



**HAL**  
open science

## Mon cerveau et ses biais dans la prise de décision

Frédéric Alexandre

► **To cite this version:**

| Frédéric Alexandre. Mon cerveau et ses biais dans la prise de décision. 2021. hal-03476195

**HAL Id: hal-03476195**

**<https://inria.hal.science/hal-03476195v1>**

Submitted on 12 Dec 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.



commune et savoir les comparer. Deuxièmement, il faut avoir entendu assez longtemps les deux parties pour se faire un avis non biaisé, mais aussi, à la fin, il faut décider ; on ne peut pas rester tout le temps dans l'indécision. Il y a à trouver ici un compromis entre la vitesse et la justesse de la décision.

Pour ces deux types de difficultés, les mathématiques développent des outils intéressants. Certains sont proposés pour coder et comparer l'information de la façon la plus objective (la moins biaisée) possible. D'autres permettent de définir, pour un niveau de précision souhaité, le seuil optimal de différence entre les avis « Pour » et « Contre » qu'il faut atteindre avant de décider. Ces modèles mathématiques ont été mis en œuvre pour des tâches de décision perceptive élémentaires : vous voyez un nuage de points en mouvement et vous devez décider si, globalement, ces points vont plutôt à droite ou à gauche. Il est possible de rendre cette tâche très difficile en programmant le mouvement de chacun des points avec des fonctions en grande partie aléatoires, ce qui rend une décision locale impossible. Or on peut montrer que ces modèles mathématiques parviennent à reproduire très fidèlement la décision humaine, aussi bien dans ses performances que dans ses caractéristiques temporelles.



Décider est souvent peser le pour et le contre, de manière non biaisée, c'est ce que l'on représente au niveau de l'allégorie de la justice avec le bandeau et la balance à fléau.

©Joseolgon - Own work, wikicommon.

[https://fr.wikipedia.org/wiki/Justice\\_\(all%C3%A9gorie\)#/media/Fichier:Estatueta\\_Justi%C3%A7a.JPG](https://fr.wikipedia.org/wiki/Justice_(all%C3%A9gorie)#/media/Fichier:Estatueta_Justi%C3%A7a.JPG)

Ces modèles mathématiques sont également intéressants car, en les analysant, on peut observer les grandeurs et les phénomènes critiques au cours de la décision. Bien sûr, on trouve dans cette liste la détection de chaque indice, mais aussi leur accumulation, la différence entre les solutions alternatives, leur comparaison au seuil de décision, etc. Des études d'imagerie cérébrale permettent d'identifier les différentes régions du cerveau impliquées dans l'évaluation de chacun de ces critères et, au cours de la décision, l'ordre

dans lequel ces régions sont activées. Notre équipe de recherche travaille dans la réalisation de réseaux de neurones artificiels qui, d'une part, calculent de façon similaire à ces modèles mathématiques et d'autre part, sont organisés selon une architecture globale reproduisant la circuiterie observée par imagerie et reproduisant également la dynamique d'évaluation. C'est à ce stade que nous pouvons constater que les modèles mathématiques évoqués plus haut ont un certain nombre de limitations et que nous pouvons modifier et adapter nos réseaux de neurones pour considérer des cas plus réalistes.

### **Quand les mathématiques ne suffisent plus.**

Au-delà, il nous faut considérer plus que des sciences formelles : des sciences humaines. Nous nous intéressons en particulier à trois types de limitation des modèles mathématiques. Premièrement, nous pouvons aller au-delà de ces modèles qui considèrent uniquement des décisions binaires (droite/gauche) en introduisant, dans les calculs neuronaux, des étapes de codage supplémentaires et des non-linéarités permettant de pouvoir considérer plusieurs catégories. Deuxièmement, les modèles mathématiques sont le plus souvent appliqués à des indices perceptifs alors que des arguments de nature différente peuvent être présentés pour emporter une décision (l'évocation de souvenirs ou de valeurs émotionnelles par exemple). Nous cherchons à bénéficier du fait que nos réseaux de neurones sont inscrits dans la circuiterie de l'architecture cérébrale et à étudier comment ajouter d'autres indices (mnésiques ou émotionnels) provenant d'autres aires cérébrales. Troisièmement, alors que nous avons parlé jusqu'à présent de décisions mathématiquement fondées (certains disent logiques ou rationnelles), il est connu que les humains sont soumis à différents types de biais quand ils font des jugements, ce qui laisse souvent penser que nous ne sommes pas rationnels.

En collaboration avec les sciences humaines et sociales qui étudient et décrivent ces biais, nous cherchons à montrer que nous pouvons les expliquer et les reproduire en manipulant certains paramètres de nos modèles. Nous prétendons aussi que ces biais ne montrent pas une faiblesse de notre jugement mais plutôt une orientation de ce jugement tout à fait pertinente pour un être vivant évoluant dans des conditions écologiques. Donnons quelques exemples. Au lieu d'intégrer tous les arguments de la même manière dans notre jugement, nous pouvons être soumis à un biais de récence ou de primauté, selon que les indices les plus récents ou les plus anciens vont jouer un rôle plus important dans la décision. Ce type de jugement, qu'on peut reproduire en modifiant certains paramètres internes du calcul neuronal, peut paraître plus adapté pour certaines situations avec des conditions changeantes (récence) ou stables (primauté). On sait par ailleurs que détecter ce type de conditions nous fait émettre des neurohormones qui modifient la nature du calcul neuronal en modifiant certains de leurs paramètres. Nous essayons de reproduire ces mécanismes dans nos modèles.

Un autre biais, appelé aversion au risque, fait que nous surestimons les indices défavorables par rapport aux indices favorables. Il a été montré expérimentalement en économie que les neurones qui codent les pertes et ceux qui codent les gains ne sont pas soumis aux mêmes non-linéarités dans leurs calculs, ce que l'on sait simplement reproduire dans nos modèles. Ici aussi, on peut comprendre que, pour un être vivant qui ne fait pas que calculer des bilans financiers, il est judicieux d'accorder plus d'importance à ce qui peut nous nuire (et

potentiellement nous blesser) qu'à ce qui est positif (et qu'on aura d'autres occasions de retrouver). De façon similaire, d'autres types de biais, appelés de référence ou de préférence, font que nous allons accorder plus d'importance à des indices faisant référence à un événement récent ou à une préférence personnelle. Ici aussi, introduire des mécanismes cérébraux connus de types mnésiques permet de reproduire ces phénomènes, dont l'intérêt adaptatif semble aussi clair. Enfin un dernier type de biais concerne la différence entre les valeurs que nous donnons à des situations et les valeurs que nous donnons aux actions pour les atteindre ou les éviter et que l'on connaît bien en addictologie : alors que l'on sait très bien que alcool, tabac et autres drogues sont mauvais pour nous, diminuer nos actions de consommation n'est pas forcément simple. On sait que dans le cerveau ces valeurs sont dissociées, autrement dit que nos actions ne reflètent pas toujours nos pensées.



L'équipe projet Mnemosyne de l'Inria, du Labri et de l'IMN à Bordeaux travaille sur la modélisation des fonctions cognitives du cerveau, de manière pluri-disciplinaire. (Crédits : Inria / Photo H.Raguét)

### **Une application concrète : aider à avoir un comportement plus éco-responsable.**

Avec l'aide de la Région Nouvelle Aquitaine, de l'Inria et des universités de Bordeaux et de La Rochelle, nous sommes actuellement impliqués dans un projet visant à mieux comprendre les décisions humaines relatives au changement climatique et à la transition écologique. Il s'agit d'un sujet de décision visant à changer nos comportements et on peut constater que la plupart des biais mentionnés plus haut s'appliquent : est-on prêts à changer d'habitude maintenant (prendre le bus plutôt que la voiture) pour des résultats à long terme (modifier la température moyenne dans 30 ans), quitter des comportements faciles et observés autour de soi (prendre sa voiture comme tout le monde) et mettre en œuvre nos convictions citoyennes (différence entre ce qu'on pense et ce qu'on fait), etc.

Ce projet va nous permettre à la fois de recueillir des informations pour affiner nos modèles et de tester leurs prédictions en interagissant avec les utilisateurs. Pour cela, nous utilisons une application que des personnes volontaires auront installée sur leurs smartphones. Cette application les aide à choisir leur mode de transport pour aller travailler et les informe sur les caractéristiques éco-responsables de leurs choix. Nous aurons ainsi accès aux décisions

qui auront été prises, en fonction des situations mais aussi des informations qui avaient été données. Nous avons l'espoir que cette approche nous permettra d'évaluer la pertinence et le poids de différents types d'arguments que nous aurons préparés en amont avec nos collègues des sciences humaines et sociales, reposant en particulier sur des hypothèses issues de nos modèles, sur des mécanismes supposés de la prise de décision.

[Dans ce projet](#), nous avons orienté notre approche selon les convictions suivantes : D'une part, si nous savons bien décrire les caractéristiques de cette prise de décision et en particulier pourquoi et quand elle est difficile, nous pouvons proposer des mises en situations qui seront plus favorables à des prises de décision responsables. D'autre part, nous faisons également le pari du citoyen informé et proactif. Si nous décrivons ces mécanismes de décision, leurs biais et leur inscription cérébrale, cela peut permettre de déculpabiliser (ce sont des mécanismes biologiques à la base de ces biais) et de donner des leviers pour travailler sur nos processus mentaux et réviser nos modes de pensée.

**Frédéric Alexandre, Directeur de Recherche Inria.**