

L'Évolution de la raison  
Sommes-nous devenus rationnels?

Stephen N. Lyle

December 21, 2022

Nous ne parlerons pas ici d'une quelconque évolution récente de la rationalité humaine. Nous adoptons un point de vue biologique de l'évolution et de la rationalité, et donc de nous-mêmes, l'espèce humaine. Qu'est-ce qui motive cela? On dit souvent que ce qui nous distingue le plus des autres animaux est notre capacité à raisonner, notre faculté de raison. Mais cela soulève une question: pourquoi notre espèce a-t-elle développé ce trait? Quel avantage nous a-t-elle donné pour que les gènes qui nous prédisposent à raisonner aient proliférer en nous?

Pour aller plus loin, nous devons décider ce que nous entendons par raisonnement. Dans son récent livre *Rationalité*, le psychologue cognitif canadien Steven Pinker nous donne cette idée: le raisonnement consiste à utiliser des connaissances pour atteindre un objectif [1]. Mais ce genre de raisonnement ne nous distinguerait guère des autres animaux. Tous les animaux collectent des informations de leur environnement avec leurs différents sens, les traitent dans leur corps et en particulier dans leur cerveau s'ils en ont un, et agissent en conséquence. En ce sens, on pourrait dire que tout traitement de l'information est une sorte de raisonnement, idée qui sera développée ici. Et notez que les animaux, y compris nous-mêmes, ne sont pas forcément conscients de ce type de traitement de l'information, donc cela pourrait bien être une sorte de raisonnement inconscient, une autre idée qui sera développée ci-dessous.

Si nous voulons une notion de raisonnement qui nous distinguerait vraiment des autres animaux, on pourrait considérer quelque chose comme: utiliser la logique pour résoudre des problèmes et s'expliquer clairement. Cela semble impliquer le langage, ce qui exclut donc la plupart des autres animaux, bien qu'il faille rester ouvert. Considérons donc brièvement ce qu'est la logique. Voici un court raisonnement sous forme générique:

Si A, alors B  
Pas B  
Donc pas A

Il y a deux prémisses et une conclusion. A et B sont des propositions, c'est-à-dire des choses que l'on pourrait dire sur ce qui se passe dans son environnement. Ceci peut être illustré par un exemple. S'il pleut pendant une semaine entière, le ruisseau déborde de ses berges. Mais le ruisseau n'a pas débordé. Par conséquent, il n'a pas plu pendant une semaine entière. Ce genre de petit raisonnement s'appelle un syllogisme et il a même un nom: *modus tollens*. Ceci est en latin, ce qui montre que les humains connaissent les syllogismes depuis longtemps!

Voici un autre syllogisme, appelé *modus ponens*:

Si A, alors B  
A  
Donc, B

Encore une fois, il y a deux prémisses et une conclusion. Nous pouvons illustrer avec le même exemple. S'il pleut pendant une semaine entière, le ruis-



Figure 1: Le grand koudou (*Tragelaphus strepsiceros*), une espèce de mammifères de la famille des bovidés qui vit dans la région de l’Afrique de l’Est à l’Afrique du Sud. Image: L0k1m0nk33 at English Wikipedia

seau déborde de ses berges. Il a en effet plu pendant toute une semaine. Par conséquent, le ruisseau sera en train de déborder de ses rives en ce moment.

Ces deux arguments courts sont clairement *valides*, quel que soit leur contenu, tel que spécifié par les propositions A et B. Ils sont valides par leur seule forme. La validité est un terme technique précis pour les logiciens: un argument est valide si, chaque fois que les prémisses sont vraies, la conclusion est nécessairement vraie. Et la logique est l’étude de la forme des arguments valides, et par extension, l’étude de la forme des arguments invalides.

Voilà pour la logique. Mais sommes-nous logiques? Et si nous le sommes, l’avons-nous toujours été? Qu’en est-il de nos lointains ancêtres *Homo sapiens*, il y a cent mille ou deux cent mille ans? Comment pourrions-nous le savoir? La figure 1 montre un koudou, un animal qui vit dans le désert du Kalahari en Afrique australe. Cet animal est chassé, voire traqué sur de grandes distances, par les San, un peuple de chasseurs-cueilleurs vivant dans la même région. Pour leur part, les San ont longtemps été traqués par les anthropologues occidentaux, précisément parce que nous pensons que ces personnes vivent comme nos lointains ancêtres, chassant et cueillant, il y a cent mille ou deux cent mille ans. Et bien sûr, une des choses qui nous intéresse est de savoir si elles sont logiques! Voici un exemple, donné par Pinker dans son livre [1], pour confirmer qu’ils sont très certainement logiques:

Le koudou peut être attrapé pendant la saison sèche car il se fatigue facilement lorsqu'il doit courir sur du sable meuble

C'est la saison sèche et l'animal qui a laissé ces traces est un koudou

En conséquence, nous devrions traquer cet animal

Il n'y a pas vraiment de surprise ici. Pinker donne de nombreux autres exemples de la rationalité chez les San. Il s'avère que leur rationalité est la même que la nôtre. Et leur irrationalité aussi. Ces choses vont clairement être des universaux, communs à toute notre espèce, et ce tout au long de son existence.

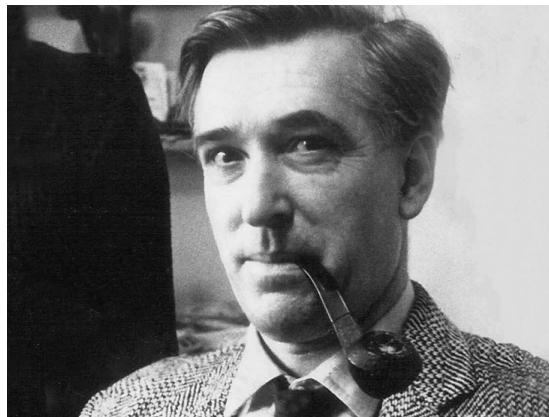


Figure 2: Peter Cathcart Wason (1924–2003) était un psychologue cognitif à l'University College, Londres, qui a été un pionnier de la psychologie du raisonnement. Il a avancé des explications sur les raisons pour lesquelles les gens commettent certaines erreurs récurrentes dans le raisonnement logique. Il a conçu des problèmes et des tests pour démontrer ces processus, comme la tâche de sélection qui porte son nom. Image: <https://medium.com/five-guys-facts/the-wason-selection-task-4521cd54cc7b>

Mais sommes-nous toujours logiques? La figure 2 montre le psychologue cognitif britannique Peter Wason, clairement inspiré par Sherlock Holmes, qui a introduit un test de logique très révélateur dans les années 1960. Le test de Wason semble très simple (voir la figure 3). On nous présente quatre cartes. Sur chacune, il y a une lettre d'un côté et un chiffre de l'autre, et on nous dit que les cartes doivent obéir à la règle suivante: s'il y a une voyelle d'un côté, il doit y avoir un nombre pair de l'autre. Alors quelles cartes doit-on retourner pour vérifier que la règle a bien été respectée? Beaucoup de gens raisonnent de la manière suivante. La première carte montre une voyelle, il faut donc vérifier qu'il y a bien un nombre pair de l'autre côté. Le second montre une consonne, et puisque la règle ne mentionne pas les consonnes, pas besoin de vérifier l'autre



Figure 3: Quatre des cartes de Wason. D'un côté il y a une lettre et de l'autre un chiffre. Et il y a une règle: si une carte a une voyelle d'un côté, elle a un nombre pair de l'autre. Quelles cartes doit-on retourner pour voir si la règle est respectée?

côté. Le troisième montre un nombre pair, et puisque la règle mentionne des nombres pairs, nous devons vérifier qu'il y a une voyelle de l'autre côté. La dernière carte affiche un nombre impair, non mentionné par la règle, donc pas besoin de vérifier.

Mais ce raisonnement est faux. Bien sûr, certaines personnes ne font pas l'erreur avec les deux dernières cartes. Les logiciens, par exemple. On n'a pas besoin de tourner la troisième carte, car quoi qu'il y ait de l'autre côté, consonne ou voyelle, la règle est respectée. Par contre, si la dernière carte avait une voyelle de l'autre côté, la règle n'aurait pas été respectée, et il faut donc bien vérifier celle-ci.

Cela montre que beaucoup de gens n'appliquent pas correctement ces simples raisonnements logiques valides. Mais regardez les quatre cartes de Wason sur la figure 4. Cette fois, chaque carte a l'âge de quelqu'un d'un côté et ce qu'il boit de l'autre. Et cette fois la règle est bien sûr qu'il faut avoir au moins dix-huit ans pour boire de l'alcool. Quelles cartes doivent être retournées maintenant? Cette fois c'est très facile, et personne, absolument personne ne fait d'erreur ici. La première carte doit être retournée pour vérifier que ce jeune de seize ans ne boit pas d'alcool. On n'a pas besoin de retourner la seconde, car à vingt-cinq ans, la personne peut boire ce qu'elle veut. La troisième personne boit du soda, donc peu importe son âge, mais la quatrième boit de la bière, nous devons donc vérifier.

Et alors? Le fait est que, logiquement parlant, le deuxième test est *exactement* le même que le premier. Alors comment se fait-il que tant de gens fassent des erreurs avec la première version, et personne avec la seconde? La réponse est que le premier critère est abstrait alors que le second porte sur des conditions déontiques, c'est-à-dire des conditions relatives au devoir et à l'obligation. En effet, nous sommes extrêmement doués pour tenir les autres responsables de leurs actes, et tout aussi doués pour trouver notre chemin dans un monde où les autres nous tiendront responsables des nôtres. Et ça, c'est parce que nous vivons dans une société de coopération. En effet, les humains sont les champions du monde des coopérateurs, bien plus que toute autre espèce. Nous pouvons coopérer sur



Figure 4: Quatre autres cartes de Wason. D'un côté, il y a l'âge de quelqu'un et de l'autre une image de ce qu'il boit. Encore une fois, il y a une règle: il faut avoir au moins dix-huit ans pour être autorisé à boire de l'alcool. Quelles cartes doit-on retourner pour voir si la règle est respectée?

tout et n'importe quoi. D'autres animaux comme les loups, les chimpanzés, les abeilles et les fourmis peuvent vivre en communautés et travailler ensemble de différentes manières, mais les humains inventent continuellement de nouvelles façons de faire les choses ensemble pour de meilleurs résultats.

Mais cela a un prix, car il est important de savoir à qui nous pouvons faire confiance, et il est important que les autres ont au moins tendance à être dignes de confiance. En d'autres termes, nous avons besoin d'un système d'éthique. La clé ici est la moralité. Alors voici quelque chose d'intéressant: nous avons entrepris de faire une théorie cognitive du raisonnement, et ici nous parlons de moralité. Ce n'est pas un hasard. Le consensus en psychologie cognitive considère aujourd'hui que le raisonnement a évolué chez l'homme à des fins de moralisation, pour nous permettre de vivre dans une société majoritairement coopérative. Les bénéfices pour les gènes conduisant à de tels comportements sont légion et ils ont proliféré. Ces gènes sont presque partout aujourd'hui dans la société humaine. On dit parfois qu'il y a eu une sorte d'auto-domestication depuis quelques dizaines de milliers d'années, c'est-à-dire une coévolution des gènes et des pratiques culturelles qui a fait que certains gènes, ces gènes de la coopération, sont devenus dominants. Nous reviendrons plus tard sur ce thème important.

Mais suivant Pinker dans son dernier livre, demandons-nous quel type de contenu A et B dans nos syllogismes nous transformerait temporairement en logiciens fiables. Selon lui, il est clair que ces contenus doivent incarner les types de défis logiques auxquels nous nous sommes habitués lorsque nous sommes devenus adultes et peut-être lorsque nous sommes devenus humains [1]. Nous avons vu que les chasseurs-cueilleurs San appliquent correctement le raisonnement syllogistique lorsqu'ils chassent ensemble. Ce serait désastreux s'ils ne le faisaient pas, étant donné l'importance de la tâche. Mais lorsqu'un problème n'est pas vraiment important pour nous, comme la première tâche de sélection de Wason, totalement abstraite, illustrée à la Fig. 3, nous avons tendance à opter pour la réponse facile, même si nous savons très bien qu'elle peut être incorrecte parce que nous avons à peine investi l'effort mental nécessaire pour être sûr. Et voici

un autre point important: travailler sur des énigmes logiques est un exercice coûteux en termes d'énergie et de temps. Essayez ceci, par exemple, également de Pinker [1]:

Huit imprimantes mettent huit minutes pour imprimer huit copies d'un document.

Combien de temps mettront vingt-quatre imprimantes pour imprimer vingt-quatre copies du même document?

Non, ce n'est pas vingt-quatre minutes! Mais j'espère que dans une situation réelle avec les imprimantes devant vous, vous obtiendrez immédiatement la bonne réponse.

## L'esprit

Au début, j'ai dit que j'adopterais un point de vue biologique de la raison, et donc de nous-mêmes. Mais ce n'est pas ainsi que nous nous comprenons habituellement lorsque nous faisons de l'introspection. Bien sûr, je peux *voir* mon corps plutôt animal, mais quand je fais de l'introspection, je trouve . . . moi, mon esprit. Le soi. Et bien sûr, c'est l'esprit ou le soi qui pense et raisonne, n'est-ce pas? C'est un dualisme tout à fait naturel: esprit et corps, ou corps et âme. Mais qu'est-ce que l'esprit? Comment ça marche? Et comment est-ce qu'une telle chose aurait pu évoluer?

Pour aborder ces questions, il est intéressant de se pencher sur les deux principaux paradigmes de l'intelligence artificielle. Car il ne faut pas oublier que l'un des objectifs de la recherche en intelligence artificielle est de comprendre l'intelligence naturelle—notre propre intelligence. A l'époque où se développaient les ordinateurs, depuis la Seconde Guerre mondiale, et particulièrement dans les années 1960 et après, il était naturel de se demander si nos propres cerveaux pouvaient fonctionner comme ces nouvelles machines informatiques. Après tout, les ordinateurs traitent l'information, et ils sont logiques, même exclusivement logiques. Cela a conduit à un paradigme d'IA qui est parfois appelé de manière plutôt péjorative la bonne vieille intelligence artificielle ou GOFAI (good old-fashioned artificial intelligence).

Ce paradigme maintient le dualisme que nous trouvons dans notre perspective ordinaire de nous-mêmes. Il y a du matériel, le hardware, et des logiciels, le software. Un robot GOFAI planifie et calcule d'abord, puis agit. Mais cela nécessite une énorme quantité de mémoire et une puissance de calcul considérable. L'une des raisons est qu'il se heurte à ce que l'on appelle le *problème du cadre* (le frame problem, en anglais). Il s'agit essentiellement du problème de la représentation de l'environnement du robot et des changements qui doivent être apportés à cette représentation chaque fois que le robot fait quelque chose. Déterminer ce qu'une action change et ce qu'elle ne change pas n'est pas simple, en particulier pour les systèmes qui ne peuvent agir qu'en appliquant des

principes préinstallés. Pour un système GOFAI, toutes les conséquences possibles d'une action doivent être programmées dans le logiciel à l'aide d'algorithmes que l'ordinateur peut exécuter. Vous devez faire une longue liste de tous les changements nécessaires dans votre représentation du monde pour chaque action possible, et une liste encore plus longue de tous les changements que vous n'avez pas besoin de faire.

Voici un premier enseignement de cette recherche en intelligence artificielle: il y a des choses qui nous sont très faciles—distinguer visuellement une pomme et une poire, marcher, attraper un objet en vol, etc.—mais ça ne veut pas dire qu'elles sont simples. Elles s'avèrent parfois très, très difficiles à programmer dans un robot. Nous reviendrons sur ce point avec une solution en partie plus tard. Mais pour l'instant, la question que cela pose est celle-ci: est-ce que les êtres humains fonctionnent vraiment comme ça? Et sinon, comment éviter d'avoir à tout calculer avant d'agir?

Voici une autre chose qui est pertinente à cette question. Il semble que les très jeunes enfants, même dès l'âge de trois ou quatre mois, ont des attentes concernant le comportement de la matière inerte. En fait, ils semblent appliquer certains principes simples de physique, parfois appelés la physique naïve. Nous le savons même si ces enfants ne sont pas encore en mesure de nous parler, car ils sont trop jeunes. L'astuce consiste à regarder de très près leurs yeux lorsqu'on leur présente différents scénarios simulés sur un écran, certains étant physiquement possibles et d'autres physiquement impossibles, pour voir s'ils sont surpris dans ce dernier cas (voir [2] pour un exemple). Ce processus, connu sous le nom de suivi oculaire (*eye tracking* or *gaze tracking* en anglais), mesure soit le point du regard (là où l'enfant regarde), soit le mouvement d'un œil par rapport à la tête. Les *eye trackers* sont largement utilisés dans la recherche sur le système visuel, mais aussi en psychologie et en psycholinguistique, et dans de nombreuses autres applications. Sans entrer dans le détail des principes de la physique naïve qui ont été établis, voici quelques exemples:

- Principe de solidité. Un objet solide ne peut pas traverser un autre.
- Permanence de l'objet. Un objet continue d'exister même lorsqu'il passe derrière un autre et n'est plus visible.
- Principe de cohésion. Un objet en mouvement conserve sa connectivité et ses limites.

Une revue générale peut être trouvée dans [3].

Ce qui nous intéresse ici, c'est l'existence de principes physiques dans l'esprit du bébé, principes qui sont très probablement innés. Ces nourrissons tirent clairement des conclusions sur les choses qu'ils regardent. Or, les principes de base de la physique naïve pourraient être représentés explicitement et appliqués de la même manière qu'ils sont appliqués dans les ordinateurs, et comme on pourrait essayer de le faire si l'on souhaitait programmer un robot. Il s'agit de l'approche computationnelle de la GOFAI, qui suppose qu'un système cognitif ne peut agir conformément à des principes que si ces principes ont été explicitement représentés au sein du système sous une forme symbolique appropriée.



Cela a été considéré comme le principe fondateur de la GOFAI, connu sous le nom de *l'hypothèse de symbole physique*: les règles et les principes sont explicitement représentés sous forme symbolique dans le cerveau. On pourrait dire quelque chose comme ceci: la syntaxe suit la sémantique. L'idée est fondamentalement que c'est la structure même de la séquence de symboles utilisée qui contient la signification de cette séquence. Dans la liste ci-dessus, les principes sont explicitement représentés en français. Cela peut être traduit en un langage informatique exprimé dans l'alphabet binaire bien connu des zéros et des uns, et cela à son tour peut être traduit en une configuration, disons d'atomes, à l'intérieur de l'ordinateur, où cette configuration contient la signification du principe dans la structure même de son agencement.

Cependant, il s'avère que l'on peut développer ce que l'on appelle des *réseaux de neurones artificiels* capables de simuler le comportement de bébés humains dans ces expériences sans encoder explicitement aucune règle ou principe. Nous avons donc ici une preuve de concept que l'approche informatique *n'est pas* la seule option. Quelque chose qui ressemble à un comportement basé sur des règles peut émerger d'un système qui n'exploite aucune règle explicitement représentée. Il peut donc exister une forme de traitement de l'information qui n'implique pas le type de manipulation de symboles régi par des règles proposé par l'hypothèse de symbole physique. En effet, aucun principe n'est stocké tel quel dans un réseau de neurones artificiels.

Le problème avec la GOFAI est qu'il s'agit d'une approche intellectuelle top-down de l'IA. Toute l'intelligence est dans la programmation. Et c'est trop lent. Nous avons juste de la chance que les ordinateurs calculent si vite, car ils doivent tout optimiser et planifier à l'avance. Imaginez les calculs nécessaires pour amener un robot GOFAI à se frayer un chemin à travers une forêt, en évitant les nombreux obstacles par l'application de ses principes préprogrammés.

Juste un mot sur ces réseaux neuronaux artificiels. L'approche ici est vraiment à l'opposé de ce qui se passe dans la GOFAI. Dans ce cas, nous avons commencé par regarder l'ordinateur et nous nous sommes demandé si notre cerveau pouvait fonctionner comme ça. Avec les réseaux neuronaux artificiels, nous commençons par le cerveau et réalisons des simulations hautement simplifiées de réseaux de neurones réels. C'est-à-dire que nous examinons certaines caractéristiques des neurones et les imitons dans le réseau artificiel. Ainsi, un tel réseau est un processeur d'informations avec certaines des caractéristiques du cerveau, mais beaucoup, beaucoup moins de neurones. Et il peut apprendre, dans un certain sens. D'où le nom anglais de *deep learning* pour ce paradigme d'IA. C'est-à-dire que le réseau neuronal artificiel est présenté avec de nombreuses données concernant la question à laquelle il est censé pouvoir répondre, dans des situations où la réponse à cette question est connue. Lorsqu'il donne une réponse inexacte, ses paramètres sont ajustés de manière bien définie avant de lui présenter plus de données, et ainsi de suite jusqu'à ce qu'il donne toujours, ou presque toujours, la bonne réponse. Une bonne introduction à l'intelligence artificielle, sans battage publicitaire, est l'excellent livre de la chercheuse en intelligence artificielle Melanie Mitchell [4].

## La Pensée

Il y a un instant, j'ai dit que c'était l'esprit ou le soi qui pense, ou bien qui raisonne. Mais je viens de parler de *neurones* artificiels, inspirés de vrais neurones dans notre cerveau. Un neurone est une cellule et il y en a environ 80 milliards dans un cerveau humain, ainsi que de nombreux autres types de cellules. On sait que l'information passe le long de ces cellules mais aussi entre elles, chaque neurone étant connecté à en moyenne 7000 autres neurones. Et maintenant, nous savons qu'en imitant quelques caractéristiques simples de ces cellules, nous pouvons construire des réseaux de neurones artificiels qui traitent l'information, voire semblent appliquer des principes qui ne sont nulle part encodés symboliquement en leur sein. Ceci suggère fortement que ce sont en quelque sorte les neurones et leur entourage d'autres cellules qui pensent et raisonnent. En effet, il y a tout lieu de penser que le soi n'est qu'une représentation générée par le cerveau pour représenter l'individu dans lequel il réside, très probablement aux fins des interactions de cet individu avec d'autres humains, en fait, avec d'autres soi.

Le dualisme naturel de l'esprit et du corps n'est alors qu'une impression que nous avons. Et si l'on espère se comprendre du point de vue biologique, il faut bien dépasser l'histoire racontée par l'introspection et adopter une perspective cognitive—ou neuronale—de l'individu, pour reprendre une tournure de phrase utilisée par le linguiste Ray Jackendoff dans son excellent livre *A User's Guide to Thought and Meaning* [5]. Fondamentalement, cela signifie partir de ce que nous pouvons observer—la matière, ses configurations et sa dynamique—et arriver d'une manière ou d'une autre à l'esprit sans ajouter aucun autre ingrédient. C'est le but du philosophe naturaliste.

Il existe une anecdote bien connue au sujet de J.M. Keynes, le célèbre économiste britannique. Bertrand Russell a déclaré que Keynes était la personne la plus intelligente qu'il ait jamais rencontrée, ce qui n'est pas une mauvaise référence. Quelqu'un a un jour demandé à Keynes s'il pensait en mots ou en images, une énigme philosophique séculaire. Keynes réfléchit un instant, puis dit: "Je pense en pensées." Au premier abord, cela ne semble pas être une remarque très utile. Cependant, s'il suggérait que nous ne pensons ni en mots ni en images, alors il était proche de la position de nombreux neuroscientifiques aujourd'hui. Car dans la perspective cognitive, il est généralement admis que *toutes nos pensées sont inconscientes*.

Cela peut sembler fou, mais considérez ceci. Lorsque vous devez communiquer une idée plutôt subtile à quelqu'un d'autre, ne trouvez-vous pas parfois que vous n'avez pas dit exactement ce que vous vouliez dire? Ce n'était pas exactement l'idée que vous vouliez exprimer? Ensuite, vous reformulez, peut-être plusieurs fois, en ajustant ce que vous dites jusqu'à ce qu'à un moment donné, cela sonne juste. Cela correspond enfin à votre idée. Pensez à ce qui se passe lorsque vous devez écrire une lettre plutôt délicate ou une dissertation. Combien de fois finissez-vous par rayer quelque chose ou l'ajuster?

Alors pourquoi disons-nous des choses qui ne signifient pas ce que nous voulions dire? Jackendoff et d'autres diront que c'est parce que cette signi-

fication est en fait cachée dans notre inconscient—c’est *l’hypothèse de la signification inconsciente* [5]. Ce que nous appelons habituellement la pensée n’est qu’une expression linguistique de notre pensée inconsciente réelle, une représentation symbolique dans notre conscience utilisant le symbolisme du langage pour représenter des pensées et des significations qui se trouvent réellement au plus profond des configurations et de la dynamique de nos neurones. La “pensée consciente” ou la “réflexion consciente” n’est alors qu’une séquence (logique?) d’interprétations linguistiques de notre pensée réelle, qui, pour sa part, s’effectue à notre insu—c’est-à-dire à l’insu de notre introspection—par nos neurones.

Cela nous ramène à la bonne vieille IA, la GOFAI, qui repose sur l’idée que nous pensons comme nous pensons que nous pensons, dans notre perspective ordinaire de nous-mêmes. Nous pensons que nous pensons en séquences logiques de phrases—ou propositions, pour utiliser un terme philosophique fortement connoté—c’est-à-dire des séquences de symboles comme ceux que nous introduisons dans nos ordinateurs, et que nos ordinateurs peuvent manipuler en utilisant des algorithmes. Mais d’un point de vue cognitif, abandonnant pour un instant les prétentions de notre moi introspectif, il est possible que la pensée ne soit pas en fait une question d’application de la logique aux prémisses et aux principes. Nous savons depuis le développement des réseaux neuronaux artificiels qu’il n’est absolument pas nécessaire que les prémisses et les principes soient codés quelque part dans notre cerveau. Dans la perspective ordinaire que nous donne l’introspection, nous avons toujours l’impression que l’information doit être encodée quelque part car notre seule prise sur ces informations—ou plutôt, la seule prise de notre introspection sur ces informations, et notamment sur leur sens, et par la “petite voix” que nous entendons jacasser presque en permanence dans nos têtes. Mais il se pourrait que ces pensées et ces significations ne soient encodées par le langage qu’à des fins de communication, avec les autres ou bien avec nous-mêmes. Et en réalité, comme nous l’avons vu, nous avons souvent du mal à raisonner de manière strictement logique, ce que nos ordinateurs font à merveille par conception.

## L’évolution et l’esprit modulaire

Nous venons de parler des neurones et du cerveau. Il semble assez clair que ce sont en fait ces entités purement biologiques qui pensent et raisonnent à notre place. Nous pouvons donc maintenant revenir à une question précédente et la reformuler: comment *notre cerveau* évite-t-il d’avoir à tout calculer avant de nous dire comment on doit agir? Pour aborder cela, il est utile de se demander *comment* nous, les humains, avons été conçus?<sup>1</sup> Mais nous connaissons la réponse: nous avons été lentement bricolés par l’évolution. Pour citer le biologiste écossais D’Arcy Wentworth Thompson: “Tout est tel qu’il est parce que

---

<sup>1</sup>Ce mot est utilisé dans le sens proposé par le philosophe Daniel Dennett pour désigner la conception par mutation génétique et sélection naturelle, où il n’y a pas de concepteur (voir, par exemple, [8]).

il est devenu comme ça.” Demandons-nous donc comment l’esprit humain, ou l’esprit de tout autre animal, aurait pu évoluer.

Une idée importante en psychologie évolutionniste, c’est-à-dire dans la manière évolutionniste de comprendre l’activité cognitive, est la théorie modulaire de l’esprit. En un mot, c’est l’idée que l’esprit n’est pas unitaire et fondé sur des principes, mais modulaire et opportuniste, un assemblage d’éléments que l’évolution a “mis ensemble” pour faire face aux problèmes spécifiques rencontrés par nos ancêtres—et pas seulement nos ancêtres humains, mais tous nos ancêtres, comme nous le verrons.<sup>2</sup> Dans cette optique, l’esprit animal est constitué de modules qui tirent des inférences, chacun dans son propre domaine spécifique, et cela sur toute une gamme de domaines très variés. Un module traite des informations, donc raisonne, au sens large évoquées au début de cette présentation, et l’animal agit en conséquence. Chaque espèce évolue, se construisant module après module, lui conférant des capacités différentes dans ses interactions avec le monde.

Cela sera illustré dans un instant, mais d’abord un peu d’histoire. En poussant un peu l’imagination, on pourrait dire que l’idée d’un esprit modulaire remonte au pratique largement discrédité de la phrénologie: l’étude de la taille et de la forme du crâne était censée nous renseigner sur le caractère et les capacités mentales d’une personne. L’idée a été réhabilitée assez différemment par le philosophe Jerry Fodor dans les années 1970 [6], laissant tomber l’idée que les modules doivent nécessairement être localisés, c’est-à-dire correspondre à des régions spécifiques du cerveau. Sa théorie suggérait que certaines activités cognitives spécifiques pourraient être gérées par un module spécialement conçu—chaque module fonctionnerait de manière plutôt autonome sur le problème spécifique qu’il avait évolué pour résoudre, par exemple, l’analyse des formes, la perception des couleurs, l’analyse des relations dans l’espace, guidage visuel des mouvements du corps, reconnaissance faciale, analyse grammaticale du langage parlé, reconnaissance vocale, et bien d’autres. Mais à son avis, les modules ne traiteraient que des problèmes spécifiques, chacun n’impliquant qu’un certain type d’informations. C’est ce qu’on appelle l’encapsulation de l’information.

Toutes ces considérations concernent la question de l’architecture de l’esprit: quels sont les différents types de traitement de l’information qui s’y déroulent, dans quels différents types de structure, et avec quelles connexions entre ces différents types de structure? Mais une chose est claire: les modules sont désormais généralement reconnus comme jouant un rôle clé à cet égard. Et les modules en question sont darwiniens au sens décrit ci-dessus. L’idée a été mise en avant en 1987 dans les travaux de l’anthropologue américain John Tooby et de la psychologue américaine Leda Cosmides [7]. Malheureusement, nous ne pouvons ici qu’effleurer ce sujet passionnant, mais regardons tout de même d’un peu plus près.

---

<sup>2</sup>Naturellement, c’est une manière de parler qui pourrait être trompeuse. Mère nature n’a fait aucun calcul ici. Ces développements sont survenus par hasard dans divers contextes où ils se sont révélés favorables aux gènes qui prédisposaient l’animal donné à les exprimer; favorable dans le sens où ces gènes pourraient ensuite passer à la génération suivante et donc proliférer.

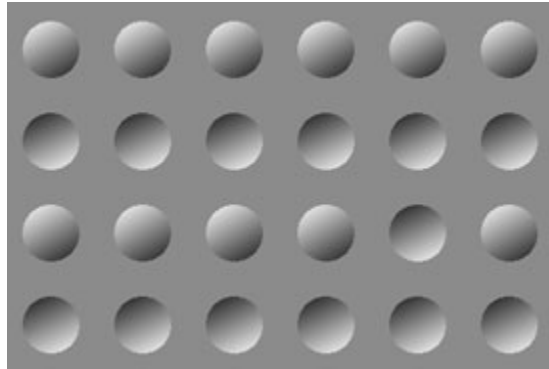


Figure 5: Que voyez-vous? La rangée du haut semblerait dépasser, mais c'est parce que votre esprit a fait une hypothèse sur la source de lumière—sans vous consulter

Selon l'hypothèse de modularité massive, l'esprit humain, ou tout esprit animal, est un ensemble de modules spécialisés, dont chacun a évolué pour résoudre un problème spécifique rencontré par nos ancêtres, humains ou non. Chaque module s'occupe d'une application spécifique, ne traitant que les informations pertinentes pour cette application. Il est autonome, rapide et inconscient. La rapidité est bien sûr essentielle dans de nombreuses situations, où l'efficacité de la réaction, voire la survie elle-même, peut être en jeu. Notez en particulier l'idée que dans de nombreux cas, nous ne sommes même pas conscients de ce que font nos modules. C'est l'une des grandes découvertes de la psychologie cognitive: nous ne sommes en fait conscients de presque rien dans lequel notre cerveau est engagé. Comme mentionné ci-dessus, il est généralement accepté aujourd'hui que nous—c'est-à-dire le soi—ne sommes qu'une construction par notre cerveau, en réponse à l'introspection, en grande partie à des fins de communication, et que nous sommes à peine impliqués dans la plupart des activités cognitives, malgré l'impression que notre propre esprit peut essayer nous donner.

Nous avons donc des modules qui font des inférences dans des domaines nombreux et variés, chacun profitant d'une certaine régularité dans la nature, c'est-à-dire quelque chose qui se présente toujours d'une certaine manière, ou quelque chose qui se produit toujours quand quelque chose d'autre vient de se produire. Et tout cela se fait sans même que nous nous en rendions compte. Regardez la figure 5, qui montre ce qui semble être une dalle de matériau avec des rangées de bosses et de creux. La plupart des gens voient la rangée du haut comme une rangée de bosses et la seconde comme une rangée de creux. C'est qu'un module de notre esprit a supposé, sans nous consulter, que la dalle était éclairée par le haut. Mais imaginez un instant—cela demande un effort—que la dalle soit éclairée par le bas. Maintenant, la rangée du haut devient une rangée de creux et la seconde une rangée de bosses. En fait, nous pouvons passer d'une perception à l'autre, avec un effort suffisant. Naturellement, tous nos ancêtres, humains ou non, auraient été habitués à ce que les choses soient éclairées d'en

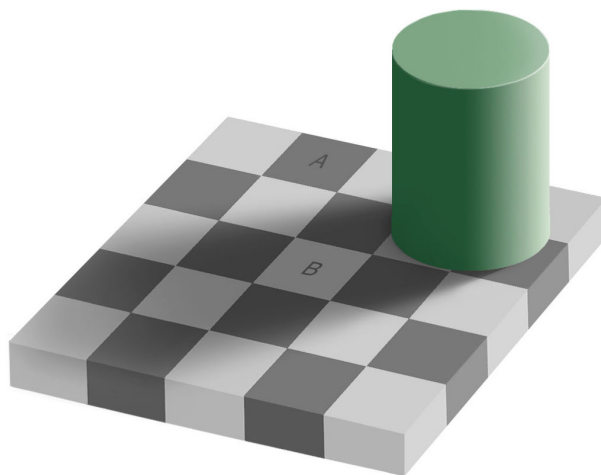


Figure 6: Cela peut être difficile à croire, mais bien que A soit évidemment un carré noir et B un blanc, le niveau de gris dans les deux carrés est la même. Jetez un œil à la figure 7 pour voir la démonstration. Image de [9], vectorisée par Pbroks13, CC BY-SA 4.0

haut, par le soleil—une régularité dans notre monde.

Voici donc un point important. Un module fait une inférence, c'est-à-dire qu'il produit de nouvelles informations sur la base d'anciennes, mais il ne nécessite pas, ni ne nous fournit, aucune raison de justifier son inférence. C'est en ce sens qu'il fonctionne automatiquement. Il raisonne sans raisonnement conscient. Ceci est illustré à nouveau dans le célèbre damier d'Adelson montré à la figure 6. Nous n'avons aucune raison *consciente* de voir le carré A plus sombre que le carré B, mais nous voyons des choses comme ça. Les deux carrés sont identiques dans l'image—le niveau de gris est la même dans les deux carrés (voir la figure 7)—mais notre cerveau est intervenu pour appliquer sa connaissance des ombres.

Une autre expérience, ou plutôt un type d'expérience, illustre magnifiquement comment notre cerveau est le juge final des informations qu'il reçoit de nos yeux, et bien en dehors de tout effort conscient de notre part. Ces expériences exploitent le paradigme de la contingence du regard (*gaze contingency paradigm* en anglais [10]): pendant que quelqu'un regarde un écran d'ordinateur, il est possible d'informer l'ordinateur en temps réel de l'emplacement exact du regard de la personne, puis d'agir sur le reste de l'image à l'écran de telle manière que la personne ne se rende pas compte. Comme nous avons dit plus haut, ces techniques de suivi oculaire sont largement utilisées en psychologie pour étudier notre façon de voir le monde. Par exemple, quelqu'un lit une nouvelle à l'écran et reste complètement inconscient du fait que la quasi-totalité du texte est brouillée à chaque instant de temps. Quelqu'un qui se tient derrière le lecteur trouvera étonnant qu'il parvienne à en extraire le moindre sens.

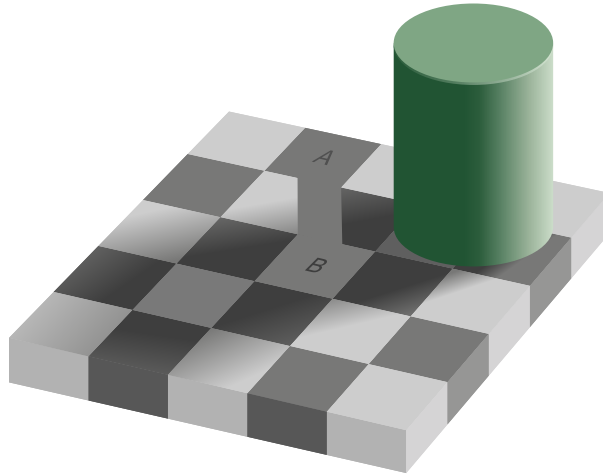


Figure 7: Le niveau de gris est le même dans les carrés A et B. Image de [9], vectorisée par Pbroks13, CC BY-SA 4.0

Alors qu'est-ce que cela a à voir avec les modules? En fait, un module construit notre image interne de ce qui est visualisé. Dans cet exemple, l'image interne ressemble peu à ce qui s'y trouve réellement. C'est parce que le module en question "suppose" que ce que nous regardions ailleurs il y a un instant sera toujours là, quasiment inchangé, une hypothèse naturelle dans le monde ordinaire de la nature. Il exploite une régularité dans le monde: les choses ne changent pas si vite.

Un autre exemple fascinant de module, qui montre que certains modules peuvent avoir évolué il y a très longtemps, ou peut-être avoir évolué plusieurs fois chez différentes espèces, concerne la façon dont les êtres vivants se guident dans des espaces étroits sans toucher aux limites de ces espaces. Lorsque nous traversons un tunnel, les murs semblent reculer à une certaine vitesse. Si le mur de droite est plus proche que le mur de gauche, il semblera reculer plus rapidement. Cette expérience, connue sous le nom de flux optique, peut nous dire quelle paroi est la plus proche, et si nous souhaitons passer par le milieu du tunnel, il suffit d'ajuster notre trajectoire pour que le flux optique soit le même de chaque côté. Ceci est bien décrit par Sloman et Fernbach dans leur excellent livre *The Knowledge Illusion* [11, 12]:

Un endroit où vous rencontrez un flux optique est sur l'autoroute. Les lignes sur la route peintes par le ministère des transports sont là pour vous garder sur le droit chemin. Tant que le flux de lignes d'un côté de vous semble se déplacer au même rythme que le flux de l'autre côté, vous resterez dans votre voie. Nous le savons grâce à des expériences réalisées dans des simulateurs de conduite. Si vous mettez quelqu'un dans un simulateur avec un écran d'ordinateur et

que les lignes vont plus vite d'un côté que de l'autre, la personne dérivera vers le côté avec les lignes plus lentes.

Imaginez un robot GOFAI faisant tous les calculs nécessaires pour voler entre deux poteaux sans les toucher. Ce n'est pas un problème évidemment si vous pouvez calculer assez rapidement. Mais le flux optique est tellement plus facile, à condition d'avoir un système pour convertir les estimations sensorielles du flux optique directement en effets moteurs.

Ce qui est fascinant, c'est que les abeilles utilisent aussi ce système [13]. On peut faire passer les abeilles d'une ruche dans un tunnel pour se rendre dans les prairies, puis projeter une image mouvante du mur sur un côté. Si le mur de droite semble reculer plus rapidement, les abeilles ne passent plus par le centre du tunnel, mais plus à gauche. Peut-être que ce module a évolué chez un ancêtre commun des abeilles et des humains, ou peut-être a-t-il évolué séparément plusieurs fois. C'est une astuce tellement utile, ce ne serait pas si surprenant.

Sloman et Fernbach encore [11, p. 100]:

Toutes les études pertinentes montrent que les gens (et les insectes) ne sont pas des constructeurs de modèles classiques qui s'engagent dans des tas de calculs, rythmés de temps en temps par l'action. Au contraire, les gens—ou du moins leurs cerveaux—utilisent leurs connaissances du monde, comme le flux optique des surfaces, pour simplifier ce qu'ils font.

Les modules évoluent ou se développent lorsqu'il existe une régularité dans le monde qui peut être exploitée pour tirer des inférences, et lorsqu'il est adaptatif de le faire. Ces inférences ne sont pas le résultat d'un raisonnement au sens où nous l'entendons normalement, comme vous pouvez le voir dans les exemples. Ce sont des conséquences directes d'une partie du cerveau sur les autres. Le cerveau n'applique aucune théorie. Cela ne nécessite aucune pensée consciente de la part du soi, ni en fait aucune implication du soi du tout.

Un autre exemple donné par Sloman et Fernbach concerne la façon dont un joueur de baseball attrape une balle [11, p. 97]. Imaginez que la balle a été frappée dans votre direction. Comment déterminez-vous où vous devez être pour l'attraper? Imaginez le robot GOFAI dans cette situation, avec les lois du mouvement de Newton programmées dedans! Comme le disent Sloman et Fernbach:

Le robot resterait là à réfléchir pendant quelques instants—avec un peu de chance pas trop longtemps—puis se déplacerait vers la bonne position (si ses calculs ont donné la bonne réponse).

La remarque entre parenthèses pourrait, par exemple, faire référence à la possibilité que la résistance au vent ne soit pas programmé dans le robot. Heureusement, le joueur de baseball n'a pas besoin de suivre cette méthode. Voici à nouveau Sloman et Fernbach:



Il s'avère qu'il existe un moyen plus simple d'attraper une balle qui n'implique aucune réflexion. Au lieu de calculer des trajectoires, il existe une astuce qui vous amènera là où la balle va atterrir. Si une balle vient dans votre direction, une chose naturelle à faire est de regarder la balle alors qu'elle s'élève dans les airs, en levant la tête pour suivre la balle avec votre regard lorsqu'elle se dirige vers vous. La direction de votre regard spécifie un angle par rapport au sol. Voici l'astuce: pour arriver là où la balle va atterrir, il suffit d'avancer ou de reculer pour que cet angle augmente toujours à un rythme constant. Afin de garder un œil sur la balle après qu'elle a été frappée, vous devrez continuellement incliner votre tête (ou vos globes oculaires) vers le haut pour suivre le mouvement de la balle. Ce qui peut vous surprendre, c'est que vous devrez continuer à lever le regard même après que la balle commence à descendre. Si vous regardez un joueur courir pour attraper une balle, vous le verrez ajuster la direction et la vitesse de son corps afin que son regard se déplace toujours vers le haut au même rythme constant. Ces ajustements le conduisent au bon endroit pour intercepter le ballon. Ensuite, tout ce qu'il a à faire est de l'attraper.

C'est quelque chose qui peut être observé sur le terrain. Mais avec le recul, on peut aussi prouver mathématiquement que ça marchera, en utilisant un peu de géométrie.

Pour résumer, les tâches sophistiquées ne sont pas accomplies en effectuant d'abord une planification et un calcul exhaustifs. Nos actions nous sont souvent faciles, non pas parce que les tâches en question sont intrinsèquement simples, mais parce qu'il existe des solutions simples à des problèmes autrement compliqués. Les tâches de haut niveau sont réalisées en combinant des compétences plus simples qui sont à leur tour des hiérarchies de compétences encore plus simples. Et on constate qu'il n'y a pas de problème de cadre. Chaque espèce fait évoluer précisément les modules dont elle a besoin pour résoudre les problèmes rencontrés par ses ancêtres, et le module concerné travaille toujours précisément sur les informations qu'il doit prendre en compte, ignorant les informations qu'il n'a pas besoin de prendre en compte. Les modules évoluent ou se développent lorsqu'il y a une régularité à exploiter dans l'inférence et lorsqu'elle est adaptative pour le faire, c'est-à-dire lorsqu'elle donne à l'animal un avantage pour sa survie et sa reproduction. C'est en ce sens que l'évolution est opportuniste, et c'est aussi en ce sens que notre esprit modulaire est opportuniste.

## Deux types de raisonnement

Les modules présentés ci-dessus confirment qu'on peut considérer tout traitement de l'information comme une forme de raisonnement, même lorsqu'aucune conscience n'est impliquée, puisque nous pouvons considérer les modules comme tirant des inférences et appliquant des connaissances sur les régularités du monde pour le faire. Mais, bien sûr, il y a aussi quelque chose que nous appelons

le raisonnement conscient. C'est quelque chose que nous pouvons écrire ou présenter sur les diapositives d'une conférence, ou encore simplement expliquer à un collègue. Dans son livre fascinant *Thinking, Fast and Slow*, Daniel Kahneman fait une telle distinction à propos de la pensée en général [14]. La *pensée rapide* est réalisée par le type de modules que nous venons de décrire: elle est rapide, automatique, inconsciente et sans effort. La *pensée lente* en revanche est lente, contrôlée, linéaire comme le langage utilisé pour la formuler, et demande un effort considérable.

La pensée rapide donne lieu à ce que l'on appelle des *biais cognitifs*, et le livre de Kahneman en présente beaucoup. Lui et son collègue Amos Tversky ont passé une grande partie de leur vie à les étudier, en particulier dans le contexte de l'économie. C'est pourquoi Kahneman, bien que psychologue cognitif, a finalement reçu le prix Nobel d'économie. Cette approche psychologique de la façon dont les investisseurs investissent *vraiment* et comment les gestionnaires financiers prennent leurs décisions *vraiment* a donné naissance à une nouvelle branche d'étude appelée *économie comportementale*, à contraster avec l'approche standard à l'étude de l'économie qui suppose que les investisseurs et les gestionnaires financiers sont purement rationnels dans un certain sens.

On en dira plus sur les biais cognitifs dans un instant. Mais tournons-nous d'abord vers une forme de raisonnement qui, sans aucun doute, nous distingue des autres animaux, ne serait-ce que parce qu'elle implique le langage.

## Raisonnement moral

Nous avons déjà noté comment le test de logique proposé par Wason a révélé quelque chose d'important sur notre capacité à être logique. Il a découvert un lien étroit avec la question de la moralité. Revenons donc à la question de l'éthique en compagnie du psychologue évolutionniste américain Jonathan Haidt. Son livre *The Righteous Mind* est une excellente introduction au sujet [15]. Notez en particulier le sous-titre *Pourquoi les gens bien sont divisés par la politique et la religion* (*Why Good People Are Divided by Politics and Religion*).

Haidt énonce cinq, ou peut-être six, fondements évolutifs de la moralité. Ce n'est pas le lieu de tous les décrire, mais il y en a un qui me semble particulièrement pertinent ici. C'est le fondement de l'équité et de la réciprocité. Chaque fondation a évolué pour relever un défi adaptatif, en l'occurrence le défi de récolter les fruits des partenariats bilatéraux, c'est-à-dire de bénéficier de la coopération. Et chaque fondation a coévolué avec certaines émotions, vraisemblablement uniques aux humains. Dans ce cas, colère contre le tricheur, gratitude envers ceux qui coopèrent et culpabilité quand nous avons nous-mêmes triché. Selon Haidt, la moralité est en grande partie une solution évoluée au problème du passager clandestin. Et bien sûr, cela implique de raisonner sur les règles, ce qui rejoint bien ce qui a été révélé par les tests Wason: notre logique est affûtée lorsqu'il s'agit de faire respecter les règles par les autres.

Nous reviendrons sur Haidt dans un instant, mais citons d'abord deux auteurs français, Hugo Mercier et Dan Sperber, et leur livre remarquable, voire

révolutionnaire, *L'énigme de la raison*, qui a le sous-titre *Une nouvelle théorie de la compréhension humaine* dans la version anglaise du livre. Il propose une approche modulaire de la raison parfaitement pertinente au sujet de cet article, mais qui serait trop longue à exposer ici. Cependant, l'idée qu'il y a quelque chose d'énigmatique dans la raison humaine, exprimée dans le titre, sera exposée plus en détail ci-dessous.

Mais notons d'abord que Mercier et Sperber sont entièrement d'accord avec Haidt: les raisons sont des constructions sociales principalement pour la consommation sociale. En effet, pour tous ces auteurs, le raisonnement moral a une fonction stratégique. Nous produisons des raisons afin de:

- justifier nos pensées et nos actions auprès des autres,
- produire des arguments pour convaincre les autres de penser et d'agir comme nous,
- évaluer les raisons que les autres produisent pour se justifier.

La clé ici est bien sûr toujours la coopération. Nous avons besoin qu'on nous fasse confiance et nous avons besoin de savoir à qui nous pouvons faire confiance.

Revenant à Haidt, nous constatons qu'il ajoute quelque chose d'autre, qui est particulièrement intéressant pour la présente discussion. Selon son évaluation, *les intuitions viennent en premier* et le raisonnement stratégique en second. L'idée de base est que toutes nos raisons morales proviennent en définitive de l'intuition. Mais qu'est-ce que l'intuition? C'est une idée dont nous prenons soudain conscience. Quelque chose qui jaillit donc de l'activité inconsciente de notre cerveau. Et c'est intéressant parce que, si Haidt a raison, même ce raisonnement moral qui nous distingue si clairement des autres animaux relève finalement de l'inconscient. Il donne plusieurs exemples d'expériences menées par des psychologues cognitifs pour illustrer cela. En voici trois:

- Dans une expérience simple pour montrer comment la familiarité peut influencer nos préférences [17], on demande aux gens d'évaluer des pictogrammes japonais sur une échelle de un à dix—comme les inventeurs de Tripadvisor vous le diront, les gens sont prêts à évaluer n'importe quoi sur une échelle de un à dix. Cependant, lors de la lecture des consignes à l'écran, certains pictogrammes s'affichent pendant un temps très court, trop court pour que l'individu puisse s'en apercevoir. Les notes les plus élevées ont tendance à être attribuées aux pictogrammes affichés de manière subliminale. D'un point de vue évolutif, il est facile de comprendre que ce serait une bonne stratégie de privilégier les choses que nous connaissons.
- Un autre facteur influençant notre jugement moral est une mauvaise odeur dans les environs [18]. Les gens sont invités à juger des situations moralement douteuses lorsqu'une mauvaise odeur est introduite dans leur voisinage. Ils ne se rendent pas compte que l'odeur a été introduite délibérément. Leurs jugements se font plus durs. De même, s'il leur a été demandé de se laver soigneusement les mains avant de commencer [19].

- Et la durée des peines prononcées par des juges à qui on a demandé d’attribuer des peines à des criminels dans des études de cas peut être influencée en les présentant de diverses manières des nombres plus ou moins grands avant qu’ils ne commencent, encore une fois de telle manière qu’ils ignorent tout lien entre les chiffres présentés et le but de l’expérience. Un aperçu détaillé de cette découverte importante peut être trouvé dans [20]. Ces effets sont connus des psychologues cognitifs sous le nom d’amorçage ou d’ancrage.

La conclusion surprenante ici est donc que, même si notre raisonnement moral s’exprime dans le langage et nous distingue donc sans aucun doute des autres animaux, nous—c’est-à-dire le soi—le contrôlons beaucoup moins que nous ne l’imaginons, car il surgit fondamentalement de l’activité inconsciente de notre cerveau sous la forme d’intuitions, ou du moins est fortement influencé par celles-ci. Il est intéressant de citer le moraliste français François VI, Duc de La Rochefoucauld, Prince de Marcillac (1613–1680):

Tout le monde se plaint de sa mémoire, et personne ne se plaint de son jugement.

Tant de choses étaient donc connues il y a longtemps. Quoi de neuf ici? Une chose est que nous avons des expériences démontrant la prévalence du phénomène, et une autre est qu’en adoptant une approche évolutive, nous arrivons à comprendre *pourquoi* nous sommes comme ça. Ou plutôt, *pourquoi nous ne sommes pas vraiment comme nous nous imaginons*. Au fond, c’est parce que le soi n’est qu’une représentation de l’animal humain produite par le cerveau pour les besoins de ses interactions avec d’autres personnes, d’autres soi. Et il n’y a aucune raison d’imaginer que cela prendra en compte, ou représentera, la totalité de l’énorme quantité d’autres activités du cerveau, telles que les *vraies* raisons pour lesquelles il fait ce qu’il fait.

Selon Mercier et Sperber, tous nos processus mentaux, tant affectifs que cognitifs, sont désormais considérés comme largement voire totalement inconscients. De plus, nous avons peu ou pas d’accès introspectif à nos propres processus mentaux et nos rapports verbaux de ces processus sont souvent de pures inventions. En effet, nous nous trompons systématiquement en supposant que nous avons une connaissance introspective directe de nos états mentaux et des processus qui les produisent.

Ainsi, pour en revenir à Haidt, les intuitions viennent en premier et, dans des circonstances normales, nous amènent à nous engager dans un raisonnement socialement stratégique. Dans son livre [15], Haidt traite assez longuement de la nature de cette opération stratégique. Voici quelques conclusions:

- La pensée morale ressemble plus à un politicien à la recherche de voix qu’à un scientifique à la recherche de la vérité.
- Ce que nous appelons le raisonnement conscient, c’est-à-dire la formulation linguistique de nos pensées, fonctionne comme un attaché de presse qui justifie automatiquement toute position prise par le président.

- En matière morale et politique, nous sommes souvent groupistes plutôt qu'égoïstes. Nous déployons nos capacités de raisonnement pour soutenir notre équipe et démontrer notre engagement envers cette équipe.
- Le raisonnement peut nous conduire à presque toutes les conclusions auxquelles nous souhaitons parvenir. Lorsque nous entendons quelque chose avec lequel nous sommes d'accord, nous demandons: Puis-je le croire? Et il est facile de trouver une raison de dire "Oui!" Et lorsque nous entendons quelque chose avec lequel nous ne sommes pas d'accord, nous demandons: Dois-je le croire? Et il est facile de trouver une raison de dire "Non!"

On peut citer ici Benjamin Franklin, scientifique, inventeur et l'un des pères fondateurs de la constitution américaine, bien qu'il ne soit jamais devenu président:

Il est si pratique d'être une créature raisonnable, car cela permet de trouver ou de fabriquer une raison pour tout ce que l'on a envie de faire.

Ainsi, lorsque le langage a évolué, pourquoi avons-nous développé un avocat intérieur, et non un juge intérieur ou un scientifique intérieur? N'aurait-il pas été plus adaptatif pour nos ancêtres de découvrir la vérité sur qui a fait quoi et pourquoi, plutôt que d'utiliser toute leur intelligence pour trouver des preuves de ce qu'ils voulaient croire? La réponse est simple du point de vue de l'évolution: ce qui était le plus important pour la survie de nos ancêtres n'était pas la vérité, mais la réputation. Est-ce qu'on appellerait ça rationnel? Eh bien, dans un sens, ça l'était. En termes de survie et de reproduction, il était essentiel d'être bien accepté et d'inspirer la confiance dans son propre groupe.

## Les biais cognitifs et le paradoxe de la raison

Il y a beaucoup de discussions sur les biais cognitifs ces jours-ci. Le livre de Kahneman [14] est une bonne introduction. Mais voici une définition tirée au hasard d'un site web (traduite de l'anglais) [21]:

Un biais cognitif est une erreur systématique de pensée qui se produit lorsque les gens traitent et interprètent les informations sur le monde qui les entoure et affecte les décisions et les jugements qu'ils prennent.

Le même site web poursuit en disant:

Le cerveau humain est puissant mais soumis à des limites. Les biais cognitifs sont souvent le résultat de la tentative de votre cerveau de simplifier le traitement de l'information. Les biais fonctionnent souvent comme des règles empiriques qui vous aident à donner un sens au monde et à prendre des décisions relativement rapidement.

Bien que contradictoire—une erreur systématique nous aidant à donner un sens au monde?—ce site web capture certainement quelque chose de la saveur de ces effets psychologiques!

Certains considèrent les biais cognitifs comme une sorte de bizarrerie parce qu'ils ne semblent pas particulièrement rationnels, ce qui a donné naissance à ce qu'on appelle le paradoxe de la raison. Car nous pouvons nous demander comment les humains ont réussi à créer un tel corpus de connaissances fiables, à se rapprocher si bien de la vérité sur le monde, comme nous pouvons le voir dans notre maîtrise de la technologie. Une solution à cette énigme, également évoquée dans le titre du livre de Mercier et Sperber [16], sera esquissée ci-dessous. Mais il faut garder à l'esprit que tous ces biais, et des dizaines d'entre eux ont été identifiés à ce jour, sont très probablement universels, c'est-à-dire communs à tous les humains, ce qui suggère qu'ils ont une base largement génétique; ce qui suggère à son tour qu'ils servent à quelque chose d'utile dans notre lutte pour survivre assez longtemps pour se reproduire. À moins bien sûr que quelque chose ait changé si radicalement dans notre monde moderne qu'ils nous rendent maintenant un mauvais service.

Prenons donc deux exemples pour illustrer ces choses. Le premier est le *biais de disponibilité*:

La tendance, sur une question donnée, à privilégier et surestimer les informations immédiatement disponibles à notre mémoire.

Pour quelqu'un qui vit dans un petit groupe de chasseurs-cueilleurs, c'est probablement une bonne stratégie, puisque presque toutes les informations disponibles concernent son entourage proche et l'environnement local. Mais la situation est très différente aujourd'hui. La couverture médiatique nous fournit d'énormes quantités d'informations, dont la plupart nous concernent à peine dans notre vie quotidienne. Et en fait, cela détermine nos évaluations de la fréquence et du risque. L'une des raisons pour lesquelles les gens ont beaucoup plus peur de voyager en avion qu'en voiture est que les accidents d'avion sont tellement couverts. Que se passerait-il si chaque jour le journal télévisé commençait par montrer tous les accidents de voiture du pays ? Est-ce que les gens conduiraient de la même manière? C'est un exemple classique, mais il y a plus à dire et Pinker le dit parfaitement [1]:

La presse est une machine à disponibilité. Elle nous sert des anecdotes qui alimentent notre impression de ce qui est fréquent [...] ceci nous laisse dans l'ignorance de la prévalence réelle d'un événement.

Les distorsions ne sont pas aléatoires, mais nous orientent vers le morbide. Les événements soudains sont généralement mauvais—une guerre, une fusillade, une famine, un effondrement financier—mais les bonnes choses tiennent de la morne plaine. C'est un pays en paix ennuyeux, une région oubliée où les gens sont en bonne santé et mangent à leur faim. Et lorsque le progrès survient, il ne se fait pas en un jour, mais se développe lentement en quelques points de pourcentage de plus par an et transforme le monde à la

dérobée. Comme le fait remarquer l'économiste Max Roser, les sites d'information auraient pu publier "137000 personnes ont échappé à l'extrême pauvreté" tous les jours au cours des vingt-cinq dernières années. Mais ils n'ont jamais publié un tel article, car il n'y a jamais eu de jeudi d'octobre où cela s'est soudainement passé. Et c'est ainsi que l'une des plus grandes évolutions de l'histoire de l'humanité—un milliard et quart de personnes échappant à la misère—est passée inaperçue.

Pinker poursuit en soulignant un risque très réel pour nos démocraties:

L'ignorance motivée par la disponibilité a de quoi être corrosive. En étant bloqué dans une boucle mentale de catastrophes et d'échecs, il est facile d'être cynique quant aux capacités qu'ont la science, la démocratie libérale et les institutions de coopération mondiale d'améliorer la condition humaine. Il peut en résulter un fatalisme léthargique ou un radicalisme inconsidéré: on appelle à casser la machine, à assécher le marécage ou à donner du pouvoir à un démagogue promettant qu'il est "le seul à pouvoir régler les choses".

Voilà pour l'heuristique de disponibilité alors. Dans le monde moderne, avec son abondance d'informations souvent non pertinentes, ce biais peut facilement nous égarer.

Le biais cognitif le plus connu est peut-être le *biais de confirmation*. Il consiste à:

- privilégier les informations qui confirment nos idées reçues,
- accorder moins d'importance aux informations qui semblent les contredire.

Vraisemblablement, toute personne honnête se reconnaîtra ici! C'est ce biais qui soulève immédiatement la question de savoir comment nous parvenons à produire des connaissances fiables—le paradoxe de la raison. Et puisque nous faisons la promotion de l'évolution en tant qu'outil explicatif par excellence pour comprendre l'activité cognitive humaine, quel pourrait être l'avantage évolutif du biais de confirmation?

Pour commencer, selon Mercier et Sperber, nous sommes plus groupistes qu'égoïstes et le biais de confirmation est en fait un biais partisan. Cela est clair dans le débat politique et peut être observé à travers toute l'activité humaine. Mais en réalité, cela peut aussi fournir un moyen efficace de partager la charge de travail cognitive lorsqu'un groupe de personnes a un problème à résoudre. Car il s'avère que nous sommes paresseux lorsqu'il s'agit d'évaluer nos propres raisons et justifications, mais extrêmement efficaces pour évaluer celles des autres. Autrement dit, personne n'est bon pour repérer ses propres erreurs, mais nous sommes très, très bons pour repérer celles des autres.

Pour comprendre comment cela pourrait être utile, nous devons nous demander quel est le contexte approprié pour le raisonnement. D'après ce que nous avons déjà dit sur le rôle social du raisonnement, il semble plausible qu'il

s'agisse d'une discussion dans laquelle les gens échangent des arguments et des justifications et où il y a un intérêt commun à trouver la vérité, ou du moins à trouver le meilleur compromis. Nous n'avons qu'à réfléchir un instant sur la myriade d'exemples de coopération sociale pour voir que de telles discussions ont lieu, même si nous avons tendance à nous concentrer sur des situations où le débat politique devient un non-débat à cause d'une partisanerie trop zélée (ce qui joue ensuite dans notre biais de disponibilité).

En effet, la raison est beaucoup plus efficace pour évaluer de bons arguments que pour les produire. Il est facile de montrer que les capacités de résolution de problèmes d'un groupe de personnes sont largement supérieures en général à celles de n'importe quel individu. Par exemple, les groupes sont bien meilleurs que les individus pour résoudre les tâches de Wason. Lorsque les gens reçoivent l'explication correcte pour la première test décrit ci-dessus, ils le reconnaissent et l'acceptent généralement immédiatement. Nous revenons au point crucial que l'esprit n'a pas évolué dans un contexte d'individus travaillant seuls pour résoudre des problèmes. Nous sommes conçus pour coopérer, et penser par soi-même n'est pas un bon moyen d'obtenir de bonnes solutions. Permettez-moi de citer à nouveau Keynes:

Il est étonnant de voir quelles bêtises on peut croire si l'on pense trop longtemps seul.<sup>3</sup>

On pourrait ajouter, si l'on réfléchit trop longtemps dans un petit groupe de personnes ayant les mêmes idées.

Alors, quand changeons-nous d'avis sur des questions morales? La réponse est que la persuasion raisonnée devient possible lorsque nous interagissons avec les autres. Notez alors le gouffre entre le raisonnement d'un individu et notre raisonnement collectif, corrigé des biais comme ceux décrits ci-dessus. Et voici la solution au paradoxe de la raison: le meilleur raisonnement vient quand nous raisonnons ensemble, un thème que je développerai dans un instant. Mais notez d'abord l'importance du langage dans ce que nous concevons généralement comme le raisonnement, et donc l'importance de la conscience. Cela nous amène à une question séculaire: le langage est-il apparu comme un enrichissement de la communication ou comme un enrichissement de la pensée? La réponse est certainement les deux à la fois.

Même seul, en exprimant un argument avec des mots, nous pouvons le voir plus objectivement en anticipant les problèmes qui pourraient être soulevés *par quelqu'un d'autre*. C'est ce qui se passe lorsque nous préparons les diapositives d'une conférence ou quand nous écrivons un livre ou un article. Et bien sûr, il est essentiel que les arguments ainsi préparés soient effectivement présentés au public et aux lecteurs pour leur critique. C'est une partie importante de la méthode scientifique, ou plutôt de la façon dont les scientifiques mènent leur travail. La science est une entreprise collective. En fait, c'est l'entreprise collective *par excellence*, peut-être la plus grande réalisation collective de l'humanité. Les scientifiques publient dans des revues internationales à comité de lecture afin

---

<sup>3</sup>Et il a conclu ici "particulièrement en économie".



que leurs travaux puissent être critiqués et corrigés si nécessaire, et ils présentent leurs travaux lors de conférences internationales dans le même but.

La science met en œuvre un processus d'autocorrection impliquant toute une communauté de personnes dont l'engagement global est de connaître la réalité du monde. Cela correspond parfaitement à l'idée ci-dessus que le contexte approprié pour le raisonnement est une discussion dans laquelle les gens échangent des arguments et des justifications et où il y a un intérêt commun à trouver la vérité, dans ce cas, la vérité sur le monde réel, sur la façon dont les choses sont vraiment dans le monde pour autant que nous puissions l'établir. Le résultat est un ensemble de connaissances fiables, voire extrêmement fiables. Il n'y a pas de paradoxe ici. Cette connaissance est le résultat d'un raisonnement collectif.

Il est intéressant de considérer une vision plutôt non conventionnelle de la découverte scientifique. La plupart des gens associent la découverte scientifique à quelques noms familiers, considérés comme des "génies", comme Galilée, Newton et Einstein. Ce n'est pas simplement une simplification, mais une déformation de ce qui se passe réellement. Ces personnes n'ont jamais travaillé seules. Chacun d'eux était entouré de livres et d'articles écrits par d'autres et par des collègues, ou des ennemis, essayant de faire la même chose. Au début, bien sûr, il n'y avait pas beaucoup de scientifiques, mais même à l'époque de Galilée, il y en avait beaucoup d'autres intéressés par des questions similaires. Les noms dont nous nous souvenons n'étaient que les chanceux qui sont tombés sur les bonnes solutions, ou en tout cas les meilleures solutions. Et si la science avance à une vitesse toujours plus grande au fil du temps, ce n'est certainement pas parce qu'il y a plus de génies à l'œuvre aujourd'hui. C'est juste qu'il y a plus de scientifiques au travail.

## L'illusion de la connaissance

Je termine cette discussion en revenant au livre de Fernbach et Sloman [11]. L'illusion à laquelle ils font référence dans le titre n'est pas une sorte d'illusion sur la connaissance produite par la science. Bien au contraire. Les auteurs sont eux-mêmes des scientifiques et soutiennent l'effort scientifique dans leur livre. Ils commentent le fait que les connaissances de tout individu sont en réalité extrêmement limitées, même si nous avons tendance à imaginer le contraire. N'importe quel universitaire peut en témoigner. Nous sommes censés être des experts dans notre domaine très restreint, mais souvent, lorsque nous recevons une question d'un étudiant ou d'un collègue, nous hésitons un instant et promettons de revenir vers eux plus tard! Nous nous précipitons alors pour essayer de nous souvenir de ce que nous sommes censés savoir sur la question qui a été soulevée. Si nécessaire, nous nous tournons vers nos notes, puis vers notre bibliothèque—les universitaires conservent toujours des centaines de livres spécialisés à cet effet—et si cela n'aide pas, nous commençons à envoyer des courriels à des collègues! Mais au final, parce que nous *sommes* experts dans notre domaine, soit nous obtenons la réponse, soit nous nous rendons compte qu'un nouveau problème a été identifié.



Figure 8: Schéma utilisé par Lawson dans l'expérience de "cycologie"

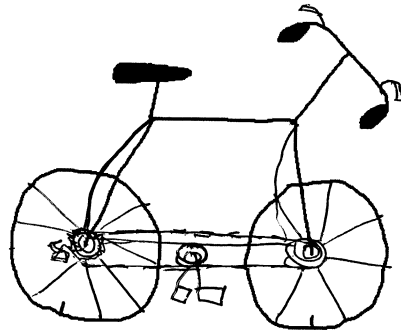


Figure 9: Exemple de schéma produit dans l'expérience de "cycologie"

La réalité est que nous ignorons en grande partie à quel point nous savons peu de choses, car nous confondons les connaissances dans nos têtes avec les connaissances auxquelles nous avons accès. Être un expert signifie savoir où trouver des réponses dans notre domaine. Cela ne veut pas dire que nous avons toujours les réponses prêtes à l'emploi. À cet égard, Fernbach et Sloman font également référence à ce qu'ils appellent *l'illusion de profondeur explicative*. Pour illustrer cela, ils décrivent des expériences menées par les scientifiques cognitifs américains Frank Keil et Leon Rozenblit [23]. Dans un premier temps, les volontaires ont été invités à évaluer leur compréhension de choses bien connues comme les fermetures éclair, les chasses d'eau, les machines à coudre et bien d'autres objets du quotidien. Dans un deuxième temps, on leur a demandé de décrire le plus précisément possible le fonctionnement de chacun de ces systèmes, précisant toutes les étapes de leur fonctionnement. Pour la plupart des candidats, cette étape n'a jamais duré très longtemps! Lorsqu'on leur a demandé à la troisième étape de réévaluer leur compréhension des différentes choses, les valeurs ont fortement chuté.

Dans un exercice similaire, la psychologue expérimentale Rebecca Lawson [24] a demandé aux candidats de compléter un schéma très sommaire d'un vélo (voir la figure 8) avec des détails sur la chaîne et le mécanisme de pédalage. Les résultats étaient amusants (voir la figure 9). Et pourtant, si le vélo est de-

vant nous, on peut bien sûr expliquer son fonctionnement, c'est-à-dire expliquer l'enchaînement causal des événements entre l'appui sur une pédale et la rotation des roues. Mais pour la plupart d'entre nous, ces objets du quotidien doivent au moins être présents pour les expliquer. Dans un sens donc, nous utilisons aussi le monde extérieur comme une extension de notre mémoire et de notre compréhension des choses. Le cerveau est efficace: il ne stocke pas plus que nécessaire.

Sloman et Fernbach ont mené des expériences similaires, mais dans un domaine très différent. Ils sont préoccupés par des questions de société, comme le niveau élevé de polarisation en politique, en particulier aux États-Unis, mais aussi la tendance souvent anti-scientifique rencontrée dans le grand public. Dans les deux cas, nous adoptons souvent des positions très fortes généralement basées sur très peu de choses que nous serions capables d'articuler. Voici une citation souvent attribuée à Winston Churchill :

Le meilleur argument contre la démocratie est une conversation de cinq minutes avec l'électeur moyen.

Que savons-nous réellement des conséquences, des conséquences détaillées, des chaînes de cause à effet, des politiques publiques que nous critiquons si durement ou que nous soutenons si généreusement? Cette question est importante parce que chacun de nous décide, au moins un peu, dans nos démocraties. Et ce problème revient au premier plan dans le débat sur la démocratie représentative et la démocratie directe. Sommes-nous vraiment capables de répondre aux questions posées lors d'un référendum?

Pour montrer que l'illusion de la profondeur explicative permet aux gens d'adopter des positions bien plus tranchées qu'ils ne puissent vraiment argumenter, Sloman et Fernbach ont mis en place des expériences très proches de celles de Rosenblit et Keil. Ils ont demandé aux candidats d'évaluer leur compréhension des questions de politique publique qui étaient des sujets brûlants à l'époque aux États-Unis. Dans la deuxième étape, ils leur ont demandé d'expliquer les conséquences de ces politiques avec le plus de détails possible, et en particulier d'en préciser les chaînes de cause à effet. Puis dans une troisième étape, les candidats ont été invités à réévaluer leur compréhension de chaque politique. Dans les mots des auteurs:

Comme dans la plupart des expériences de ce type, les participants étaient assez mauvais pour générer des explications. À de très rares exceptions près, ils n'avaient tout simplement pas grand-chose à dire lorsque nous leur avons demandé d'expliquer le fonctionnement d'une politique. Ils ne connaissaient pas les mécanismes des politiques d'une manière qu'ils pouvaient articuler. Et conformément à leur incapacité à expliquer, ils ont évalué leur compréhension la deuxième fois plus bas que leur compréhension la première fois. Ils ont manifesté une illusion de profondeur explicative. Leur tentative d'expliquer le problème leur a révélé qu'ils ne le comprenaient pas aussi bien qu'ils le pensaient. Nous concluons que, tout comme

les gens surestiment leur propre compréhension des toilettes et des ouvre-boîtes, ils surestiment également leur compréhension des politiques publiques.

Cependant, ce qu'ils voulaient vraiment savoir dans cette expérience n'était pas si les gens souffraient d'une illusion. Ce qu'ils voulaient savoir, c'était si la tentative d'explication les rendrait moins extrêmes dans leur position sur la question. Ils leur ont donc également demandé d'évaluer leur soutien ou leur opposition sur une échelle de 1 à 7, où 1 indiquait fortement pour et 7 fortement opposé, tandis que 4 signifiait qu'ils n'avaient aucune opinion dans un sens ou dans l'autre. Et encore une fois, ils leur ont demandé à la fois avant et après avoir expliqué quelles seraient les conséquences de la politique. Voici la conclusion des auteurs:

Nous avons constaté que tenter d'expliquer comment une politique fonctionnait non seulement réduisait le sentiment de compréhension de nos participants, mais réduisait également l'extrémité de leur position. Si l'on considère l'ensemble du groupe, le fait que les gens soient en moyenne moins extrêmes signifie que le groupe dans son ensemble est moins polarisé après l'exercice d'explication. La tentative d'explication a fait converger leurs positions.

Mais il y avait une autre expérience dans la série. Sloman et Fernbach encore:

Habituellement, lorsque les gens réfléchissent et parlent de politiques, ils ne sont pas engagés dans une explication causale. La plupart des discours sur la politique portent sur les raisons pour lesquelles nous croyons ce que nous croyons: qui est d'accord avec nous, pourquoi nous tenons à la valeur que la politique aborde, ce que nous en avons entendu aux informations l'autre jour. Notre expérience demandait aux gens de faire quelque chose de difficile et d'inhabituel, en l'occurrence expliquer causalement les effets d'une politique. Cette tâche nécessite d'aborder les détails de la politique et d'expliquer comment la politique interagirait avec un monde compliqué.

Alors que se passerait-il si, au lieu de demander une analyse causale détaillée, on demandait simplement aux candidats de donner les raisons pour lesquelles ils s'opposent à ou soutiennent une politique? Dans les mots des auteurs:

On leur a demandé d'indiquer précisément pourquoi ils pensaient ainsi de la politique. Au lieu de sortir de leurs propres intérêts en pensant à la politique selon ses propres termes, nous leur avons spécifiquement demandé de réfléchir à la politique de leur propre point de vue. De cette manière, nous leur avons demandé de faire ce que les gens font normalement lorsqu'ils réfléchissent à des questions de politique publique. Les participants ont répondu aux mêmes questions que ceux de la première expérience: ils ont évalué leur

compréhension de la question et leur position sur celle-ci avant et après avoir généré des raisons.

Générer des raisons plutôt qu'une explication causale a conduit à un comportement assez différent: les participants n'ont montré aucune diminution de leur sens de la compréhension, et ils n'ont pas non plus modéré leurs positions.

La conclusion de tout cela est simple: si vous voulez réduire la polarisation dans un groupe, ne demandez pas aux gens les raisons pour lesquelles ils adoptent la position qu'ils adoptent. Demandez-leur de produire une explication causale détaillée. La beauté de l'explication causale est qu'elle emmène ceux qui doivent expliquer la chose en dehors de leurs propres systèmes de croyance—pas seulement les croyances religieuses, mais les systèmes de croyance en général. Briser l'illusion de la compréhension réduit la polarisation parce que les gens se rendent compte qu'ils ne comprennent pas vraiment les conséquences d'une politique. Le problème est que, quand on n'est pas qualifié, on ne sait pas ce qu'on ne sait pas, et on ne sait même pas qu'on ne le sait pas. Le problème n'est pas l'ignorance en elle-même—nous sommes tous ignorants. Le problème est que nous oublions trop vite à quel point nous sommes ignorants.

Une autre question soulevée par Sloman et Fernbach est la forte tendance anti-scientifique dans nos démocraties occidentales, non seulement dans certains cercles intellectuels, mais aussi parmi le grand public. Pourquoi persiste-t-elle et même augmente-t-elle malgré toutes nos tentatives d'éducation? Encore une fois, lorsqu'il s'agit de questions importantes et complexes comme l'utilisation d'OGM, de vaccins, d'antibiotiques, d'énergie nucléaire, etc., nous n'en savons pas assez individuellement pour nous faire une opinion nuancée et éclairée sur ces nouvelles technologies et ces développements scientifiques. Et ce n'est pas seulement une question de savoir si nous comprenons la science qui est impliquée dans ces domaines. C'est aussi la question des impacts positifs et négatifs et la complexité de l'évaluation des risques et des bénéfices. En fin de compte, en tant que citoyens ordinaires, nous n'avons d'autre choix que d'adopter les positions de ceux en qui nous avons confiance. Et cela soulève la question: à qui devons-nous faire confiance?

Très souvent, ce sont les gens de notre communauté immédiate. Et comme le soulignent Sloman et Fernbach, nos croyances ne sont pas seulement nos croyances. Ils sont partagés avec notre communauté et cela les rend vraiment difficiles à changer. Les croyances ne sont pas des données isolées que les gens peuvent adopter et rejeter à volonté. Ils sont profondément liés à d'autres croyances, à des valeurs culturelles partagées et à nos identités. Rejeter une croyance implique souvent de rejeter toute une série d'autres croyances, d'abandonner notre communauté, d'aller à l'encontre de ceux en qui nous avons confiance et que nous aimons, et de remettre en question notre propre identité. Dans la confrontation entre culture et cognition, le pouvoir de la culture sur la cognition l'emporte sur toute tentative d'éducation.

Donc, encore une chose à garder à l'esprit lorsque nous nous interrogeons sur la rationalité humaine: l'importance de notre sentiment d'appartenance à un groupe, que ce soit un groupe religieux, une tendance politique ou autre.

Comme nous l'avons dit précédemment, il est parfaitement rationnel de vouloir être bien vu au sein de sa communauté, et en même temps, pour de nombreuses questions sur lesquelles nous aimerions tous nous prononcer aujourd'hui, les positions que nous adoptons sont à peine rationnelles lorsqu'elles sont vues en termes du fonctionnement réel du monde, que ce soit le monde physique, le monde politique ou tout autre aspect de notre existence.

## Conclusion

La position adoptée ici est celle prônée par une *philosophie naturaliste*. L'objectif est de comprendre les humains sans l'ajout d'ingrédients mystérieux. De ce point de vue, la clé est l'idée d'évolution. La seule façon de comprendre qui nous sommes et pourquoi nous fonctionnons comme nous le faisons est de comprendre comment nous en sommes arrivés là. Pour prendre un exemple, on peut parler d'une sorte d'auto-domestication—une coévolution des gènes et des pratiques culturelles—qui a conduit notre espèce à vivre dans des sociétés coopératives.

Concernant la vie animale en général, cela donne l'idée que les esprits des animaux peuvent être composés d'une hiérarchie de modules qui tirent des inférences basées sur des régularités dans le monde, chacun généré il y a fort longtemps par les interactions des ancêtres de l'animal avec leur environnement. La plupart de ces modules fonctionnent inconsciemment et beaucoup sont probablement partagés entre de nombreuses espèces. Ceux-ci opèrent déjà une forme de raisonnement, au sens de traitement de l'information, et la plupart de nos propres modules de ce type là ne nous distingueraient pas par nature des autres animaux.

Le raisonnement qui nous distinguerait des autres animaux implique le langage, la conscience et nos efforts collectifs, principalement axés sur la coopération. Car nous sommes bien les champions du monde de la coopération, capables d'inventer de nouvelles activités en un clin d'œil et d'y travailler ensemble. Mais pour que la coopération se développe à ce niveau, l'évolution a dû (sans réfléchir) inventer la moralité. C'était le but non planifié de l'auto-domestication évoquée plus haut, qui permet aux gènes nous prédisposant à la coopération de proliférer face à la menace des gènes nous prédisposant à profiter des autres sans contrepartie, ceux qui façonnent en nous le resquilleur. Mais même dans le domaine du raisonnement moral, nous sommes encore fortement influencés par des facteurs dont nous ne sommes tout simplement pas conscients, un héritage sans doute du long processus évolutif qui nous a amenés ici.

Alors, sommes-nous devenus rationnels? Individuellement, dans un certain sens. Nos biais cognitifs sont universels et servent des finalités dont nous venons tout juste de prendre conscience, grâce à notre approche évolutive de la compréhension de nous-mêmes. D'autre part, cette rationalité n'est pas celle à laquelle on a tendance à penser, celle qui passe par l'expression logique et soignée des arguments. Ce dernier est né une fois de plus à des fins de coopération. Collectivement, nous sommes certainement capables de ce genre de rationalité, lorsque nous travaillons ensemble pour éliminer autant que possible certains de

ces biais. La science en donne le meilleur exemple, produisant un énorme corpus de connaissances fiables pour la première fois dans la préhistoire et l'histoire humaines. Ce type de raisonnement se produit lorsque nous nous asseyons ensemble pour partager des points de vue opposés, mais avec un objectif commun, pour parvenir à la vérité sur quelque chose ou pour parvenir à un compromis. Dans le cas de la science, ce but commun est bien de connaître et d'expliquer autant que possible la réalité du monde réel que nous supposons exister autour de nous. C'est-à-dire se rapprocher toujours plus de la vérité sur notre monde.

## References

- [1] S. Pinker, *Rationalité*, Les Arènes, Paris (2021)
- [2] R. Baillargeon, Object permanence in 3 1/2 and 4 1/2 month-old infants. *Developmental Psychology* **23** (5), 655 (1987)
- [3] E. Spelke, Initial knowledge: six suggestions, *Cognition* **50**, 431–445 (1994)
- [4] M. Mitchell, *Intelligence artificielle. Triomphes et déceptions*, Dunod, Paris (2021)
- [5] R. Jackendoff, *A User's Guide to Thought and Meaning*, Oxford University Press (2015)
- [6] J.A. Fodor, *The Modularity of Mind: An Essay on Faculty Psychology*, MIT Press, Cambridge, Massachusetts (1983)
- [7] See, for example, L. Cosmides, J. Tooby, From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupré (Ed.), *The Latest on the Best: Essays on Evolution and Optimality* (pp. 276–306), The MIT Press (1987)
- [8] D. Dennett, *From Bacteria to Bach and Back. The Evolution of Minds*, W.W. Norton and Co., New York, London (2017)
- [9] [https://en.wikipedia.org/wiki/Checker\\_shadow\\_illusion](https://en.wikipedia.org/wiki/Checker_shadow_illusion)
- [10] [https://en.wikipedia.org/wiki/Gaze-contingency\\_paradigm](https://en.wikipedia.org/wiki/Gaze-contingency_paradigm)
- [11] S. Sloman, P. Fernbach, *The Knowledge Illusion. The Myth of Individual Thought and the Power of Collective Wisdom*, Pan Books, London (2018)
- [12] A.P. Duchon, W.H. Warren Jr, A visual equalization strategy for locomotor control. Of honeybees, robots, and humans, *Psychological Science* **13** (3), 272–278 (2002)
- [13] M.V. Srinivasan, M. Lehrer, W.H. Kirchner, S.W. Zhang, Range perception through apparent image speed in freely flying honeybees, *Visual Neuroscience* **6** (5), 519–535 (1991)

- [14] D. Kahneman, *Système 1, système 2: Les deux vitesses de la pensée*, Flammarion, Paris (2016)
- [15] J. Haidt, *The Righteous Mind. Why Good People Are Divided by Politics and Religion*, Penguin (2013)
- [16] H. Mercier, D. Sperber, *L'Énigme de la raison*, Odile Jacob, Paris (2021)
- [17] R.B. Zajonc, Attitudinal effects of mere exposure, *Journal of Personality and Social Psychology* **9**, 1–27 (1968)
- [18] S. Schnall, J. Haidt, G.L. Clore, A.H. Jordan, Disgust as embodied moral judgment, *Personality and Social Psychology Bulletin* **34**, 1096–1109 (2008)
- [19] C.B. Zhong, B. Strejcek, N. Sivanathan, A clean self can render harsh moral judgment, *Journal of Experimental Social Psychology* **46**, 859–62 (2010)
- [20] B. Englich, Blind or biased? Justitia's susceptibility to anchoring effects in the courtroom based on given numerical representations, *Law and Policy* **28**, no. 4, 497–514 (October 2006)
- [21] <https://www.verywellmind.com/what-is-a-cognitive-bias-2794963>
- [22] <https://twitter.com/MaxCRoser/status/919921745464905728?s=20>
- [23] L. Rozenblit, F. Keil, The misunderstood limits of folk science: an illusion of explanatory depth, *Cognitive Science* **26** (5), 521–562
- [24] R. Lawson, The science of cycology: failures to understand how everyday objects work, *Memory and Cognition* **34** (8), 1667–1675