmettrice qui régule l'activité de populations de neurones dans le cerveau. Elle est étudiée en neurosciences depuis plus d'un siècle. Elle est produite par les neurones dopaminergiques, situés dans des noyaux profonds du cerveau appelés la substance noire. Ces neurones dopaminergiques sont connectés massivement au reste du cerveau (en particulier au striatum et au cortex préfrontal) de façon à pouvoir y moduler l'activité nerveuse grâce à la dopamine émise. Enfin, des travaux dans de nombreuses équipes de recherche en neurosciences ont montré que la dopamine joue un rôle dans de nombreuses fonctions différentes : plaisir, récompense, effort, motivation, mais également mouvement, etc.

Dans les années 90, une série d'expériences ont été menées dans l'équipe du neurobiologiste Wolfram Schultz, à l'Université de Fribourg en Suisse, pour étudier le rôle de la dopamine dans l'apprentissage par renforcement. Ces travaux ont montré que les réponses phasiques des neurones dopaminergiques (c'est-à-dire des réponses ponctuelles et non pas soutenues dans le temps) ressemblaient fortement à un calcul très particulier : celui d'une erreur de prédiction de la récompense, telle que définie par l'algorithme TD-RL présenté plus haut (Schultz, Dayan, & Montague, 1997). Plus précisément, ces expériences mesuraient l'activité des neurones dopaminergiques dans la substance noire dite pars compacta de singes macaques pendant que ces singes étaient soumis à une expérience de conditionnement pavlovien. Initialement, les neurones ne répondaient qu'à l'apparition soudaine et inattendue de la récompense. Mais après le conditionnement, l'activité des neurones dopaminergiques avait changé : ceux-ci étaient fortement activés au moment de l'arrivée imprévisible du CS (un son), mais n'étaient plus activés au moment de l'obtention de la récompense (comme si celle-ci était maintenant parfaitement prévisible et ne générait donc plus d'erreur de prédiction). Ce profil d'activité

est cohérent avec une erreur de prédiction : avant d'avoir appris, le singe est surpris par l'arrivée de la récompense, ce qui entraîne une erreur de prédiction positive. Après avoir appris, le singe est surpris par l'arrivée du stimulus (car aucun signal n'annonce sa présentation), ce qui entraîne une erreur de prédiction (même en l'absence de renforcement) parce que la *valeur future* de ce stimulus n'était pas prédite avant qu'il soit présenté. La récompense, par contre, est parfaitement prédite (le singe sait qu'elle arrive à peu près toujours au bout du même délai après l'apparition du stimulus), et ne suscite donc pas d'erreur de prédiction. Enfin, un dernier cas permet de parachever cette démonstration : celui de l'omission. Dans cette situation, après apprentissage, on présente le CS à l'animal, qui s'attend donc à recevoir une récompense, mais la récompense n'est pas donnée. Le modèle prédit alors une erreur de prédiction négative, car la situation est moins bien que prévue. C'est exactement ce qui a été observé au niveau de la réponse des neurones dopaminergiques : ceux-ci montrent une baisse d'activité au moment où l'animal s'attendait à recevoir une récompense mais n'en a pas obtenu (**Figure 6**).

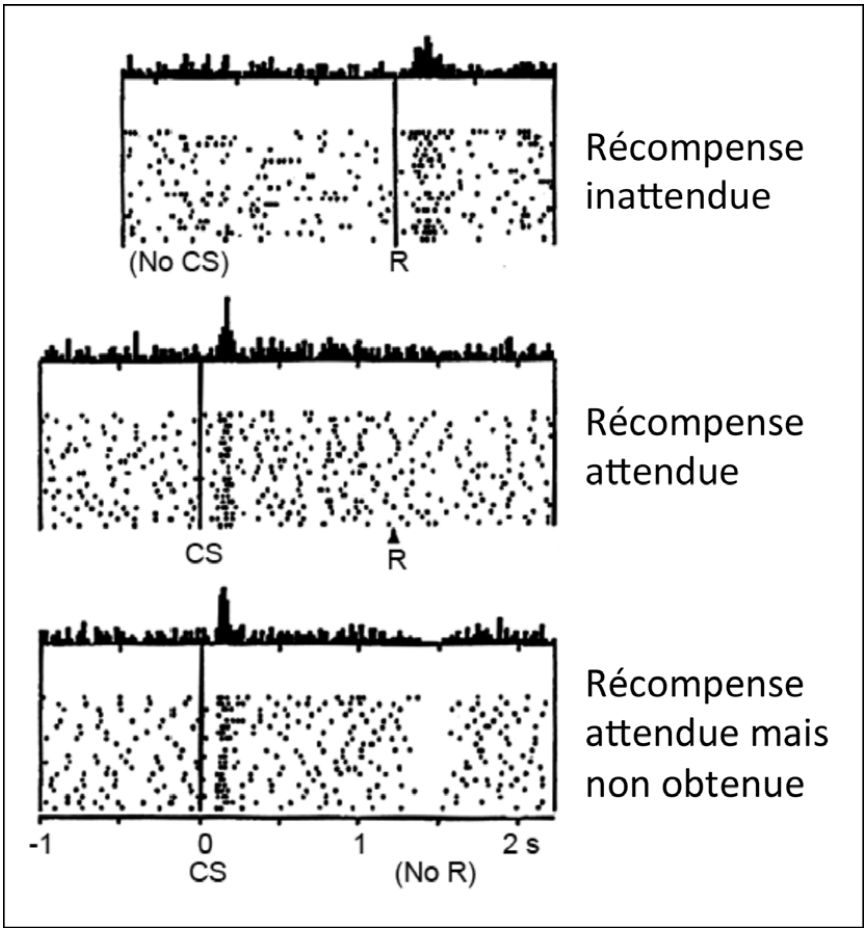


Figure 6: Illustration du patron d'activité phasique des neurones dopaminergiques enregistrés par (Schultz et al., 1997) chez le primate non-humain pendant une tâche de conditionnement Pavlovien. Lorsqu'il n'y a pas de Stimulus Conditionné (CS), la récompense est imprévisible et les neurones dopaminergiques y répondent. Après conditionnement, la récompense est attendue et les neurones dopaminergiques n'y répondent plus. En revanche, ils répondent au CS car celui-ci est inattendu tout en étant annonciateur de récompense. Enfin, en cas d'omission après présentation du CS, les neurones dopaminergiques montrent une diminution d'activité à peu près au moment où l'animal s'attendait à obtenir une récompense. Figure tirée de Pessiglione, M. (2014). Comment le cerveau motive le comportement: du circuit de la récompense au système des valeurs. Bulletin de l'Académie Nationale de Médecine, 198(7), 1283-1296.

Les résultats de cette expérience sont considérés comme une des découvertes majeures des neurosciences cognitives de ces dernières années, en plus de la découverte des cellules de lieux de l'hippocampe et des cellules de grille du cortex entorhinal, qui jouent un rôle dans la navigation et la localisation dans l'espace (voir le **chapitre 5**), et de la découverte des neurones miroirs qui jouent un rôle important dans la représentation des actions d'autrui, et donc de la théorie de l'esprit (voir le **chapitre 4**).

Depuis cette série d'expériences séminales sur les neurones dopaminergiques chez le primate non humain, de nombreuses autres ont permis de comprendre le fonctionnement de tout un réseau d'aires cérébrales qui semble utiliser ce signal dopaminergique pour apprendre la valeur des états et des actions, et pour implémenter une forme d'apprentissage par renforcement. En particulier, on trouve dans le striatum des neurones dont l'activité semble représenter la valeur estimée des stimuli présentés (Khamassi, Mulder, Tabuchi, Douchamps, & Wiener, 2008), ce signal étant ensuite envoyé aux neurones dopaminergiques dans la substance noire pour leur permettre de calculer une erreur de prédiction. En retour, le signal dopaminergique est renvoyé au striatum où il module la plasticité synaptique (Steinberg et al., 2013). Ceci permet de renforcer les connections cortico-striatales qui soutiennent un choix menant à une récompense, et à diminuer les autres. On peut même activer artificiellement (par des méthodes de stimulation optogénétique) les neurones dopaminergiques pour renforcer certaines actions chez l'animal, ou diminuer la probabilité que l'animal effectue d'autres actions (Kravitz, Tye, & Kreitzer, 2012). Tous ces résultats suggèrent que l'évolution a progressivement abouti vers des mécanismes neurophysiologiques reposant sur des calculs d'erreurs de prédiction pour permettre d'apprendre à préférer les actions menant à la récompense, et à éviter les actions menant à des dangers. Les modèles d'apprentissage par renforcement ont contribué à mieux comprendre la façon dont les mécanismes d'apprentissage par renforcement sont structurés dans un réseau autour des ganglions de la base, un ensemble de noyaux sous-corticaux dont le striatum constitue la principale structure d'entrée (Khamassi & Humphries, 2012; Collins & Frank, 2014).

Ces résultats ont aussi des conséquences importantes pour comprendre de multiples troubles neurologiques et psychiatriques liés à la fonction dopaminergique du cerveau. Par exemple, la maladie de Parkinson est une maladie où les neurones dopaminergiques dégénèrent progressivement. Bien que cette maladie ait longtemps été considérée comme une maladie du système moteur, elle inclut également des symptômes cognitifs, comme des difficultés d'apprentissage. L'étude du rôle de la dopamine, et des effets de certains médicaments à base de dopamine (comme la Lévodopa), a permis d'expliquer en partie ces effets comme étant liés à une altération des mécanismes d'apprentissage par renforcement (Frank, Seeberger, & O'Reilly, 2004). Les signaux dopaminergiques ont aussi un rôle essentiel pour les questions d'addiction, et des troubles de l'humeur. Les modèles d'apprentissage par renforcement permettent de relier les neurosciences à la psychologie, et de tenter d'expliquer pourquoi ces symptômes surviennent. Ceci a donné naissance à un nouveau domaine interdisciplinaire de recherche appelé la *psychiatrie computationnelle*, qui utilise les modèles (notamment d'apprentissage par renforcement) pour mieux caractériser les altérations du comportement dans certaines pathologiques, et mieux prédire les effets des médicaments (Huys, Maia, & Frank, 2016).

### 3.4 Autres

Malgré l'énorme succès de ces méthodes d'apprentissage par renforcement, elles ont aussi d'importantes limites. Par exemple, ces méthodes supposent que nous fondons nos décisions uniquement sur la valeur de nos choix, que nous avons apprise de manière incrémentale par expérience répétée. Cependant, il est évident que cette hypothèse est limitée, et ceci de plusieurs manières :

- Les humains peuvent parfois apprendre très vite, en une seule expérience : la plupart d'entre nous n'appuierons plus jamais sur un bouton qui nous délivre un fort choc électrique. Nous n'avons donc pas toujours besoin de multiples répétitions.

- Les humains peuvent même parfois apprendre sans faire l'expérience d'un phénomène spécifique : plus précisément, nous pouvons généraliser ce que nous avons appris à des situations totalement nouvelles. Par exemple, lorsqu'on acquiert un nouveau téléphone, nous pouvons transférer tout ce que nous savons des téléphones en général pour très vite s'adapter a la nouvelle situation, sans avoir besoin de nouvelle information.

- Les humains utilisent d'autres informations que la valeur passée pour apprendre à prendre des décisions. Par exemple, même si on a l'habitude de prendre tous les jours la même route pour aller au travail, si on est avertis qu'il y a des travaux sur cette route, on tiendra compte de cette information pour prendre une autre route ce jour-là.

Une manière de résumer les limites actuelles de l'algorithme TD RL est de dire qu'il n'intègre que peu d'informations disponibles. Un agent a en général davantage d'informations : il peut construire un modèle mental de l'environnement (quels états sont similaires les uns aux autres ; quels aspects d'une situation sont pertinents pour une décision ; quel est mon objectif actuel, etc.). La famille des modèles d'apprentissage par renforcement peut être largement étendue pour intégrer ce type d'information et modéliser de manière plus précise le comportement humain. Ces algorithmes, cependant, sont plus complexes, et nous ne comprenons pas encore bien s'ils permettent de rendre compte correctement du fonctionnement du cerveau.

# 4 Méthode

La modélisation mathématique est un outil très puissant en psychologie. Elle permet de définir de manière rigoureuse et quantitative des théories, qui peuvent ainsi générer des prédictions précises et falsifiables. Cependant, comme tout outil puissant, il est aussi dangereux (au sens où il peut conduire à de mauvaises conclusions et interprétations s'il n'est pas manipulé avec précaution), et il est donc essentiel d'être méthodologiquement solide pour l'utiliser. Nous présentons ici les points les plus importants de la méthode, et invitons le lecteur à lire des publications plus spécialisées pour approfondir (e.g., (Wilson & Collins, 2019)).

## 4.1 Définir un modèle mathématique

Comment définit-on un modèle mathématique pertinent pour la psychologie ? Nous avons donné deux exemples, mais en général, comment décide-t-on quelles sont les variables et la formulation mathématique adéquates pour représenter un phénomène cognitif particulier ? Il n'y a pas de réponse parfaite ni unique à cette question. Mais en général, les modélisateurs utilisent une des approches suivantes :

- **Inspiration de l'intelligence artificielle**. Plusieurs domaines de recherche - les mathématiques appliquées, l'informatique, les statistiques, l'apprentissage automatique (machine learning en anglais) et l'intelligence artificielle - produisent des algorithmes qui tentent de résoudre des problèmes similaires aux phénomènes psychologiques qu'un chercheur pourrait vouloir modéliser. Ces algorithmes ont souvent des propriétés importantes - par exemple des garanties théoriques indiquant dans quelles conditions l'algorithme est garanti de fonctionner, et dans lesquelles il peut ne pas marcher. Cette littérature scientifique est une source importante d'inspiration pour les modélisateurs. Elle apporte souvent une solution optimale à un problème spécifique, et permet de comparer le comportement humain à cette jauge d'optimalité. Une méthode fréquente consiste à prendre un modèle optimal, et à modifier ses composants pour le rendre sous-optimal d'une manière qui ressemble à la performance humaine. En effet, si une machine peut être programmée de façon optimale pour un problème dédié, l'humain a au contraire une capacité de s'adapter à de nombreuses situations, quitte à être très bon sans être parfait sur chacun des problèmes rencontrés.
- **Inspiration heuristique**. Beaucoup de modèles n'ont néanmoins pas de sources dans la théorie mathématique. Dans ce cas, les modèles sont plutôt inspirés de manière heuristique de la théorie psychologique : l'observation d'un phénomène amène le modélisateur à proposer une formulation mathématique qui reproduirait ce phénomène, sans qu'il puisse se référer à une théorie préalablement existante. Le modèle ainsi produit peut par la suite être testé psychologiquement, et potentiellement étudié pour ses propriétés mathématiques. Ceci peut parfois, à terme, conduire les modélisateurs à développer une nouvelle théorie mathématique du phénomène étudié.
- **Inspiration des neurosciences**. Certains modèles dérivent du champ des neurosciences computationnelles. L'observation du transfert d'information entre différents neurones, dans un réseau de neurones, ou entre différentes régions du cerveau peut offrir des indices sur comment certains phénomènes psychologiques émergent, et ainsi inspirer des modèles mathématiques correspondants.

Le processus de création d'un modèle n'est pas linéaire. Les modèles mathématiques sont en permanence testés : des expériences sont faites pour tester leurs prédictions. Lorsque de nouvelles données montrent les limites d'un modèle, le modèle peut être disqualifié, mais peut aussi être modifié - transformé en un nouveau modèle avec une nouvelle composante qui permet d'expliquer les nouvelles données. Ce modèle amélioré doit alors lui aussi être testé.

## 4.2 "Model-fitting" : relier le modèle aux données

Comment teste-t-on un modèle ? L'étape essentielle consiste à comparer le modèle aux données expérimentales. Le but est de vérifier que les prédictions faites par le modèle sont en effet réalisées. Pour ce faire, il y a (au moins) deux étapes importantes.

D'abord, le modèle doit être simulé et son comportement analysé. En pratique, on utilise un ordinateur pour créer un agent artificiel qui est mis dans la même exacte situation qu'un agent réel - par exemple, la même expérience de psychologie. Cette simulation produit un comportement, et ces données peuvent être analysées comme des données réelles. Les modélisateurs explorent le comportement d'un modèle en faisant un grand nombre de simulations, en variant les paramètres du modèles, et en analysant le comportement du modèle pour différentes valeurs de ces paramètres. Cela produit des prédictions qualitatives qui peuvent être comparées aux données expérimentales.

Ensuite, le modèle est relié aux données de manière plus quantitative. En particulier, la plupart des modèles ont des paramètres libres (par exemple le taux d'apprentissage $\alpha$ ou la température inverse $\beta$). Une étape importante consiste

à *"fitter"* le modèle, ce qui veut dire identifier la valeur des paramètres qui permet au comportement du modèle de s'approcher au plus près des données. Plusieurs méthodes existent pour fitter un modèle. La méthode la plus simple consiste à identifier *une mesure d'erreur* entre le comportement du modèle et les données, et de trouver les paramètres qui minimisent cette erreur. Plusieurs erreurs peuvent être considérées comme pertinentes pour ce processus - par exemple la distance entre les données et la prédiction (utilisée pour les régressions linéaires, cf. **Figure 1**), ou des mesures plus complexes aux propriétés mathématiques différentes (par exemple la distance de Kullbach-Leibler). Une méthode similaire consiste à calculer la probabilité que le modèle puisse générer les données observées avec une valeur des paramètres, et de sélectionner les paramètres qui optimisent cette probabilité. Cette méthode, qui repose sur la théorie Bayésienne, est la plus répandue pour les modèles cognitifs.

### 4.3 Comparer des modèles

Comment interpréter les résultats ? Mon modèle est-il un bon modèle ? Que faire avec ce modèle ? Deux étapes importantes, après le "model fitting", consistent à faire une *comparaison de modèles*, puis une *validation du modèle gagnant*.

Il est essentiel de ne pas considérer seulement un seul modèle, mais d'en considérer plusieurs. En particulier, il faut s'assurer que les prédictions qualitatives faites par un modèle sont distinctes de celles faites par un autre modèle. Dans l'étape de validation d'un modèle, un expérimentateur simule plusieurs modèles qu'il considère avec les paramètres qui ont été fittés aux données, et comparent qualitativement les prédictions faites par le modèle. Un modèle ne peut être considéré comme un bon modèle des données que s'il rend compte de manière adéquate des effets observés dans les données, en contraste avec d'autres modèles. En effet, un modèle seul peut donner l'impression de bien rendre compte des données (par exemple quand sa prédiction est une droite qui passe à peu près au milieu du nuage de points), alors qu'il existe un autre modèle considéré qui en rend compte de manière plus précise.

Une comparaison quantitative des modèles est aussi importante - elle ne remplace pas l'étape de validation qui est la seule à indiquer si un modèle est bon en soi - mais apporte des informations supplémentaires. En particulier, elle permet de considérer le problème de la complexité du modèle. En effet, un risque majeur quand on construit un modèle est le risque d'*overfitting* : créer un modèle trop compliqué - par exemple avec beaucoup d'éléments et de paramètres - qui reproduit tous les détails des données, même si ces détails sont en fait du bruit statistique (par exemple, la machine utilisée par l'expérimentateur pour mesurer ces données est imparfaites et introduit une petite marge d'erreur au moment de l'acquisition des données). Dans ces situations, en général, le modèle a toutes les chances de ne pas généraliser à d'autres données similaires. Par exemple, si un laboratoire disposant d'une machine plus précise fait une nouvelle expérience similaire, alors les nouvelles données mesurées pourront s'avérer beaucoup moins variables que ce que le modèle prédit. Il est donc important de pénaliser la complexité d'un modèle par rapport à la qualité de son fit quant on compare des modèles - c'est le principe du rasoir d'Ockham (mentionné dans l'introduction). Des mesures qui donnent une bonne approximation de cette pénalisation d'une manière mathématiquement fondée existent et permettent de limiter ce risque.

### 4.4 Interpréter les résultats

Lorsque ces étapes ont été effectuées, plusieurs types de conclusions peuvent être tirées de la modélisation.

D'abord, l'étape de comparaison des modèles est souvent importante. C'est une étape qui compare de manière formalisée et quantitative les prédictions faites par deux théories exprimées sous les formes d'un modèle quantitatif. À ce titre, cette étape permet parfois d'arbitrer entre différentes théories qui cherchent à expliquer un phénomène. Un exemple particulier consiste à identifier si le modèle optimal est un bon modèle des données, et si non, quelle partie du modèle est sous-optimale. En effet, si le modèle est optimal, on peut alors interpréter que l'évolution a pu aboutir à des mécanismes qui produisent le comportement le plus adapté à une situation donnée, car c'est ce qu'il y a de mieux pour bien faire face à cette situation. Si le modèle n'est pas optimal, c'est peut-être parce qu'on ne l'évalue que dans une situation donnée. Or, si on considérait un ensemble de situations différentes, on se rendrait peut-être compte que ce modèle correspond à ce qui est le plus optimal pour obtenir la meilleure performance moyenne sur l'ensemble de ces situations. Mais cela nécessite alors qu'une nouvelle théorie mathématique soit développée de ce qui est optimal sur l'ensemble de ces situations.

Ensuite, les modélisateurs considèrent souvent avec attention les paramètres obtenus pour le modèle *gagnant*. Ces paramètres sont différents pour différents sujets d'une même expérience, par exemple, et pourraient donc indiquer d'où viennent des différences comportementales observées (par exemple, est-ce que la performance est plus faible à cause d'un taux d'apprentissage plus lent, ou à cause d'un choix plus stochastique ?). Cette approche est souvent utilisée dans l'étude des différences inter-individuelles, par exemple dans le contexte du développement, ou de maladies psychiatriques ou neurologiques.

Enfin, un dernier usage fréquent des modèles consiste à considérer les variables mêmes du modèle - par exemple l'erreur de prédiction, ou la valeur d'un choix - et d'utiliser ces variables pour analyser d'autres données. La méthode de *model-based fMRI*, par exemple, utilise des variables dérivées d'un modèle, et les compare aux signaux observés dans différentes régions du cerveau, pour tenter de mieux comprendre comment celui-ci met en oeuvre un processus cognitif par des mécanismes similaires au modèle. En effet, si l'activité d'une partie du cerveau comme le striatum ventral montre systématiquement, au travers de nombreuses expériences répliquées par différents laboratoires, une corrélation avec une erreur de prédiction de la récompense telle que calculée par les modèles d'apprentissage par renforcement, c'est peut-être qu'avec ces modèles on touche du doigt un mécanisme fondamental de calcul d'erreurs de prédiction que le cerveau utilise pour permettre à l'individu de s'adapter à son environnement.

## 5 Conclusion

En résumé, la modélisation computationnelle est un type d'approches qui est de plus en plus utilisée en psychologie et en neurosciences cognitives. Elle permet de décrire de façon précise des hypothèses dans un langage mathématique, ce qui permet ensuite de simuler et prédire de nouvelles observations d'un phénomène. Grâce à cela, de nouvelles idées de protocoles expérimentaux peuvent être développées pour tester si les prédictions issus des modèles sont valides ou pas. Quand deux modèles font des prédictions différentes, un nouveau protocole expérimental peut même avoir pour objectif de trancher entre les deux modèles, et ainsi nous permettre de clarifier quels sont les mécanismes les plus plausibles pour rendre compte de telle ou telle observation du comportement humain ou de l'activité cérébrale sous-jacente.

Il y a néanmoins des limites et des précautions à garder en tête. Car avec beaucoup de paramètres, on peut faire dire presque tout ce que l'on veut à un modèle. Il convient donc d'adopter des méthodes rigoureuses de fit des données expérimentales avec les modèles ou de comparaison de modèles, dont les aspects principaux ont été résumés ici.

Ce chapitre a présenté les éléments essentiels à retenir sur la modélisation, illustrés par quelques exemples. Quelques suggestions de lectures sont données ensuite, pour celles et ceux qui souhaitent aller plus loin. Nous espérons que cela donnera envie à certains étudiants de psychologie et de neurosciences de se former à la modélisation pour pouvoir eux-mêmes programmer, simuler, et tester des modèles. Pour celles et ceux qui ne souhaitent par faire de la modélisation eux-mêmes, nous pensons que ce chapitre résume les connaissances principales qui vous permettront d'interagir avec des modélisateurs, de comprendre leur démarche, et parfois même de collaborer avec eux en leur proposant de se confronter à vos propres données expérimentales.

## References


Collins, A. G., & Frank, M. J. (2014). Opponent actor learning (opal): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological review*, *121*(3), 337.

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.

Huys, Q. J., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature neuroscience*, *19*(3), 404.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

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*, *6*, 79.

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), 1849–1866.

Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature neuroscience*, *15*(6), 816.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information, henry holt and co* (Vol. 2) (No. 4.2). Inc., New York, NY.

Schultz, W., Dayan, P., & Montague, P. R. (1997, mar). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599.

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature neuroscience*, *16*(7), 966.

Steverson, K., Chung, H.-K., Zimmermann, J., Louie, K., & Glimcher, P. (2019). Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. *Scientific reports*, *9*(1), 1–14.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *Elife*, *8*, e49547.