

**Stanislas Dehaene**

# Apprendre !

**Les talents du cerveau,  
le défi des machines**



STANISLAS DEHAENE

# Apprendre !

Les talents du cerveau,  
le défi des machines



© ODILE JACOB , SEPTEMBRE 2018  
15, RUE SOUFFLOT , 75005 PARIS

[www.odilejacob.fr](http://www.odilejacob.fr)

ISBN : 978-2-7381-4543-7

Le code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5 et 3 a, d'une part, que les « copies ou reproductions strictement réservées à l'usage du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4). Cette représentation ou reproduction donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

*Ce document numérique a été réalisé par [Nord Compo](#) .*

*Pour Aurore, qui vient de naître,  
et pour toutes celles et tous ceux  
qui ont été bébés un jour.*

« Commencez donc par mieux étudier vos élèves ; car très assurément vous ne les connaissez point. »

Jean-Jacques ROUSSEAU ,  
*Émile ou De l'éducation* (1762).

« Chose étrange et presque stupéfiante, on connaît tous les recoins du corps humain, on a catalogué tous les animaux de la planète, on a décrit et baptisé tous les brins d'herbe, et on a laissé durant des siècles les techniques psychologiques à leur empirisme, comme si elles étaient de moindre importance que celles du guérisseur, de l'éleveur ou du cultivateur. »

Jean PIAGET ,  
*La Pédagogie moderne* (1949).

# INTRODUCTION

---

En septembre 2009, la rencontre d'un enfant hors du commun m'a forcé à réviser mes idées sur l'apprentissage. J'étais en visite à l'hôpital Sarah de Brasilia, un centre de soins à l'architecture blanche inspirée d'Oscar Niemeyer, dédié à la rééducation neurologique et avec qui mon laboratoire collabore depuis une dizaine d'années. Sa directrice, Lucia Braga, me propose de rencontrer l'un de ses patients : Felipe, un gamin de 7 ans, dont la moitié passée à l'hôpital. Elle m'explique que cet enfant, à 4 ans, a reçu une balle perdue (ce qui n'est hélas pas rare au Brésil). Le projectile lui a sectionné la moelle épinière : il est presque totalement paralysé des quatre membres (tétraparétique). La balle a également fauché les aires visuelles du cortex : il est aveugle. Pour l'aider à respirer, une ouverture a été pratiquée dans sa trachée à la base du cou. Et depuis trois ans, il habite une chambre d'hôpital, enfermé dans le cercueil de son corps inerte.

Dans le couloir qui mène à sa chambre, je me prépare mentalement à découvrir un grand handicapé. Et je rencontre... Felipe, un petit garçon comme tous les enfants de 7 ans, au visage plein de vie, bavard, curieux de tout. Il parle à la perfection, avec un vocabulaire riche, et me questionne avec espièglerie sur les mots du français. J'apprends qu'il est passionné par les langues et qu'il ne manque jamais une occasion d'enrichir son vocabulaire trilingue (portugais, anglais et espagnol). Bien qu'il soit aveugle et cloué au lit, il s'évade en écrivant des contes. L'équipe de l'hôpital Sarah l'a encouragé dans cette voie. En

quelques mois, il a appris à dicter ses histoires à un assistant, puis à les écrire à l'aide d'un clavier connecté à un ordinateur et à une carte son. Les pédiatres et les orthophonistes se sont relayés à son chevet pour transformer ses œuvres en vrais livres tactiles illustrés d'images en relief qu'il palpe avec fierté, avec le peu de sensibilité qui lui reste. Ses livres parlent de héros et d'héroïnes, de montagnes et de lacs que jamais il ne verra, mais dont il rêve comme n'importe quel petit garçon.

La rencontre de Felipe m'a bouleversé et m'a persuadé de me pencher sur ce qui est sans doute le plus grand talent de notre cerveau : l'acte d'apprendre. En effet, cet enfant offrait à la fois une belle leçon d'espoir et un défi pour le neuroscientifique. Comment les facultés cognitives résistent-elles à un tel bouleversement de l'environnement ? Comment l'apprentissage parvient-il à converger vers un état stable, le même chez tous les êtres humains, quelles que soient les vicissitudes de la vie ? De nombreux neuroscientifiques sont empiristes : ils estiment, avec John Locke, que le cerveau tire ses connaissances de son environnement. Selon eux, la principale propriété des circuits corticaux est leur plasticité, leur capacité de s'adapter. De fait, les cellules nerveuses ajustent en permanence leurs synapses en fonction des entrées qu'elles reçoivent. Mais dans ce cas, privé d'entrées visuelles et motrices, Felipe aurait dû devenir un être profondément différent. Par quel miracle était-il parvenu à développer des facultés cognitives strictement normales ?

Le cas de Felipe est loin d'être isolé : chacun connaît l'histoire de Helen Keller ou de Marie Heurtin, sourdes et aveugles, mais qui ont toutes deux appris à parler en langue des signes, jusqu'à devenir des adultes engagées dans la société<sup>1</sup>.

Au fil des pages, nous ferons d'autres rencontres qui, je l'espère, chambouleront vos idées reçues sur l'apprentissage. Vous ferez la connaissance d'Emmanuel Giroux, aveugle depuis l'âge de 11 ans, mais mathématicien émérite, pour qui « en géométrie, l'essentiel est invisible avec les yeux, on ne voit bien qu'avec l'esprit ». Comment parvient-il à se promener au sein de la géométrie algébrique, à manipuler des plans, des sphères et des polyèdres, sans jamais les voir ? Nous verrons qu'il utilise les mêmes circuits que d'autres

mathématiciens, mais que son cortex visuel, loin d'être inactif, s'est également recyclé pour faire des mathématiques.

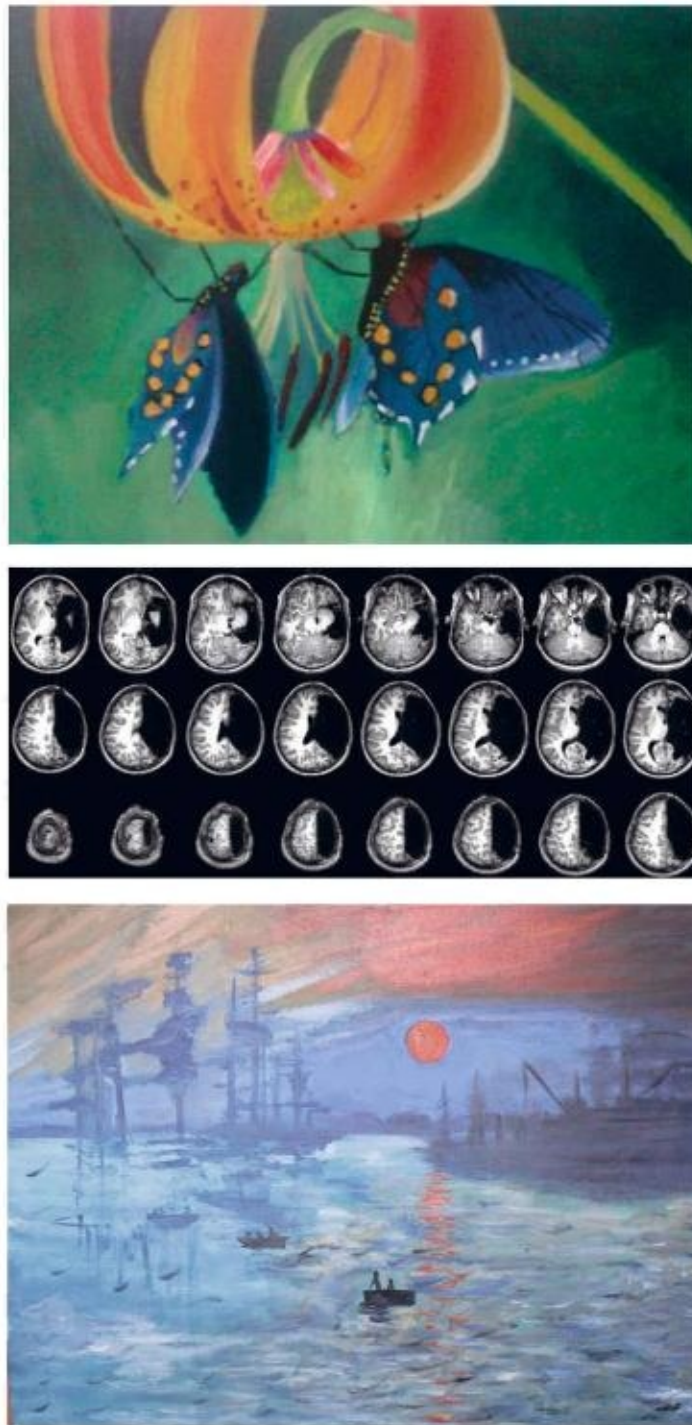
Je vous présenterai également Nico, ce jeune peintre qui, en une visite au musée Marmottan, est parvenu à faire une excellente copie du célèbre tableau de Monet, *Impression soleil levant* (figure 1). Quoi d'exceptionnel ? Rien, si ce n'est qu'il ne possède qu'un seul hémisphère, le gauche – on lui a enlevé la quasi-totalité du cerveau droit à l'âge de 3 ans ! Nico a donc appris à loger au sein d'un seul hémisphère cérébral tous ses talents : parole, écriture et lecture, dessin, peinture, informatique et même escrime, dont il est l'un des champions dans les compétitions internationales en fauteuil roulant. Oubliez tout ce que vous croyez savoir sur les rôles respectifs des deux hémisphères, car la vie de Nico prouve qu'il est tout à fait possible de devenir un artiste sans le secours de l'hémisphère droit : la plasticité cérébrale semble faire des miracles.

Nous visiterons aussi les orphelinats de Bucarest, de sinistre mémoire, où les enfants étaient laissés dès la naissance dans un quasi-abandon – et pourtant, des années plus tard, certains d'entre eux, adoptés avant l'âge de 1 ou 2 ans, ont un parcours scolaire quasi normal.

Tous ces exemples illustrent l'extraordinaire résilience du cerveau humain : même un traumatisme majeur, comme la cécité, la perte d'un hémisphère ou l'isolement social, ne parvient pas à éteindre l'étincelle de l'apprentissage. Langage, lecture, mathématiques, création artistique : tous ces talents singuliers de l'espèce humaine, qu'aucun autre primate ne possède, résistent à la perte d'un hémisphère, de la vue ou de la motricité. Apprendre est un principe vital. Mais nous découvrirons également des contre-exemples dramatiques, où l'apprentissage semble figé. Prenez l'exemple de l'alexie pure, l'incapacité de lire le moindre mot. J'ai personnellement étudié plusieurs adultes, excellents lecteurs, qu'un minuscule accident vasculaire cérébral, restreint à une toute petite région du cerveau, a rendus incapables de déchiffrer même des mots aussi simples que « ton » ou « lac ». Je me souviens d'une femme très brillante, trilingue, fidèle lectrice du *Monde* et qui se désolait qu'à la suite de sa lésion cérébrale, chaque page de son quotidien ressemble désormais à des hiéroglyphes indéchiffrables. Sa motivation à réapprendre à lire était à la mesure du



déracinement qu'elle avait subi. Pourtant, deux années d'efforts ne lui permirent pas de dépasser le niveau de lecture d'un débutant de cours préparatoire, qui ânonne lettre après lettre et bute sur chaque mot. Pourquoi ne pouvait-elle plus apprendre ? Et pourquoi certains enfants, dyslexiques ou dyscalculiques ou dyspraxiques, présentent-ils le même aveuglement dans un domaine particulier de la lecture, du calcul ou du geste ?



**Figure 1.** La plasticité du cerveau parvient parfois à compenser d'impressionnants déficits. Nico, depuis l'âge de 3 ans, ne possède plus qu'un seul hémisphère, le gauche. Cela ne l'a pas empêché de devenir un artiste accompli, capable de peindre aussi bien d'excellentes copies (en bas) que des œuvres originales (en haut).

La plasticité cérébrale semble capricieuse : tantôt elle surmonte des déficits massifs, tantôt elle laisse de côté des enfants et des adultes motivés, intelligents, mais atteints d'un déficit restreint et qui semble permanent. Dépend-elle de circuits particuliers ? Se ferme-t-elle à l'âge adulte ? Peut-elle se rouvrir ? Quelles sont les règles qui la gouvernent ? Comment le cerveau de l'enfant fait-il pour être tellement efficace, dès la naissance et pendant toute la jeunesse ? Quels algorithmes l'évolution a-t-elle implantés dans nos circuits cérébraux afin de leur permettre de se façonner une représentation du monde ? Les comprendre nous permettrait-il d'apprendre mieux et plus vite ? Et pourrait-on s'en inspirer pour construire des machines plus performantes, des intelligences artificielles qui nous imitent, voire nous dépassent ? Telles sont quelques-unes des questions auxquelles ce livre tente de répondre, dans une perspective résolument pluridisciplinaire, en mobilisant les recherches des sciences cognitives et des neurosciences, mais aussi de l'intelligence artificielle et des recherches en éducation.

## *Pourquoi l'apprentissage ?*

Pourquoi devons-nous apprendre ? L'existence même de la faculté d'apprentissage pose question. Ne vaudrait-il pas mieux que nos enfants sachent parler et réfléchir dès le premier jour, telle Athéna dont la légende dit qu'elle sortit toute armée et casquée du crâne de Zeus, en poussant son cri de guerre ? Pourquoi ne naissons-nous pas précâblés, avec un logiciel préprogrammé et doté de toutes les connaissances nécessaires à notre survie ? Dans la lutte pour la vie que décrit Darwin, un animal qui naît mature, avec plus de savoir que les autres, ne devrait-il pas finir par l'emporter ? Pourquoi l'évolution a-t-elle donc inventé l'apprentissage ?

Ma réponse est simple : le précâblage complet du cerveau n'est ni possible ni souhaitable. Impossible, vraiment ? Oui, car si notre ADN devait spécifier tous les détails de nos connaissances, il n'aurait simplement pas la capacité de

stockage nécessaire. Nos 23 paires de chromosomes contiennent 3 milliards de paires de « lettres » A, C, G, T – les molécules adénine, cytosine, guanine et thymine. Combien cela fait-il d'informations ? L'information se mesure en bits : une décision binaire, 0 ou 1. Comme chacune des 4 lettres du génome code 2 bits (on peut les coder comme 00, 01, 10 et 11), cela fait un total de 6 milliards de bits – mais attention, dans les ordinateurs actuels, on compte en octets, qui sont des séquences de 8 bits. Le génome humain se réduit donc à environ 750 mégaoctets – le contenu d'un CD-ROM ou d'une petite clé USB ! Et ce calcul élémentaire ne tient même pas compte des nombreuses redondances qui parsèment notre ADN.

À partir de cette modeste somme d'informations héritées des millions d'années d'évolution, notre génome, initialement confiné à une seule cellule, l'ovule fécondé, parvient à organiser l'ensemble du corps – chaque molécule de chacune des cellules de notre foie, de nos reins, de nos muscles, et bien sûr de notre cerveau : 86 milliards de neurones, 1 000 milliards de milliards de connexions... Comment pourrait-il les définir une par une ? En supposant que chaque connexion ne code qu'un seul bit, ce qui est très certainement une sous-estimation, la capacité de notre cerveau est de l'ordre de 100 téraoctets (environ  $10^{15}$  bits), soit cent mille fois plus que l'information contenue dans notre génome. Paradoxe : le palais fantastique qu'est notre cerveau contient cent mille fois plus de détails que le plan de l'architecte ! Je n'y vois qu'une seule explication : le gros œuvre se construit en suivant les lignes directrices de l'architecte (notre génome), tandis que les détails sont laissés au maître d'œuvre, qui les adapte au terrain (l'environnement). Précâbler un cerveau humain dans tous ses détails serait rigoureusement impossible, c'est pourquoi l'apprentissage doit prolonger l'œuvre des gènes.

Ce simple argument comptable, toutefois, ne suffit pas à expliquer pourquoi l'apprentissage est universellement répandu dans le monde animal. En effet, même des organismes simples et dépourvus de cortex, comme le ver de terre, la mouche drosophile ou le concombre de mer, apprennent bon nombre de leurs comportements. Prenez le petit ver qu'on appelle « nématode », ou *C. elegans*, et qui est rapidement devenu une star de laboratoire. Cet organisme est

incroyablement précâblé : la plupart des individus comprennent exactement 959 cellules, dont 302 neurones, dont toutes les connexions sont connues et reproductibles. Et pourtant, il apprend<sup>2</sup> . Les chercheurs le considéraient initialement comme une sorte d'automate tout juste capable de ramper en avant ou en arrière, mais ils se sont ensuite aperçus qu'il possédait au moins deux formes d'apprentissage : par habitude et par association. L'habitude signifie que l'organisme s'adapte à la présence répétée d'une stimulation (par exemple une molécule dans l'eau) et finit par ne plus y répondre. L'association, quant à elle, consiste à découvrir et à retenir en mémoire quels aspects de l'environnement prédisent les sources de nourriture ou de danger. Le ver nématode s'avère être un champion de l'association, capable de se souvenir que tel goût, telle odeur ou telle température ont été associés par le passé à de la nourriture (des bactéries) ou à des molécules repoussantes (l'odeur d'ail) et d'utiliser cette information pour choisir son chemin dans son environnement.

Avec son petit nombre de neurones, le nématode aurait très bien pu être précâblé. S'il ne l'est pas, c'est parce qu'il est avantageux, pour sa survie, de s'adapter aux conditions spécifiques dans lesquelles il naît. Même des organismes génétiquement identiques ne naissent pas forcément dans le même environnement. Tous ont intérêt à s'adapter rapidement à des conditions fondamentalement imprévisibles. La sélection naturelle, qui est l'algorithme découvert par Darwin, parvient certes à adapter chaque organisme à sa niche écologique, mais elle le fait avec une lenteur affligeante : il faut que des générations meurent, faute d'être adaptées, avant qu'une mutation favorable ne puisse augmenter la survie. La faculté d'apprentissage, elle, agit bien plus vite : elle modifie le comportement en quelques minutes. Et c'est ce qui fait tout l'intérêt de l'apprentissage : s'adapter, le plus vite possible, à des conditions imprévisibles.

C'est pourquoi l'évolution a inventé la faculté d'apprendre. Au fil des générations, elle a découvert qu'il était utile de laisser certains paramètres de l'organisme libres de se modifier pour mieux s'ajuster aux aspects les plus changeants de son environnement. Certains aspects de la physique du monde sont strictement invariables : la gravitation est universelle, la propagation de la

lumière ou des sons dans l'air ne changent pas du jour au lendemain, et c'est pourquoi nous n'avons – Dieu merci ! – pas besoin d'apprendre à faire pousser nos oreilles, nos yeux, ou les labyrinthes de notre système vestibulaire qui mesurent l'accélération de notre corps : toutes ces propriétés de notre corps et de notre cerveau sont codées génétiquement. Par contre, l'espacement de nos yeux, le poids et la longueur de nos membres, la hauteur de notre voix varient, et c'est pourquoi notre cerveau doit les apprendre. Notre pensée est le résultat d'un compromis : énormément d'inné (toutes les grandes catégories intuitives à l'aide desquelles nous subdivisons le monde en images, sons, mouvements, objets, animaux, personnes, causes...), mais encore plus d'acquis qui raffine ces compétences précoces.

Notre espèce a fait de l'apprentissage sa spécialité. Dans notre cerveau, des milliards de paramètres sont libres de s'adapter à notre milieu, notre langue, notre culture, nos parents, notre nourriture... Ces paramètres sont choisis avec soin : au sein de notre cerveau, l'évolution a défini, avec précision, quels circuits sont précâblés et lesquels sont ouverts à l'environnement. Dans notre espèce, la part d'apprentissage est particulièrement vaste, car notre enfance se prolonge pendant de longues années. Par le biais du langage et des mathématiques, nos espaces d'hypothèses se démultiplient en une combinatoire potentiellement infinie – même s'ils s'appuient toujours sur des fondations fixes et invariables, héritées de notre évolution.

## Homo docens

S'il fallait résumer d'un mot le talent particulier de notre espèce, je retiendrais donc le verbe « apprendre ». Plus que des *Homo sapiens* nous sommes des *Homo docens* – car ce que nous savons du monde, pour la plus grande part, ne nous a pas été donné : nous l'avons appris de notre environnement ou de notre entourage. Aucun autre animal n'a su, comme nous, découvrir les secrets du monde naturel. Grâce à l'extraordinaire flexibilité de ses

apprentissages, notre espèce est parvenue à quitter sa savane natale pour traverser déserts, montagnes, océans et, en quelques milliers d'années seulement, conquérir les îles les plus lointaines, les grottes les plus profondes, les banquises les plus glaciales, et jusqu'à la Lune. Depuis la conquête du feu et la fabrication des outils jusqu'à l'invention de l'agriculture, de la navigation ou de la fission atomique, l'histoire de l'humanité n'est que constante réinvention. À la source de tous ces triomphes, un seul secret : l'extraordinaire faculté de notre cerveau à formuler des hypothèses et à les sélectionner pour transformer certaines d'entre elles en connaissances solides sur notre environnement.

Cette remarquable capacité d'apprentissage, l'humanité a découvert qu'elle pouvait encore l'augmenter grâce à une institution : l'école. La pédagogie active est l'apanage de notre espèce : aucun autre animal ne prend le temps d'enseigner de nouveaux talents à ses enfants, activement, en prêtant attention à leurs difficultés et à leurs erreurs. L'invention de l'école, en systématisant l'instruction informelle présente dans toutes les sociétés humaines, a décuplé notre potentiel cérébral. Nous avons compris qu'il fallait profiter de cette exubérante plasticité du cerveau de l'enfant pour lui inculquer un maximum d'informations et de talents. Au fil des siècles, notre système scolaire n'a cessé de progresser en efficacité, commençant toujours plus tôt, dès la maternelle, et se prolongeant pendant une quinzaine d'années, voire plus : un nombre toujours croissant de cerveaux bénéficient d'un enseignement supérieur à l'université, véritable raffinerie neuronale où nos circuits cérébraux acquièrent leurs meilleurs talents.

Aujourd'hui, l'Éducation nationale peut être considérée comme le principal accélérateur de notre cerveau. Sa place de choix, parmi les tout premiers postes de dépenses de l'État, se justifie aisément : sans elle, nos circuits corticaux resteraient des diamants bruts. La complexité de nos sociétés contemporaines ne doit son existence qu'aux multiples améliorations que l'éducation a apportées à notre cortex : lecture, écriture, calcul, algèbre, musique, sens du temps et de l'espace, raffinement de la mémoire... Sait-on, par exemple, que la capacité de mémoire immédiate d'un analphabète, le nombre de syllabes ou de chiffres qu'il

peut répéter, est près de deux fois plus faible que celle d'une personne scolarisée ?

## *Apprendre à apprendre*

L'éducation démultiplie les facultés déjà considérables de notre cerveau – mais pourrait-elle faire mieux encore ? À l'école, à l'université et au travail, contraints de nous adapter toujours plus vite, nous jonglons avec nos algorithmes cérébraux d'apprentissage. Cependant, nous le faisons d'une façon intuitive, sans avoir jamais appris à apprendre. Personne ne nous a expliqué les règles qui font que notre cerveau mémorise et comprend ou, au contraire, oublie et se trompe. C'est dommage, car les données abondent. Un excellent site anglais, l'Education Endowment Fund (EEF), recense les interventions pédagogiques qui marchent<sup>3</sup>. Et l'une des plus efficaces, selon lui, c'est la métacognition, c'est-à-dire le fait de mieux connaître son propre fonctionnement cognitif. Savoir apprendre est l'un des plus importants facteurs de réussite scolaire.

Au cours des trente dernières années, d'importants progrès ont été réalisés dans la compréhension des principes fondamentaux de la plasticité cérébrale et de l'apprentissage. Le fonctionnement de la mémoire, le rôle de l'attention, l'importance du sommeil sont autant de découvertes riches de conséquences pour chacun d'entre nous. Lorsque vous refermerez ce livre, j'espère que vous en saurez beaucoup plus sur vos propres processus d'apprentissage. Il me paraît fondamental que chaque enfant, que chaque adulte prenne la pleine mesure du potentiel de son propre cerveau et aussi, bien sûr, de ses limites. Les sciences cognitives contemporaines, par la dissection systématique qu'elles pratiquent de nos algorithmes mentaux et de leurs mécanismes cérébraux, revisitent le célèbre adage socratique « Connais-toi toi-même ». Aujourd'hui, il ne s'agit plus de pratiquer l'introspection, mais de mieux connaître la subtile mécanique neuronale qui engendre nos pensées, afin de mieux la maîtriser et de la mettre au service de nos goûts et de nos besoins.



Et je pense aussi, bien entendu, aux professionnels de l'apprentissage que sont les enseignants. Je suis profondément convaincu qu'on ne peut pas enseigner convenablement sans posséder, implicitement ou explicitement, un modèle mental de ce qui se passe dans la tête de l'enfant : quelles sont ses intuitions, correctes ou erronées, quelles sont les étapes par lesquelles il doit passer pour progresser, et quels facteurs l'aident à développer ses compétences.

Tous les enfants démarrent dans la vie avec une architecture cérébrale similaire. Il existe donc des principes fondamentaux que tout enseignement, s'il se veut efficace, se doit de respecter. Dans ce livre, nous en donnerons de nombreux exemples. Les compétences des très jeunes enfants pour le langage, l'arithmétique, la logique ou l'estimation des probabilités démontrent l'existence d'intuitions précoces et abstraites sur lesquelles l'enseignement doit s'appuyer. La réalité de pathologies du développement telles que la dyslexie, la dyscalculie, la dyspraxie ou les troubles de l'attention ne fait plus de doute, et des stratégies existent pour les détecter et les compenser. L'un des objectifs de ce livre est de mieux diffuser ces connaissances, afin que chaque enseignant, mais aussi chaque parent, puisse en tirer les conséquences et adapter sa manière d'enseigner.

## *Le défi des machines*

L'intelligence humaine est aujourd'hui confrontée à un nouveau défi : elle n'est plus la seule à savoir apprendre. Dans tous les domaines du savoir, des algorithmes défient notre espèce, car ils apprennent, parfois mieux que nous, à reconnaître les visages ou les voix, à transcrire la parole, à traduire des langues étrangères, à contrôler des machines, et même à jouer aux échecs ou au jeu de go. Les algorithmes de *machine learning* nourrissent aujourd'hui une industrie multimilliardaire qui, de plus en plus, s'inspire de notre cerveau. Comment fonctionnent-ils ? Leurs principes peuvent-ils nous aider à comprendre ce qu'est l'apprentissage ? Parviennent-ils déjà à imiter le fonctionnement de notre cerveau, ou ont-ils encore beaucoup à apprendre ?

Dans ce livre, je consacre une section entière à ces questions. Si les avancées actuelles de l'informatique sont fascinantes, leurs limites sont nettes. Les algorithmes d'apprentissage profond conventionnels ne font qu'imiter une petite partie du fonctionnement de notre cerveau : celle qui correspond aux premières étapes de traitement sensoriel, les fameux 200 ou 300 millisecondes pendant lesquelles notre cerveau opère dans un mode non conscient. Cela ne veut pas dire que ce traitement reste superficiel : en une fraction de seconde, notre cerveau peut non seulement reconnaître un visage ou un mot, mais aussi l'associer à un contexte, le comprendre, l'intégrer à une petite phrase... Cependant, il demeure strictement ascendant, *bottom-up* en anglais, c'est-à-dire sans vraie capacité de réflexion.

Car c'est seulement dans une seconde étape, beaucoup plus lente, consciente et réfléchie, que notre cerveau parvient à déployer toutes ses capacités de raisonnement, d'inférence, de flexibilité, que les machines d'aujourd'hui sont encore loin d'égaliser. Seules certaines architectures informatiques commencent à se construire des modèles abstraits du monde.

Même au sein de leur domaine de prédilection, la reconnaissance rapide des formes, les algorithmes actuels rencontrent un second problème : ils sont beaucoup moins efficaces que notre cerveau. L'état de l'art du *machine learning* consiste à faire tourner un ordinateur sur des millions, voire des milliards d'essais d'entraînement. Le moindre bébé fait beaucoup mieux – il n'a pas besoin de plus d'une ou deux répétitions pour apprendre un mot nouveau. Toute une économie de la donnée échappe encore aux ordinateurs. Pour eux, *machine learning* est synonyme de *big data* : sans données massives, les algorithmes ont bien du mal à extraire des connaissances abstraites et qui généralisent à des situations nouvelles. En bref, ils ne font pas le meilleur usage des données.

Notre cerveau, lui, s'approche de l'optimum : il parvient souvent à extraire la substantifique moelle de la moindre observation. S'ils souhaitent atteindre les mêmes performances, les chercheurs en informatique devront s'inspirer des nombreuses astuces d'apprentissage que l'évolution a intégrées à notre cerveau : l'attention, par exemple, qui nous permet de sélectionner et d'amplifier une information pertinente ; ou bien le sommeil, un algorithme par lequel notre

cerveau fait la synthèse des apprentissages de la journée. Des machines nouvelles, pourvues de ces propriétés, commencent à voir le jour, et leurs performances ne cessent de croître – ce seront elles qui, demain, concurrenceront notre cerveau.

Mais, pour parvenir à une véritable intelligence artificielle générale, c'est probablement le principe même de nos architectures informatiques qu'il faut revoir, parce que le cerveau humain, à la différence des autres espèces animales, semble posséder des algorithmes très spéciaux pour formuler des théories du monde extérieur. Il parvient à explorer systématiquement toutes les combinaisons de toutes les règles possibles, à l'aide d'une sorte de « langage de la pensée », en s'appuyant sur la théorie des probabilités.

Un théorème mathématique le prouve : seule la manipulation des probabilités, c'est-à-dire des incertitudes sur ce que l'on a appris, permet de tirer le maximum de chaque information. Il semble bien que notre cerveau ait découvert cette astuce qui consiste à garder constamment trace de l'incertitude associée à chaque information, et à la mettre à jour lors de chaque apprentissage.

C'est ce que l'on appelle la théorie du cerveau statisticien : en manipulant les probabilités, notre cerveau optimise sa capacité d'apprentissage. De nombreuses données expérimentales viennent à l'appui de cette hypothèse. Même les bébés comprennent les probabilités, et celles-ci semblent profondément inscrites dans nos circuits cérébraux. Chaque enfant semble agir comme un petit scientifique en herbe : à son insu, son cerveau formule des hypothèses, de véritables théories scientifiques que chaque expérience met à l'épreuve. Le raisonnement sur les probabilités, inconscient lui aussi, mais profondément inscrit dans la logique de nos apprentissages, permet de rejeter progressivement les hypothèses fausses et de ne conserver que les théories qui marchent.

Aujourd'hui, des algorithmes nouveaux, qu'on appelle bayésiens, du nom du révérend Thomas Bayes qui en a esquissé la théorie dès le XVIII<sup>e</sup> siècle, commencent à formaliser et à implémenter cette vision nouvelle de l'apprentissage. Je suis prêt à parier qu'ils vont révolutionner le *machine*

*learning* , tant ils s'avèrent déjà capables d'extraire des informations abstraites avec une efficacité proche de celle d'un scientifique humain.

\*

Partons donc, ensemble, explorer ce que nous comprenons aujourd'hui de l'apprentissage. Je vous propose un voyage en trois étapes.

Dans la première partie, intitulée « Qu'est-ce qu'apprendre ? », nous ferons le point sur les théories actuelles de l'apprentissage, à la lumière de leur implémentation concrète dans les ordinateurs. Ce sera l'occasion de formaliser ce qu'apprendre veut dire. En comparant les performances des algorithmes informatiques avec ceux de notre cerveau, *in silico* versus *in vivo* , nous commencerons à entrevoir comment l'apprentissage, pour être optimal, doit s'appuyer sur un usage raisonné des probabilités et des statistiques. Entre le tout inné et le tout acquis, un nouveau modèle émerge : celui du cerveau bayésien, véritable statisticien neuronal. Selon cette théorie, nos gènes, au sein du cerveau en développement, mettent en place de vastes espaces d'hypothèses *a priori* , ainsi que les mécanismes qui permettent de les adapter aux entrées du monde extérieur – et l'environnement *sélectionne* , parmi ces hypothèses, celles qui correspondent le mieux au monde extérieur.

Cette théorie correspond-elle vraiment au fonctionnement de notre cerveau ? Dans une deuxième section, « Comment notre cerveau apprend », nous nous pencherons sur le berceau du bébé humain, une véritable machine à apprendre, souvent imitée mais jamais égalée. Les données récentes démontrent que l'enfant est bien ce statisticien en herbe que prédit la théorie bayésienne. Ses intuitions fulgurantes, dans le domaine du langage, de la géométrie, des nombres ou des statistiques, confirment qu'il n'a rien d'une ardoise vierge, une *tabula rasa* . Dès la naissance, les circuits neuronaux de l'enfant sont bien organisés et projettent des hypothèses sur le monde extérieur. Mais ils possèdent aussi une marge considérable de plasticité, qui se traduit par un bouillonnement permanent de modifications cellulaires. Au sein de cette machine statistique, l'inné et l'acquis, loin de s'opposer, se combinent. Il en résulte un système structuré mais

plastique, aussi bien capable de s'autoréparer en cas de lésion cérébrale que de recycler ses circuits pour apprendre à lire ou à faire des mathématiques.

Dans la troisième partie, « Les quatre piliers de l'apprentissage », je détaillerai quelques-unes des astuces qui font de notre cerveau le plus efficace dispositif d'apprentissage que l'on connaisse aujourd'hui. Quatre mécanismes essentiels modulent massivement notre capacité d'apprendre. En premier vient l'attention : un ensemble de circuits neuronaux qui sélectionnent, amplifient et propagent les signaux auxquels nous accordons de l'importance – et multiplient par cent ou par mille leur représentation en mémoire. En deuxième, l'engagement actif : un organisme passif n'apprend pratiquement rien, car l'acte d'apprendre exige que le cerveau génère activement des hypothèses, avec curiosité. Troisième volet, et complément naturel de l'engagement actif : les signaux d'erreur et de surprise. Ce sont eux qui, en se propageant dans tout le cerveau, viennent corriger nos modèles mentaux, éliminer les hypothèses inappropriées et stabiliser les plus justes. Enfin, quatrième facteur, la consolidation : au fil du temps, notre cerveau compile ce qu'il a acquis et le transfère en mémoire à long terme, afin de libérer les ressources pour d'autres apprentissages. La répétition joue un rôle essentiel dans cette consolidation, et même le sommeil, loin d'être une période d'inaction, constitue un moment privilégié au cours duquel le cerveau se répète et recode les acquis de la journée.

Ces facteurs sont universels : bébé, enfant ou adulte, quel que soit notre âge, ils continuent d'exercer leur pouvoir sur notre capacité d'apprendre. C'est pourquoi nous devons apprendre à les maîtriser. Dans la conclusion, je reviendrai sur les conséquences pratiques de ces avancées scientifiques. Changer nos pratiques, à l'école, en famille ou au bureau, n'est pas forcément aussi compliqué qu'on le pense. Des idées très simples, sur le jeu, le plaisir, la curiosité, la socialisation, la concentration ou encore le sommeil, peuvent augmenter encore ce qui est déjà le plus grand talent de notre cerveau : apprendre.

PREMIÈRE PARTIE

# QU'EST-CE QU'APPRENDRE ?

---

---

« L'intelligence peut être considérée comme la capacité de convertir des informations brutes en connaissances utiles et exploitables. »

Demis HASSABIS ,  
fondateur de la société DeepMind (2017).

Qu'est-ce qu'apprendre ? Ce verbe possède la même racine latine qu'*appréhender* : prendre, attraper, saisir. Apprendre, c'est donc saisir par la pensée : emporter en soi une parcelle de réalité, un modèle de la structure du monde. Comme le dit bien Demis Hassabis, le directeur de la société DeepMind et l'un des chercheurs les plus actifs en intelligence artificielle, apprendre consiste à transformer les informations qui nous parviennent en un jeu de connaissances utiles et exploitables. Grâce à l'apprentissage, les données brutes qui frappent nos sens se muent en idées abstraites, raffinées et suffisamment générales pour que nous puissions les exploiter dans des situations nouvelles. En sciences cognitives, nous les appelons des *modèles internes* .

Dans les pages qui suivent, nous allons passer en revue ce que l'intelligence artificielle et les sciences cognitives nous ont appris de ces modèles internes et de la façon dont ils se transforment avec l'apprentissage. Nous commencerons par examiner les réseaux de neurones conventionnels, ces modèles informatiques qui s'inspirent du cerveau. Comment font-ils pour ajuster leurs paramètres afin de modéliser le monde extérieur ? Nous verrons qu'en dépit de leurs succès, ils ne capturent, pour l'instant, qu'une fraction des capacités du cerveau humain. Le langage, les mathématiques exigent bien plus qu'un réseau de neurones : un véritable langage intérieur, capable de combiner les concepts et de sélectionner parmi ces combinaisons en fonction de leur plausibilité statistique. La vision qui en émergera est celle d'un cerveau statisticien, qui formule des hypothèses

comme un scientifique, et les adopte ou les rejette en fonction des données qu'il reçoit.



# CHAPITRE 1

## Comment un réseau de neurones apprend

---

Apprendre, c'est affiner un modèle du monde. Notre cerveau porte en lui des milliers de modèles du monde extérieur. Métaphoriquement, ce sont comme des maquettes, des modèles réduits plus ou moins fidèles à la réalité qu'ils représentent. Nous avons tous dans la tête, par exemple, un plan de notre quartier, ou une carte mentale de notre maison ou de notre bureau – nous pouvons fermer les yeux et le voir par la pensée.

La richesse de ces représentations mentales, pour la plupart inconscientes, dépasse l'imagination. Notre cerveau d'adulte abrite des milliers de modèles réduits du monde. Nous disposons par exemple d'un vaste modèle mental de la langue française, qui nous permet de comprendre les mots que vous lisez en ce moment même et de deviner que *plastovski* n'est pas un mot du français, tandis que *sextant* l'est et que *blascon* pourrait l'être. Notre cerveau comprend également plusieurs modèles de notre corps : il s'en sert pour coder où se trouvent nos membres, comment les diriger, à quelle vitesse ils se meuvent, comment garder l'équilibre... D'autres modèles mentaux représentent nos

connaissances des objets et de nos interactions avec eux : savoir tenir un stylo, écrire ou faire du vélo. D'autres encore emportent en nous l'esprit des autres : un vaste catalogue mental des personnes qui nous sont proches, de leur aspect, de leur voix, de leurs goûts et de leurs tics.

Ces modèles mentaux sont d'un réalisme impressionnant. Pour vous en convaincre, songez que votre cerveau fait parfois tourner des simulations confondantes de vérité, dans lesquelles vous marchez, bougez, dansez, visitez des lieux nouveaux, entretenez de brillantes conversations, ressentez des émotions fortes... : ce sont vos rêves ! Il est fascinant de prendre conscience que toutes les pensées, parfois fort complexes, qui nous viennent en rêve ne sont que le produit de nos modèles internes qui tournent en boucle. Mais nous rêvons aussi la réalité : notre cerveau projette en permanence, sur le monde extérieur, des hypothèses, des cadres d'interprétation qui donnent du sens au flux de données qui nous parvient des sens. En l'absence d'un modèle interne, ces entrées brutes resteraient impénétrables.

Apprendre, c'est construire, au sein de notre cerveau, un nouveau modèle du monde. L'apprentissage permet à notre cerveau de saisir une parcelle de réalité qui lui échappait auparavant. Réalité extérieure, lorsque nous apprenons l'histoire, la botanique ou le plan d'une ville. Mais également réalité interne à notre organisme, lorsque nous apprenons à coordonner nos gestes et à concentrer nos pensées afin de jouer du violon. Dans tous les cas, notre cerveau *internalise* un aspect nouveau de la réalité : il ajuste ses circuits afin de s'approprier un domaine qu'il ne maîtrisait pas auparavant.

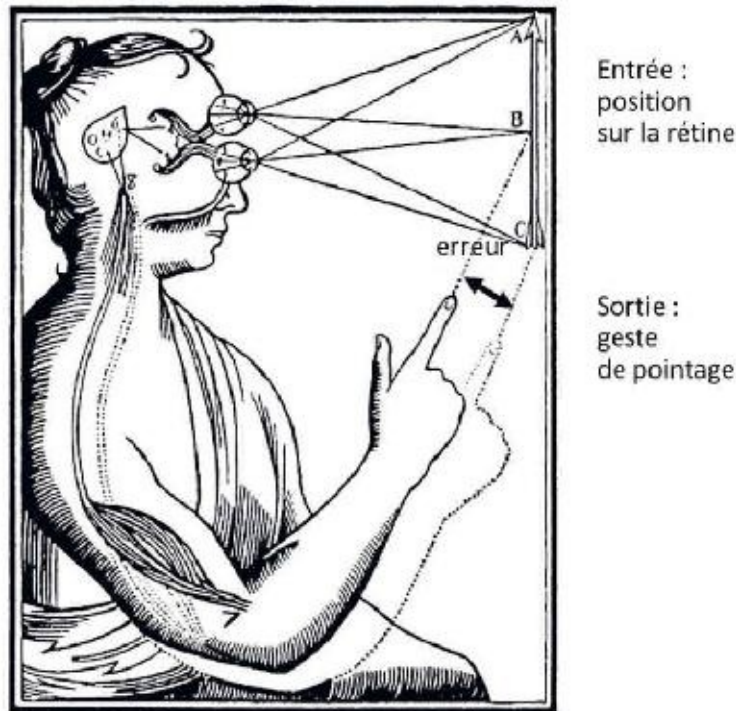
## *Apprendre, c'est ajuster les paramètres d'un modèle mental*

Ajuster un modèle mental, c'est parfois très simple. Comment faisons-nous, par exemple, pour contrôler nos mouvements afin de tendre la main vers un objet

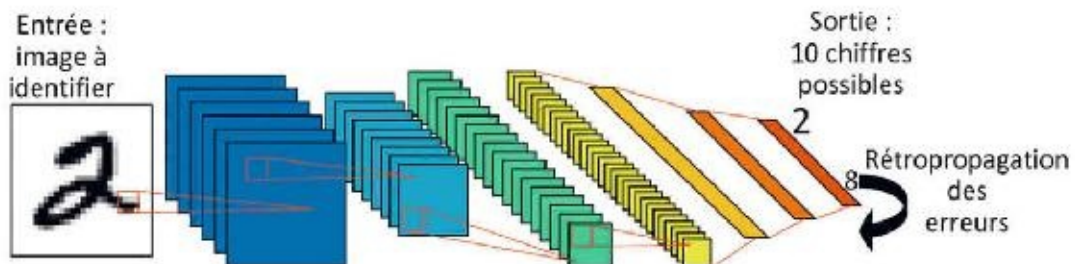
que nous voyons ? René Descartes l'avait déjà pressenti : nous devons apprendre à transformer les entrées visuelles en commandes musculaires ([figure 2](#)). Vous pouvez l'expérimenter sur vous-même en quelques secondes : essayez d'attraper un bâton alors que vous portez les lunettes de quelqu'un d'autre, si possible très myope. Mieux encore, si vous le pouvez, munissez-vous de lunettes à prismes qui décalent la vision d'une dizaine de degrés vers la gauche, et tentez d'attraper un objet – vous pouvez faire cette expérience à l'exposition *C3RV34U* que j'ai organisée à la Cité des sciences de Paris. Vous verrez que votre premier geste est faux : à cause des prismes, votre main atterrit largement à droite du bâton que, pourtant, vous visez. Progressivement, vous ajustez vos mouvements vers la gauche. Par tâtonnements successifs, vos gestes deviennent de plus en plus précis : votre cerveau a appris à corriger le décalage de vos yeux. Enlevez à présent les lunettes, et saisissez le bâton : vous serez surpris de voir votre main se diriger du mauvais côté, bien trop à gauche !

Que s'est-il passé ? Au cours de ce bref apprentissage, votre cerveau a réajusté son modèle interne de la vision. Un paramètre de ce modèle, qui correspond au décalage entre la scène visuelle et l'orientation du corps, a été recalibré. Votre cerveau s'est comporté comme un tireur d'élite qui effectue d'abord un tir d'essai, puis ajuste la hausse de son viseur afin de tirer plus juste. Cet apprentissage est très rapide : quelques essais suffisent à corriger le décalage entre le geste et la vision. Cependant, le nouveau réglage n'est pas compatible avec l'ancien – d'où l'erreur systématique que nous faisons tous lorsque nous ôtons les prismes et que nous revenons à une vision normale.

### Ajustement d'un seul paramètre : le décalage entre la vision et l'action



### Ajustement de millions de paramètres : les connexions qui permettent la vision



**Figure 2.** Apprendre, c'est ajuster les paramètres d'un modèle du monde. Apprendre à viser du doigt, par exemple, consiste à ajuster le décalage entre la vision et l'action : chaque erreur permet de corriger le tir. Dans les réseaux de neurones artificiels, le principe est le même, mais les réglages sont infiniment plus nombreux. Reconnaître quel chiffre est présent dans une image exige d'ajuster des millions de connexions. Là encore, toute erreur – ici, un surcroît d'activation du chiffre 8 – permet de rectifier leur valeur et, ainsi, d'améliorer les performances au prochain essai.

Indéniablement, cet apprentissage est un peu particulier, car il ne requiert que l'ajustement d'un seul paramètre : l'angle de visée. C'est pourquoi il est aussi rapide. La plupart de nos apprentissages sont bien plus élaborés et

demandent d'ajuster plusieurs dizaines, centaines, voire milliers de millions de paramètres – les synapses qui déterminent l'activité de nos circuits. Cependant, le principe est toujours le même : il s'agit de rechercher, parmi une myriade de réglages possibles de notre modèle interne, ceux qui correspondent le mieux à l'état du monde extérieur.

Considérons l'apprentissage d'une langue comme le japonais. Lorsque le cerveau d'un bébé japonais apprend sa langue maternelle, il ajuste progressivement son modèle de la langue à chacune des caractéristiques de la langue japonaise. Essayez d'imaginer une machine dotée de millions de réglages à tous les niveaux. Certains de ces réglages, au niveau de l'entrée auditive, déterminent l'inventaire des consonnes et des voyelles qu'utilise le japonais, et des règles qui permettent de les composer. Le bébé qui naît dans une famille japonaise doit découvrir quelles catégories de sons y sont utilisées en japonais, et où placer les frontières entre les sons. L'un des paramètres, par exemple, concerne la distinction entre les sons *R* et *L* : celle-ci est cruciale en français, mais pas en japonais, qui ne fait pas de nuance entre une élection et une érection... Chaque bébé doit donc fixer un jeu de paramètres qui, collectivement, précisent quelles catégories sont pertinentes pour sa langue maternelle.

À tous les niveaux, il en va de même : apprendre, c'est fixer les paramètres d'un modèle hiérarchique de la réalité. Prenons un exemple de plus haut niveau, celui des règles de grammaire. Une autre différence entre le japonais et le français, que le bébé doit apprendre, concerne l'ordre des mots. Dans une phrase canonique, avec un sujet, un verbe et un complément, la langue française énonce d'abord le sujet, puis le verbe, et enfin son objet : « Jean mange une pomme. » En japonais, par contre, l'ordre le plus courant est sujet, puis objet, puis verbe : « Jean pomme mange. » Ce qui est remarquable, c'est que l'ordre s'inverse également pour les prépositions (qui deviennent, très logiquement, des postpositions), les possessifs et bien d'autres groupes de mots. La phrase : « mon oncle veut habiter à Paris » devient ainsi un sabir digne du Yoda de *La Guerre des étoiles* : « Oncle mon Paris à habiter veut. »

Tous ces chamboulements ne sont pas indépendants les uns des autres. Les linguistes pensent qu'ils relèvent d'un seul et même paramètre qu'on appelle la « position de la tête » : le mot le plus important d'un groupe, sa tête verbale, se place toujours en premier en français (à Paris, mon oncle, veut habiter), mais en dernier en japonais (Paris à , oncle mon , habiter veut ). Ce paramètre binaire distingue d'ailleurs de très nombreuses langues, même si elles n'ont pas de lien historique entre elles (la langue navajo, par exemple, suit les mêmes règles que le japonais). Il suffit donc que l'enfant ajuste le paramètre « position de la tête » dans son modèle interne de la langue afin de s'adapter au français ou au japonais.

## *Apprendre, c'est exploiter le potentiel de la combinatoire*

Est-il vraisemblable que l'apprentissage des langues se réduise ainsi à la sélection de quelques paramètres ? Si cela nous paraît difficile à croire, c'est parce que nous n'imaginons pas le nombre extraordinaire de possibilités qui s'ouvrent lorsque l'on accroît, ne fût-ce qu'un peu, le nombre de paramètres ajustables. Supposez que la grammaire des langues du monde puisse être décrite par une cinquantaine de paramètres binaires, comme le postulent certains linguistes. Cela donne  $2^{50}$  combinaisons, c'est-à-dire plus de 1 million de milliards, ou 1 suivi de 15 zéros ! Les règles syntaxiques des 3 000 langues du monde rentrent aisément dans ce gigantesque espace des langues possibles. Or, dans notre cerveau, ce n'est pas une cinquantaine, mais un nombre mirifique de paramètres qui sont ajustables : 86 milliards de neurones, chacun pourvu d'une dizaine de milliers de contacts synaptiques dont la force peut varier. L'espace de représentations qui s'ouvre est pratiquement infini.

Les langues humaines ne se privent pas d'exploiter cette combinatoire à tous les niveaux. Prenons l'exemple du lexique mental, c'est-à-dire l'ensemble des

mots que nous connaissons et dont nous emportons en nous le modèle. Chacun d'entre nous a appris, dans sa langue maternelle, environ 50 000 mots aux significations les plus diverses. C'est un grand nombre, certes, mais qui reste faible face aux quantités prodigieuses qu'offre la combinatoire. Si l'on considère que ces 50 000 mots font en moyenne deux syllabes, chacune formée d'environ 3 phonèmes, pris parmi les 36 phonèmes du français, le codage binaire de tous ces mots requiert moins de 2 millions de choix binaires élémentaires (des *bits*, dont la valeur est 0 ou 1). Autrement dit, toute notre connaissance du dictionnaire tiendrait dans un petit fichier informatique de 250 kilo-octets (chaque octet comprenant 8 bits).

Encore serait-il possible de comprimer ce lexique mental à une taille bien plus petite, si l'on tenait compte des nombreuses redondances qui régissent les mots. Tirer 6 lettres au hasard, comme « xfdrga », n'engendre pas un mot français. Les vrais mots sont composés d'une pyramide de syllabes qui s'assemblent selon des règles strictes. Et c'est vrai à tous les niveaux : les phrases sont des assemblages réguliers de mots, qui sont des assemblages réguliers de syllabes, qui sont des assemblages réguliers de phonèmes. À chaque niveau, la combinatoire est à la fois vaste (car on choisit parmi plusieurs dizaines ou centaines d'éléments) et bornée (car seules certaines combinaisons sont légales).

Apprendre une langue, c'est découvrir les paramètres qui régissent cette combinatoire à tous les niveaux. Notre cerveau décompose le problème de l'apprentissage en se créant un modèle hiérarchique, multiniveaux, de la structure du langage, depuis le son élémentaire jusqu'à la phrase entière ou même au discours – et ce principe de décomposition hiérarchique se reproduit dans tous les systèmes sensoriels. Certaines aires cérébrales capturent les régularités de bas niveau : elles voient le monde à travers une toute petite fenêtre temporelle et spatiale, et elles en analysent les plus infimes régularités. Ainsi, dans l'aire visuelle primaire, la première région du cortex à recevoir des entrées de la rétine, chaque neurone n'analyse qu'une toute petite portion du champ visuel, un peu comme s'il ne voyait le monde que par un trou d'épingle, et il n'est sensible qu'à des régularités de bas étage, par exemple la présence d'une

ligne oblique qui se déplace. Des millions de neurones font le même travail en différents points de la rétine, et leurs sorties deviennent les entrées du niveau suivant, qui va donc détecter des « régularités de régularités », et ainsi de suite. À chaque niveau, l'échelle augmente : le cerveau recherche les régularités à des échelles de plus en plus vastes, dans le temps comme dans l'espace. Au fil de cette hiérarchie apparaît la capacité de détecter des objets ou des concepts de plus en plus complexes : une ligne, un doigt, une main, un bras, un corps humain, non, deux, ce sont deux personnes qui se font face, c'est une poignée de main... c'est la première rencontre Trump-Macron.

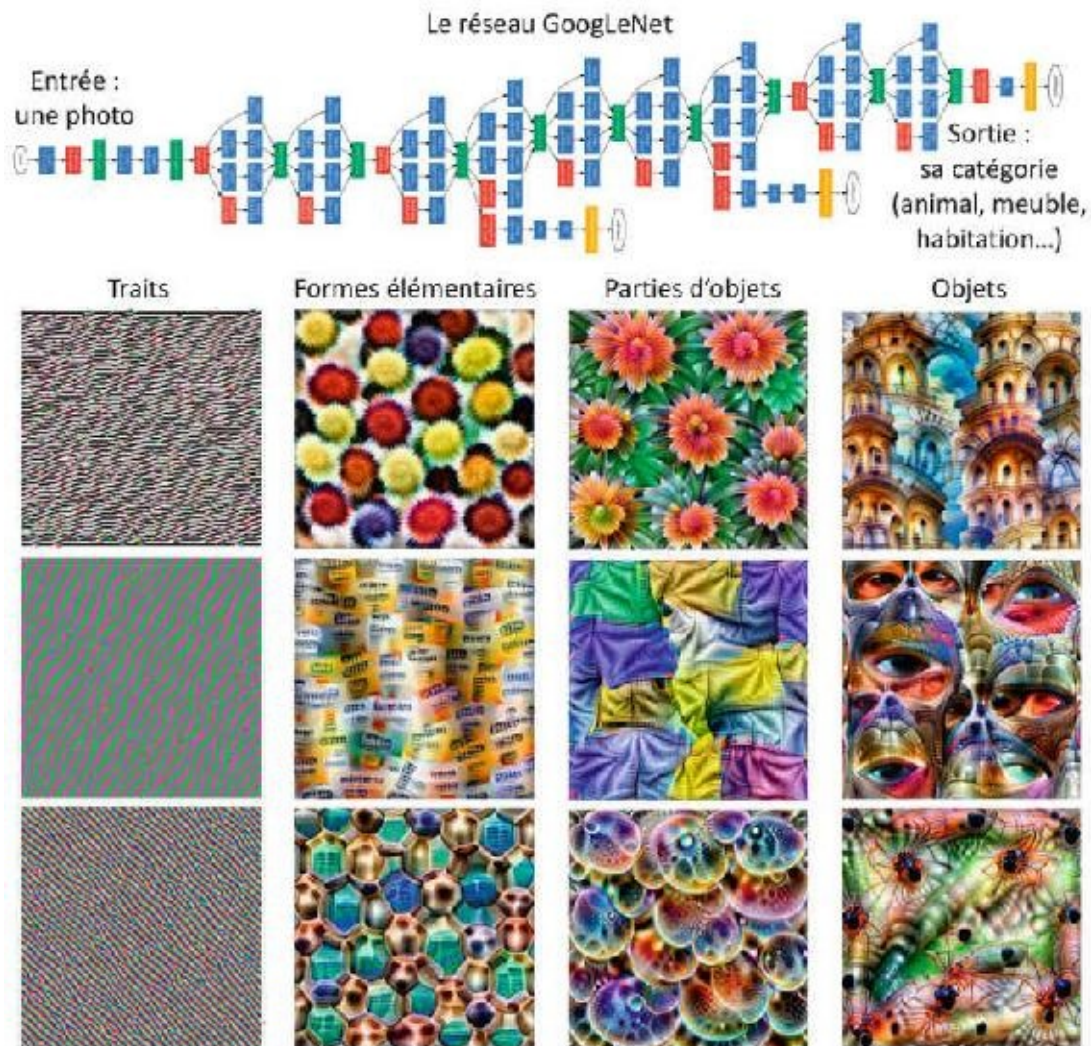
## *Apprendre, c'est minimiser ses erreurs*

Les algorithmes informatiques que nous appelons réseaux de neurones artificiels s'inspirent directement de l'organisation hiérarchique du cortex. On les appelle également réseaux profonds, parce qu'ils sont formés d'une pyramide de couches successives dont chacune analyse les régularités de la couche précédente. Une seule de ces couches ne peut guère découvrir que des aspects extrêmement simples de la réalité extérieure (les mathématiciens parlent de problème linéairement séparable, c'est-à-dire que le système ne parvient à séparer deux catégories A et B qu'en tirant une ligne droite au sein des données). Assemblez plusieurs de ces couches, en revanche, et vous obtenez un dispositif d'apprentissage extrêmement puissant, capable de s'ajuster à des problèmes extrêmement divers. Les réseaux d'aujourd'hui, qui profitent des avancées des puces électroniques, sont profonds au sens où ils contiennent des dizaines de couches successives, de plus en plus éloignées de l'entrée sensorielle, et toujours plus sagaces, capables d'identifier des propriétés de plus en plus abstraites de leurs entrées.

Prenons l'exemple de l'algorithme LeNet, créé par le pionnier français des réseaux de neurones, Yann LeCun ([figure 3](#))<sup>4</sup>. Dès les années 1990, ce réseau de neurones atteint des performances remarquables dans la reconnaissance des



caractères manuscrits. Pendant des années, la poste canadienne l'utilisa pour le traitement automatique des codes postaux. Comment fonctionne-t-il ? L'algorithme reçoit en entrée l'image d'un caractère sous forme de pixels et il en propose, en sortie, une interprétation : l'un des 10 chiffres ou l'une des 26 lettres possibles. Il possède toute une hiérarchie d'unités de traitement qui ressemblent un peu à des neurones et qui forment des couches successives. Les premières couches sont connectées en direct avec l'image : elles appliquent des filtres qui reconnaissent des fragments de lignes et de courbes. Plus on s'élève dans la hiérarchie et plus ces filtres deviennent larges et complexes. Les unités les plus élevées apprennent à reconnaître des portions de plus en plus vastes de l'image : la courbe d'un 2, la clôture d'un O ou les lignes parallèles d'un Z... jusqu'à parvenir, en sortie, à des neurones artificiels qui répondent à un caractère indépendamment de sa position et de son écriture. Toutes ces propriétés ne sont pas imposées par un programmeur : elles découlent entièrement des millions de connexions qui relient les unités entre elles. Ces connexions, une fois ajustées par un algorithme automatisé, définissent le filtre que chaque neurone applique à ses entrées : ce sont elles qui font qu'un neurone répond au chiffre 2 et un autre au chiffre 3.



**Figure 3.** Apprendre, c'est se doter d'une hiérarchie de représentations appropriées au problème posé. Dans le réseau GoogLeNet, qui apprend à identifier des images, des millions de paramètres s'ajustent afin que chaque niveau de la hiérarchie reconnaisse un aspect utile de la réalité. Au plus bas niveau, les neurones artificiels sont sensibles aux traits et aux textures. Plus on monte dans la hiérarchie, plus les neurones répondent à des formes complexes.

Comment s'effectue le réglage de ces millions de connexions ? Exactement comme dans le cas des lunettes à prismes : le réseau constate qu'il a fait une erreur et essaie d'ajuster ses paramètres afin de réduire cette erreur. Chaque réponse erronée fournit une information extrêmement précieuse. Par son signe (comme un geste trop à droite ou trop à gauche), l'erreur nous indique ce qu'il

aurait fallu faire pour réussir. En remontant aux sources de l'erreur, la machine parvient à découvrir comment il aurait fallu régler les paramètres afin d'éviter de se tromper.

Reprenons l'exemple du chasseur qui ajuste le viseur de son fusil. La procédure d'apprentissage est élémentaire. Le chasseur tire et constate qu'il vise 5 centimètres trop à droite. Il dispose à présent d'une information essentielle, à la fois sur l'amplitude (5 centimètres) et sur le signe de l'erreur (trop à droite). Cette information lui permet de corriger le tir. S'il est un peu malin, il sait dans quelle direction effectuer la correction : si la balle a dévié vers la droite, il faut tourner le viseur d'un poil vers la gauche. Même s'il n'est pas aussi futé, il peut faire des essais au petit bonheur et constater que, s'il tourne le viseur vers la droite, l'erreur augmente, alors que s'il le tourne de l'autre côté, elle diminue. Par tâtonnements, le chasseur peut ainsi déterminer comment réduire la taille de l'erreur. Il calcule, implicitement, ce que les mathématiciens appellent la dérivée ou encore le gradient du système. En ajustant son viseur pour diminuer l'erreur, notre brave chasseur, tel Monsieur Jourdain, applique un algorithme d'apprentissage sans le savoir : il effectue ce que l'on appelle une descente de gradient, c'est-à-dire qu'il détermine dans quelle direction il doit ajuster son tir afin de réduire la probabilité de se tromper.

La plupart des réseaux de neurones, en dépit de leurs millions d'entrées, de sorties et de paramètres ajustables, ne fonctionnent pas autrement. Ils observent leurs erreurs, et s'en servent pour ajuster leur état interne dans la direction qui leur paraît le mieux à même de faire diminuer cette erreur. Dans de nombreux cas, tout est connu : d'une part, on dit exactement au réseau quelle réponse il aurait dû activer en sortie (« c'est un 1, pas un 7 ») ; d'autre part, on sait précisément dans quelle direction ajuster les paramètres s'il fait une erreur (un calcul mathématique permet de savoir exactement quelles connexions il faut ajuster lorsque le réseau active un peu trop la sortie « 7 » en réponse à une image du chiffre 1). C'est ce que l'on appelle l'apprentissage supervisé (parce que quelqu'un, qu'on pourrait appeler le superviseur, connaît la bonne réponse que doit donner le système) et par rétropropagation d'erreurs (parce que les erreurs sont renvoyées dans le réseau afin d'en modifier les paramètres). La procédure

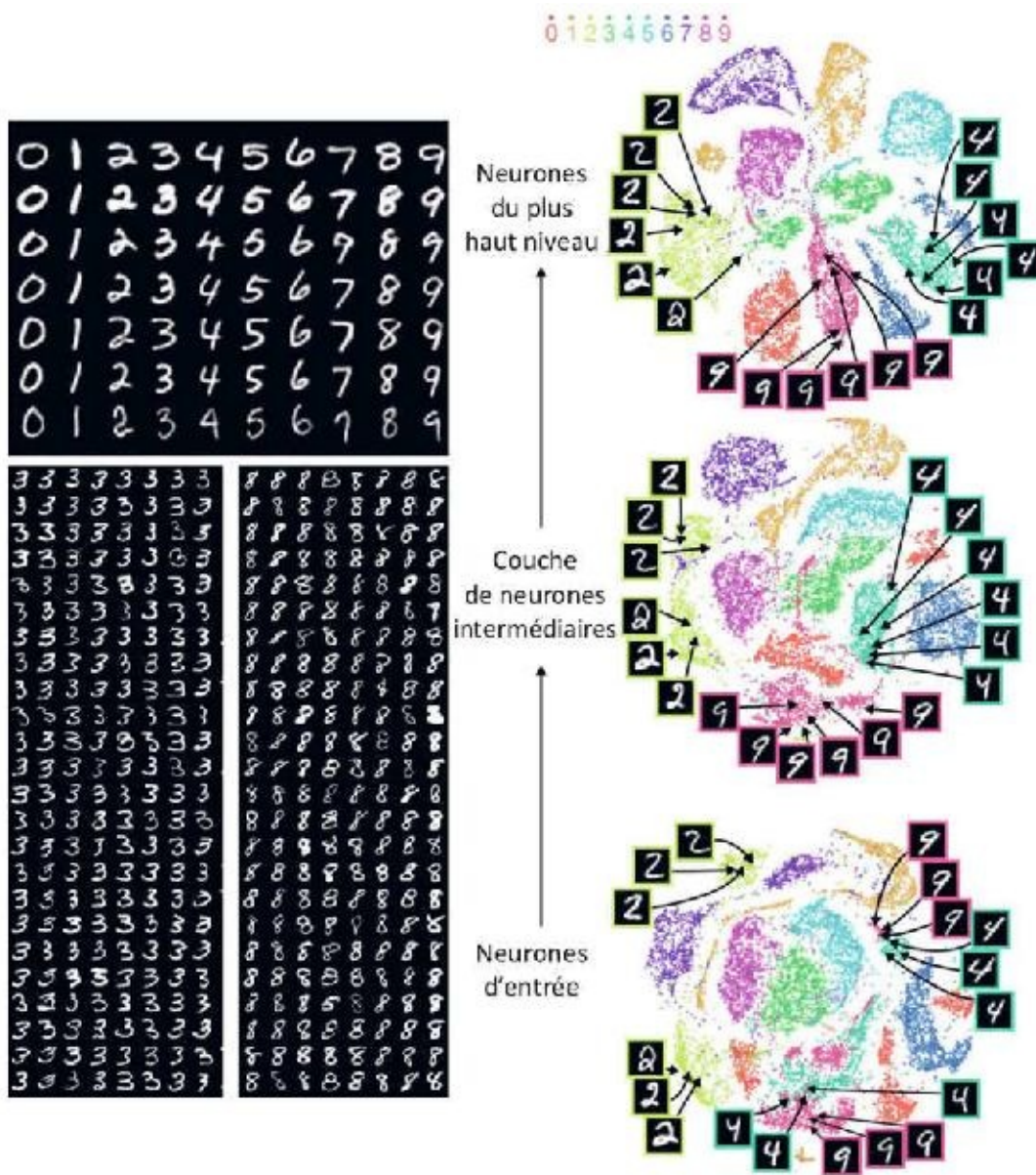
est simple : je tente une réponse, on me dit ce que j'aurais dû répondre, je mesure mon erreur, et je corrige tous mes paramètres afin de la réduire. À chaque étape, je ne fais qu'un tout petit pas, une petite correction dans la bonne direction. C'est pourquoi apprendre une activité complexe, comme jouer à Tetris, exige d'appliquer cette recette plusieurs milliers, millions, voire milliards de fois. Dans un espace qui comprend une multitude de paramètres ajustables, il peut être très long de découvrir le réglage optimal de chaque vis et de chaque boulon.

Les tout premiers réseaux de neurones, dans les années 1980, fonctionnaient déjà sur ce principe de la correction graduelle des erreurs. Les progrès de l'informatique ont permis d'étendre cette idée à des réseaux de neurones gigantesques, qui comprennent des centaines de millions de connexions ajustables. Ces réseaux de neurones profonds sont composés d'une succession d'étapes qui, à chaque pas, s'ajustent au problème posé. À titre d'exemple, la figure 3 montre le système GoogLeNet, dérivé de l'architecture LeNet proposée par Yann LeCun, et qui a gagné l'une des plus importantes compétitions internationales en reconnaissance d'images. Exposé à des milliards d'images, ce système a appris à les séparer en près de 1 000 catégories distinctes : visages, paysages, bateaux, voitures, chiens, insectes, fleurs, panneaux de signalisation, *etc.* Chaque niveau de sa hiérarchie s'est ajusté à un aspect utile de la réalité : les unités de bas niveau répondent sélectivement à des traits ou à des textures, et plus on monte dans la hiérarchie, plus les neurones répondent à des formes complexes : figures géométriques (cercles, courbes, étoiles...), parties d'objets (poche de pantalon, poignée de voiture, paire d'yeux...), voire objets tout entiers (bâtiments, visages, araignées...) <sup>5</sup>.

C'est l'apprentissage profond qui, en tentant de minimiser les erreurs, a découvert que ces formes sont les plus utiles à la catégorisation des images. Mais le même réseau, s'il avait été exposé à des textes ou à des partitions musicales, aurait appris à reconnaître les formes les plus utiles à ce nouvel environnement. La figure 4, par exemple, montre comment un réseau de ce type s'organise lorsqu'on lui demande de se spécialiser pour la reconnaissance de milliers de chiffres manuscrits <sup>6</sup>. Au plus bas niveau, les données sont mélangées : il existe

des formes très semblables, comme un 3 et un 8, qu'il faudrait pourtant distinguer et, inversement, des formes très différentes, comme plusieurs versions du 8 avec la boucle du haut ouverte ou fermée, qu'il faudrait pourtant regrouper. À chaque étape, le réseau de neurones progresse en abstraction, jusqu'à regrouper correctement tous les exemplaires d'un même chiffre. Grâce à la procédure de réduction des erreurs, il a découvert une hiérarchie d'indices qui résout la tâche demandée. Apprendre, c'est découvrir une hiérarchie d'indices appropriés au problème posé.

À l'heure actuelle, le concept d'apprentissage par rétropropagation des erreurs reste au cœur de bon nombre d'applications informatiques. Si votre smartphone reconnaît votre voix, ou si certaines voitures commencent à savoir distinguer les piétons des panneaux de signalisation, c'est grâce à lui. Cependant, le domaine de l'apprentissage artificiel a énormément progressé en trente ans, et les chercheurs ont découvert énormément d'astuces supplémentaires qui facilitent l'apprentissage. En voici quelques-unes en vrac ; nous allons voir qu'elles nous en disent beaucoup sur nous-mêmes et sur nos propres fonctions d'apprentissage.



**Figure 4.** Comment un réseau de neurones profond apprend-il à catégoriser les chiffres manuscrits ? C'est un problème difficile, car le même chiffre peut s'écrire de centaines de manières différentes. Au plus bas niveau de la hiérarchie neuronale (en bas à droite), tous les chiffres sont mélangés, et ceux qui se ressemblent, comme les 9 et les 4, sont confondus. Plus on monte dans la hiérarchie, plus les neurones parviennent à regrouper toutes les images d'un même chiffre et à les séparer par des frontières nettes.

## *Apprendre, c'est explorer l'espace des possibles*

Un des problèmes avec la procédure de correction d'erreurs que je viens de décrire, c'est qu'elle peut rester coincée sur un jeu de paramètres qui n'est pas le meilleur. Imaginez une balle de golf qui roule sur le gazon, toujours selon la ligne de plus grande pente : il se peut qu'elle reste bloquée dans une petite dépression du sol, sans forcément atteindre le point le plus bas de tout le paysage, l'optimum absolu de l'apprentissage. De même, l'algorithme de descente de gradient se retrouve parfois bloqué en un point d'où il ne peut sortir. C'est ce que l'on appelle un minimum local : un puits dans l'espace des paramètres, un piège d'où l'algorithme ne parvient pas à se sortir parce qu'il lui semble impossible de faire mieux. À cet endroit, l'apprentissage se bloque, parce que tous les changements semblent contre-productifs : chacun d'eux augmente le taux d'erreurs. Le système estime avoir été jusqu'au bout de ce qu'il pouvait apprendre. Il ignore qu'il existe en réalité, un peu plus loin dans l'espace des paramètres, d'autres combinaisons bien meilleures. L'algorithme de descente de gradient ne les « voit » pas, car il refuse de remonter la pente pour mieux redescendre de l'autre côté du trou. Myope, il ne regarde qu'à petite distance de son point de départ.

Le problème vous paraît-il abstrait ? Pensez à une situation concrète : vous partez faire vos emplettes au marché, où vous êtes prêt à passer un peu de temps pour trouver les produits les moins chers. Vous vous promenez dans une allée, passez le premier vendeur qui vous paraît hors de prix, évitez le deuxième qui est toujours très cher, et vous arrêtez enfin au troisième stand, nettement meilleur marché que les précédents. Mais qui vous dit que, dans l'allée voisine, ou peut-être même dans le village d'à côté, les prix ne seraient pas encore plus intéressants ? S'arrêter sur le meilleur prix *local* ne garantit pas de trouver le minimum *global*.

Fréquemment confrontés à cette difficulté, les informaticiens ont trouvé une panoplie d'astuces. La plupart consistent à introduire une dose de hasard dans la

recherche des bons paramètres. Au petit bonheur, l'algorithme va parfois chercher plus loin, en sorte que si une meilleure solution est à portée de main, il ait une chance de la trouver. Au lieu de ne chercher que dans une allée du marché, on se balade au hasard. Au lieu de laisser la balle de golf descendre sagement la pente, on la secoue, ce qui diminue ses chances de rester coincée dans un creux. En pratique, on abandonne donc au hasard le choix d'une partie des paramètres. On peut aussi introduire une part de hasard en diversifiant l'ordre des exemples, en ajoutant du bruit aux entrées du système, ou encore en n'utilisant qu'une fraction aléatoire des connexions – toutes ces idées améliorent la robustesse de l'apprentissage.

Certains algorithmes de *machine learning* s'inspirent également de l'évolution des espèces : ils introduisent, au cours de la recherche, des mutations et des croisements aléatoires. Comme en biologie, le taux de ces mutations est savamment contrôlé afin d'explorer des solutions nouvelles sans pour autant perdre trop de temps en tentatives hasardeuses.

Un autre algorithme s'inspire des forges, où les ouvriers ont appris à optimiser les propriétés du métal en le « recuisant ». Quand on veut forger une épée exceptionnellement dure, la méthode du recuit consiste à la réchauffer plusieurs fois, à des températures de moins en moins fortes, afin d'augmenter les chances que les atomes de fer se disposent dans une configuration régulière. Le procédé a été transposé à l'informatique : l'algorithme de recuit simulé s'autorise à introduire des changements aléatoires des paramètres, avec une température virtuelle qui varie progressivement : le niveau de hasard est élevé au départ, mais on le réduit progressivement jusqu'à geler le système dans un réglage optimal.

Toutes ces astuces sont si efficaces que notre cerveau les a adoptées : l'exploration au hasard, la curiosité, la génération aléatoire de décharges neuronales jouent un rôle essentiel dans l'apprentissage chez *Homo sapiens*. Lorsqu'il passe en mode « apprentissage », c'est-à-dire lorsqu'il joue, l'enfant explore des dizaines de possibilités avec une bonne dose de hasard – et pendant la nuit, son cerveau continue de jongler avec les idées jusqu'à trouver la combinaison optimale. Dans la troisième section de ce livre, je reviendrai sur ce



que l'on sait de l'algorithme semi-aléatoire qui gouverne l'extraordinaire curiosité des enfants – et des rares adultes qui ont su garder leur esprit gamin.

## *Apprendre, c'est optimiser une fonction de récompense*

Souvenez-vous du réseau de Yann LeCun, qui reconnaît la forme des chiffres. Pour apprendre, ce type d'intelligence artificielle a besoin qu'on lui fournisse, pour chaque image, la bonne réponse : un chiffre parmi les dix possibles. Le réseau ne parvient à se corriger qu'en calculant la différence entre la réponse correcte et sa tentative à lui. On parle d'*apprentissage supervisé* : quelqu'un, en dehors du système, connaît la solution et tente de l'enseigner au système. C'est efficace, mais force est de constater que les apprentissages de ce type, où l'on connaît à l'avance la bonne réponse, sont plutôt rares. Lorsqu'un enfant apprend à marcher, personne ne lui dit exactement quels muscles contracter – on se contente de l'encourager encore et encore, jusqu'à ce qu'il ne tombe plus. Il apprend uniquement sur la base d'une évaluation du résultat – je suis tombé, ou bien, au contraire, j'ai enfin réussi à traverser la pièce debout !

L'intelligence artificielle est confrontée au même problème d'*apprentissage non supervisé*. Lorsqu'une machine apprend à jouer à un jeu vidéo, la seule chose qu'elle connaît au départ est qu'elle doit essayer de faire le meilleur score. Personne ne peut lui dire à l'avance quelles actions précises effectuer pour y parvenir. Comment peut-elle découvrir seule quelle est la bonne manière de faire ?

C'est en réponse à ce défi que les chercheurs ont inventé l'*apprentissage par récompense*, où on donne au système, non pas tous les détails de ce qu'il doit faire (on ne les connaît pas), mais une « récompense », une évaluation sous forme d'un score quantitatif<sup>7</sup>. Parfois même, le score n'est donné qu'après un très long délai. C'est sur ce principe que la société DeepMind a créé une

machine qui joue au jeu de go. Dans ce cas, comme d'ailleurs aux échecs ou aux dames, ce n'est qu'à la fin de la partie qu'on sait si elle est gagnée ou perdue. Comment, dès lors, remonter dans le temps et ajuster, rétrospectivement, tous les paramètres du jeu ?

L'astuce qu'ont trouvée les informaticiens consiste à apprendre deux choses en même temps : agir et s'autoévaluer. Une moitié du système, qu'on appelle le critique, apprend à prédire le score final. À chaque instant, ce réseau de neurones évalue l'état du jeu et tente de prédire la récompense : suis-je plutôt en train de gagner la partie ou de la perdre ? Mon équilibre est-il stable, ou bien suis-je sur le point de tomber ? Grâce au critique qu'il se forge au fil des essais, le système dispose d'une évaluation de ses actes à chaque instant, et non plus seulement à la fin de la partie. L'autre moitié du réseau, l'acteur, peut alors utiliser cette évaluation pour se corriger : Attention ! Je ferais mieux d'éviter telle ou telle action, parce que le critique pense qu'elle augmente le risque de perdre. Au fil des essais, l'acteur et le critique progressent ensemble : l'un apprend à agir à bon escient, en se focalisant sur les actions les plus efficaces, tandis que l'autre apprend à évaluer, toujours plus finement, les conséquences de ces actes.

La combinaison acteur-critique est l'une des stratégies les plus efficaces de l'intelligence artificielle contemporaine. Adossée à un réseau de neurones hiérarchique, elle fait des merveilles. Dès les années 1980, c'est elle qui a permis à un réseau de neurones de gagner la Coupe du monde de backgammon. Plus récemment, c'est elle qui a permis à la société anglaise DeepMind, filiale de Google, de créer un réseau de neurones multifonctions, capable d'apprendre à jouer à toutes sortes de jeux vidéo tels que Super Mario ou Tetris<sup>8</sup>. On donne simplement au système, en entrée, les pixels de l'image, en sortie, les actions possibles et, comme fonction de récompense, le score du jeu. La machine apprend tout le reste. Lorsqu'elle joue à Tetris, elle découvre que l'écran est composé de formes, dont certaines sont plus importantes que d'autres, que les actions modifient la position de ces formes, que certains gestes font immédiatement perdre la partie, et ainsi de suite – jusqu'à converger vers un joueur artificiel d'une efficacité redoutable. Et lorsqu'elle joue à Super Mario, le changement d'entrées et de récompense lui fait apprendre d'autres aspects :

quels pixels forment l'image de Super Mario, comment il se déplace, où sont les ennemis, les murs, les portes, les pièges et les bonus... et comment agir face à chacune de leurs configurations. En ajustant ses paramètres, c'est-à-dire les millions de connexions qui relient les couches entre elles, le même réseau parvient à s'adapter à toutes sortes de jeux et apprendre à reconnaître indifféremment les formes de Tetris, de Pac-Man, d'Astérix ou d'Alien.

Quel est l'intérêt, pour une société comme Google, d'apprendre à jouer aux jeux vidéo ? Deux ans plus tard, les ingénieurs de DeepMind demandaient au même réseau de neurones de résoudre un problème économique d'intérêt vital : comment optimiser la gestion du parc de serveurs informatiques de Google ? Le réseau de neurones était le même, seules changeaient les entrées (nombre de personnes connectées à Google, jour et heure de la journée, météo, activité sur Twitter, etc.), les sorties (allumer ou éteindre tel ou tel serveur sur divers continents) et la fonction de récompense (consommer moins d'énergie). Résultat : une chute immédiate de la consommation électrique de près de 40 %, faisant économiser des dizaines de millions de dollars – et cela alors même que des myriades d'ingénieurs spécialisés avaient déjà tenté d'optimiser les serveurs. L'intelligence artificielle a vraiment atteint des niveaux de réussite capables de bouleverser toute l'industrie.

La société DeepMind continue d'engranger les succès. Comme chacun le sait sans doute, elle est parvenue à battre à plate couture les spécialistes mondiaux du jeu de go, considéré jusqu'à très récemment comme l'Everest de l'intelligence artificielle<sup>9</sup>. Ce jeu se joue sur un vaste damier (un *go-ban*) de 19 cases de côté, soit  $19 \times 19 = 361$  coups possibles. Le nombre de combinaisons y est tellement vaste qu'il est rigoureusement impossible d'explorer systématiquement tout l'arbre des possibilités qui s'offrent à chaque joueur. Et pourtant, l'apprentissage par récompense a permis au logiciel AlphaGo d'y reconnaître les combinaisons favorables ou défavorables mieux que n'importe quel joueur humain. L'une des nombreuses astuces a consisté à faire jouer le système contre lui-même, exactement comme un joueur d'échecs s'entraîne en jouant à la fois les blancs et les noirs. L'idée est simple : à la fin de chaque

partie, le logiciel gagnant renforce ses actions, tandis que le perdant les affaiblit – mais tous deux ont également appris à mieux évaluer leurs coups.

On se moque volontiers du Baron de Münchhausen qui tentait stupidement de s'envoler en tirant sur les lanières de ses bottes (*bootstraps* en anglais). En intelligence artificielle, pourtant, la folie de Münchhausen a donné naissance à une stratégie plutôt maligne, le *bootstrapping* : petit à petit, partant d'une architecture minimale, dépourvue de connaissances, un réseau de neurones peut devenir champion du monde simplement en jouant contre lui-même.

Cette idée d'augmenter la vitesse de l'apprentissage en faisant collaborer deux réseaux entre eux ou, au contraire, en les opposant continue de susciter des avancées majeures en intelligence artificielle. L'une des astuces les plus récentes, qu'on appelle l'apprentissage contre un adversaire (*adversarial learning*)<sup>10</sup>, consiste à entraîner deux systèmes : l'un qui apprend à faire le tri, par exemple entre les tableaux de Van Gogh et ceux des autres peintres, et l'autre dont le seul objectif est de faire échouer le premier, en concevant de faux tableaux. La récompense du premier, c'est d'identifier les vrais Van Gogh. Celle du second, c'est de faire en sorte que le premier n'y parvienne pas. Résultat, l'algorithme fait émerger non pas une, mais deux intelligences artificielles : une spécialiste mondiale de Van Gogh, férue du moindre détail qui permette d'authentifier un tableau du maître ; et une géniale faussaire, capable de produire des tableaux que même les experts jugent authentiques.

Revenons à notre cerveau. Lui aussi comprend toute une collection d'experts qui se combattent, se coordonnent ou s'évaluent mutuellement. Certaines de nos aires cérébrales apprennent à simuler ce que font les autres : elles nous permettent d'anticiper et d'imaginer le résultat de nos actions, parfois avec un réalisme digne des meilleurs faussaires : notre imagination nous fait voir, avec beaucoup de réalisme, la crique où nous nous sommes baignés l'été dernier, ou la poignée de la porte que nous saisissons dans le noir. D'autres aires apprennent à critiquer les autres : elles évaluent en permanence nos capacités et savent prédire les récompenses ou les punitions que nous risquons d'obtenir. Ce sont elles qui nous poussent à agir ou à nous taire. Nous verrons que la métacognition, cette capacité à se connaître soi-même, à s'autoévaluer, à simuler

mentalement ce qui se passerait si nous agissions de telle ou telle manière, joue un rôle fondamental dans les apprentissages humains. L'opinion que nous nous forgeons de nous-mêmes nous aide à progresser ou, au contraire, nous enferme dans le cercle vicieux de l'échec.

## *Apprendre, c'est restreindre l'espace de recherche*

L'intelligence artificielle contemporaine reste confrontée à un problème majeur. Plus le modèle interne est riche en paramètres, plus il est difficile de trouver comment l'ajuster. Or, dans les réseaux de neurones actuels, l'espace de recherche est immense. On parle d'explosion combinatoire : à chaque étape, des millions de choix sont possibles, et leurs combinaisons sont tellement nombreuses qu'il est impossible de les explorer toutes. Résultat, l'apprentissage est parfois excessivement lent : il faut des milliards d'essais pour orienter le système dans la bonne direction au sein de ce paysage immense.

Le nombre immense de paramètres que possèdent les réseaux de neurones entraîne souvent une seconde difficulté, que les Anglo-Saxons appellent *overfitting* ou « surapprentissage ». Le système dispose de tant de degrés de libertés qu'il lui est plus facile d'apprendre par cœur tous les détails de chaque exemple que d'identifier une propriété plus générale qui permette de les expliquer.

Comme l'a dit John von Neumann, le père de l'informatique : « Donnez-moi quatre paramètres, et je reproduis la forme d'un éléphant – et avec un cinquième, je lui fais remuer la trompe ! » Disposer de trop de paramètres nuit à l'abstraction. Le système apprend certes facilement, mais il est incapable de généraliser à des situations nouvelles, auxquelles son éducation ne l'a pas exposé. Or cette faculté de généralisation est la clé de l'apprentissage. Quel serait l'intérêt d'une machine qui ne gagnerait que les parties de go auxquelles

elle a déjà joué ? Il faut qu'elle atteigne le meilleur niveau face à n'importe quel joueur, que celui-ci joue des coups connus ou nouveaux.

Si l'on veut à la fois accélérer l'apprentissage et améliorer la généralisation, l'une des interventions les plus efficaces consiste à simplifier le modèle : on réduit au minimum le nombre de paramètres à ajuster, ce qui force le système à trouver des solutions générales. C'est sur ce principe que Yann LeCun a inventé les *réseaux à convolution*, ce qu'il y a de meilleur aujourd'hui dans le domaine de la reconnaissance d'images<sup>11</sup>. L'idée est simple : pour reconnaître une image, il faut faire à peu près le même travail partout. Sur une photo, par exemple, les visages sont susceptibles d'apparaître n'importe où. Pour les reconnaître, il suffit donc d'appliquer en chaque point le même algorithme (rechercher un ovale, une paire d'yeux, etc.). Ainsi, il n'est pas nécessaire d'apprendre un modèle différent en chaque point de la rétine : ce que l'on apprend en un endroit peut être réutilisé partout ailleurs.

Au cours de l'apprentissage, les réseaux à convolution de Yann LeCun recopient donc, dans l'ensemble du réseau, les apprentissages qu'ils effectuent en un endroit donné – et ils effectuent cette opération à tous les niveaux, à des échelles de plus en plus larges. Ils possèdent donc un nombre restreint de paramètres à apprendre : un seul filtre qui s'applique partout, plutôt qu'une pléthore de connexions différentes. Cette simple astuce améliore massivement les performances, et tout particulièrement la généralisation à des images nouvelles – tout simplement parce que l'algorithme qui s'applique à une image nouvelle bénéficie d'une expérience immense, celle de tous les points de toutes les photos jamais vues auparavant. Elle accélère également l'apprentissage, parce que la machine n'explore qu'un sous-ensemble des modèles de la vision. On peut dire qu'avant même d'apprendre, elle sait déjà quelque chose d'important sur le monde : le fait que le même objet peut apparaître n'importe où dans l'image.

Cette astuce se généralise à pléthore de situations similaires. Reconnaître la parole, par exemple, c'est faire abstraction de la voix du locuteur. On y parvient en forçant un réseau de neurones à utiliser les mêmes connexions dans diverses bandes de fréquences, que la voix soit grave ou aiguë. Réduire le nombre de

paramètres à apprendre accélère l'apprentissage et conduit à une meilleure généralisation à des voix nouvelles : le bénéfice est double, et c'est ainsi que votre smartphone répond à votre voix.

## *Apprendre, c'est projeter des hypothèses a priori*

La stratégie de Yann LeCun fournit un bon exemple d'une notion bien plus générale : l'exploitation des connaissances innées. Si les réseaux à convolution apprennent mieux et plus vite que d'autres types de réseaux de neurones, c'est parce qu'ils n'apprennent pas tout. Ils incorporent, dans leur architecture même, une hypothèse forte : ce que j'apprends en un endroit peut se généraliser partout ailleurs. Le problème principal de la reconnaissance d'image, c'est l'invariance : je dois reconnaître un objet, quelles que soient sa position et sa taille, même s'il bouge de droite ou de gauche ou qu'il s'éloigne. C'est un défi, mais c'est aussi une contrainte très forte : je m'attends à ce que les mêmes indices permettent de reconnaître un visage en tout point de l'espace. En répliquant le même algorithme partout, les réseaux à convolution exploitent avec efficacité cette contrainte : ils l'intègrent dans leur structure même, de façon innée, préalable à tout apprentissage. Le système n'apprend pas l'invariance : il la suppose vraie *a priori* et l'utilise pour réduire l'espace d'apprentissage – astucieux !

La morale de cette histoire, c'est que l'apprentissage ne peut pas tout – il s'appuie toujours, d'une manière ou d'une autre, sur des *a priori* . Plutôt que d'essayer de tout apprendre de zéro, il est bien plus efficace de s'appuyer sur des hypothèses par défaut, qui délimitent clairement le domaine de recherche et intègrent, dans leur architecture même, les lois qui le gouvernent. Plus il y a d'hypothèses innées, plus l'apprentissage est rapide (à condition, bien sûr, que ces hypothèses soient justes !). Il serait faux, par exemple, de croire que le logiciel AlphaGo Zéro, qui est devenu champion de go en jouant contre lui-

même, est parti de rien : sa représentation initiale inclut, entre autres, la connaissance des symétries du jeu, ce qui divise par huit le nombre d'états à étudier.

Notre cerveau, lui aussi, est pétri d'*a priori* de toutes sortes. Nous verrons dans un instant que, dès la naissance, le cerveau des bébés est très organisé et connaît déjà bien des choses sur le monde extérieur. Son cerveau sait déjà, de façon implicite, que le monde est fait d'objets solides qui se déplacent seulement lorsqu'on les pousse, sans jamais s'interpénétrer – et aussi d'autres entités plus étranges, qui se déplacent d'elles-mêmes en fonction d'intentions et de croyances, et qui parlent (les personnes). Pas besoin d'apprendre ces lois : puisqu'elles sont vraies partout où vivent les humains, notre génome les a inscrites en dur, afin de réduire l'apprentissage à la portion congrue : ne restent à préciser que les paramètres spécifiques qui varient à chaque génération, comme la couleur des yeux, la tonalité de la voix ou les goûts singuliers des personnes qui nous entourent.

Bien sûr, si notre cerveau connaît d'emblée des choses aussi abstraites que la différence entre les personnes et les objets, c'est qu'il les a apprises – non pas pendant les quelques semaines de sa courte vie, mais au cours des millions d'années de son évolution. La sélection darwinienne peut être considérée comme un algorithme d'apprentissage – un programme incroyablement puissant, parce qu'il tourne depuis des centaines de millions d'années, en parallèle sur des milliards de machines (chaque être vivant)<sup>12</sup>. Nous sommes les héritiers d'une infinie sagesse : notre génome a internalisé le savoir des générations ancestrales qui nous ont précédés. Ces connaissances innées sont d'une nature différente des faits que nous apprenons au cours de notre vie : elles sont bien plus abstraites, car elles confèrent à nos réseaux de neurones des biais qui facilitent l'apprentissage des lois de la nature.

En résumé, dans notre cerveau, les gènes spécifient, au cours de la grossesse, une architecture générale qui, en restreignant l'espace de recherche, facilite et accélère les apprentissages ultérieurs. Les informaticiens parlent d'*hyperparamètres* – ce sont les grandes variables de haut niveau qui spécifient le nombre de couches, le type de neurones, la forme générale de leurs



interconnexions, le fait qu'elles soient dupliquées en tout point de la rétine, et ainsi de suite. Dans notre cerveau, ces variables font partie de notre génome, nous n'avons plus besoin de les apprendre – notre espèce les a internalisées au cours de son évolution.

Notre cerveau n'est donc pas simplement soumis, de façon passive, aux entrées sensorielles. Au contraire, il possède d'emblée un jeu d'hypothèses abstraites héritées de son évolution, qu'il projette sur le monde extérieur. La philosophie empiriste qui sous-tend bon nombre de réseaux de neurones actuels, selon laquelle nous naissons avec des circuits complètement désorganisés, vierges de toute connaissance, et qui reçoivent l'empreinte de l'environnement, n'est pas la bonne. Apprendre, c'est partir d'un jeu d'hypothèses *a priori* (même s'il peut être très vaste), projeter ces hypothèses sur les données et sélectionner celles qui conviennent le mieux. Comme le dit Jean-Pierre Changeux dans *L'Homme neuronal* : « Apprendre, c'est éliminer. »

## CHAPITRE 2

# Pourquoi notre cerveau apprend mieux que les machines

---

Les succès récents de l'intelligence artificielle peuvent laisser croire que nous avons enfin compris comment imiter l'apprentissage et l'intelligence de l'espèce humaine dans des machines – au point que, selon certains prophètes autoproclamés, les machines seraient sur le point de nous dépasser. Rien n'est plus faux. En fait, la plupart des chercheurs en sciences cognitives, même s'ils admirent les progrès récents des réseaux de neurones artificiels, savent très bien que ces machines demeurent limitées. Dans un article récent, j'ai argumenté que les réseaux de neurones conventionnels correspondent étroitement aux opérations que notre cerveau réalise inconsciemment, en deux dixièmes de seconde, lorsqu'il perçoit une image : il la reconnaît, la catégorise et accède à son sens<sup>13</sup>. Cependant, notre cerveau, lui, va beaucoup plus loin : il est capable de l'explorer consciemment, avec attention, point par point, pendant plusieurs secondes. Il formule des représentations symboliques, des théories explicites du monde que nous pouvons partager avec d'autres personnes par le biais du langage.

Les opérations de cette nature, lentes, raisonnées, symboliques, demeurent pour l'instant l'apanage de notre espèce. Les réseaux de neurones actuels les modélisent mal, même si, chaque année, des progrès sont faits dans la compréhension du langage, la traduction automatique ou le raisonnement logique. C'est un reproche que l'on fait souvent aux réseaux de neurones artificiels : ils essaient de tout apprendre au même niveau, comme si chaque problème revenait à une question de classification automatique. Pour celui qui ne possède qu'un marteau, tout ressemble à un clou ! Notre cerveau, lui, est bien plus flexible. Il parvient très vite à hiérarchiser les informations et, lorsque c'est possible, à en extraire des principes généraux, logiques, explicites.

## *Ce qui manque encore à l'intelligence artificielle*

Il est intéressant d'essayer de préciser ce qui manque encore à l'intelligence artificielle, car c'est une manière de définir, le plus précisément possible, ce qu'il y a d'unique dans notre propre capacité d'apprentissage. Voici une petite liste, sans doute incomplète, de fonctions que même un très jeune enfant possède, et qui font échouer lamentablement la plupart des réseaux actuels :

- *L'apprentissage de concepts abstraits.* La plupart des réseaux de neurones contemporains ne modélisent convenablement que la toute première passe de traitement de l'information pendant laquelle, en moins d'un cinquième de seconde, les aires visuelles analysent une image. Ces algorithmes connexionnistes sont loin d'être aussi profonds qu'on le dit. Selon l'un de leurs inventeurs, Joshua Bengio, en réalité, « les réseaux dits "profonds" ont tendance à apprendre des régularités statistiques superficielles dans les données, plutôt que des concepts abstraits de haut niveau<sup>14</sup> . » Pour reconnaître la présence d'un objet, ils s'appuient sur des éléments anecdotiques de l'image, tels que la couleur ou la forme. Changez ces

détails, et leurs performances s’effondrent : ils sont absolument incapables de reconnaître ce qui fait l’essence d’un objet et de concevoir qu’une chaise reste une chaise même si elle est faite de verre, d’un seul pied de métal plié ou de plastique gonflable. Cette propension à ne faire attention qu’à la surface des choses les rend susceptibles d’erreurs massives. Il existe toute une littérature sur la manière de tromper un réseau de neurones : prenez une banane, modifiez-lui quelques pixels ou collez-lui un autocollant bien particulier, et le réseau de neurones la prendra pour un grille-pain !

Il est vrai que, si l’on flashe une image à une personne pendant une fraction de seconde, elle fait parfois le même genre d’erreur que la machine, et peut confondre par exemple un chien et un chat<sup>15</sup>. Cependant, dès qu’on lui laisse un peu plus de temps, le cerveau humain ne s’y trompe pas. Contrairement à la machine, il s’interroge, réanalyse, porte son attention sur tel ou tel aspect de l’image qui ne correspond pas à sa première impression. Cette seconde analyse, consciente, intelligente, fait appel à nos capacités de raisonnement et d’abstraction. Les réseaux de neurones négligent un point essentiel : apprendre, c’est se former un modèle abstrait du monde, pas juste un filtre de reconnaissance de formes. En apprenant à lire, par exemple, nous avons acquis un concept abstrait de chaque lettre de l’alphabet, qui nous permet de la tracer aussi bien que de la reconnaître sous tous ses déguisements :



Douglas Hofstadter, informaticien et cogniticien, a dit un jour que le vrai challenge pour l’intelligence artificielle consistait à reconnaître la lettre A... Boutade, certes, mais boutade profonde. L’intelligence abstraite que les humains déploient même dans cette situation triviale est à l’origine d’un amusant objet de

la vie quotidienne : le CAPTCHA, cette petite chaîne de lettres que certains sites Internet vous demandent de reconnaître pour prouver que vous êtes un être humain, et non une machine.

Pendant des années, les CAPTCHA ont résisté aux machines. Mais l'informatique évolue vite : en 2017, un système artificiel est parvenu à reconnaître les CAPTCHA à peu près aussi bien qu'un humain<sup>16</sup>. Sans surprise, cet algorithme imite bon nombre d'aspects de notre cerveau. Véritable tour de force, il extrait le graphe de chaque lettre, l'essence d'un A, et il utilise toutes les ressources du raisonnement statistique pour vérifier à quel point cette idée abstraite s'applique à l'image actuelle. Cependant, cet algorithme informatique sophistiqué ne s'applique qu'aux CAPTCHA. Notre cerveau, lui, applique cette faculté d'abstraction à tous les aspects de notre vie quotidienne.

- *La vitesse d'apprentissage.* Tout le monde s'accorde à dire que les réseaux de neurones actuels apprennent bien trop lentement : il leur faut des milliers, des millions, voire des milliards de données pour acquérir l'intuition d'un domaine. Cette lenteur, nous en avons des preuves expérimentales. Il faut pas moins de 900 heures de jeu pour que le réseau de neurones conçu par DeepMind atteigne un niveau raisonnable sur une console Atari – alors qu'un être humain atteint le même niveau en 2 heures<sup>17</sup> !

Autre exemple : l'apprentissage du langage. Le psycholinguiste Emmanuel Dupoux estime que, dans la plupart des familles françaises, un enfant entend environ 500 à 1 000 heures de parole par an, ce qui lui suffit à apprendre sa langue maternelle. Encore s'agit-il certainement d'une surestimation, car chez les Chimane, une population indigène de l'Amazonie bolivienne, les enfants n'entendent que 60 heures de parole par an, ce qui ne les empêche pas de devenir d'excellents locuteurs de la langue chimane. En comparaison, les meilleurs systèmes informatiques actuels d'Apple, de Baidu ou de Google nécessitent entre vingt et mille fois plus de données. Dans le domaine de l'apprentissage, l'efficacité du cerveau humain reste inégalée : « *Machines are data hungry, but*

*humans are data efficient.* » L'apprentissage, dans notre espèce, sait tirer le meilleur parti de la moindre donnée.

- *L'apprentissage social.* Notre espèce est la seule à pratiquer le partage d'informations : nous apprenons énormément d'informations des autres êtres humains, par imitation ou par le biais du langage. Cette capacité est, pour l'instant, hors de portée des réseaux de neurones actuels. Chez eux, la connaissance est cryptée, diluée dans les valeurs de centaines de millions de poids synaptiques, où elle demeure implicite. Il est impossible de l'extraire pour la communiquer à d'autres. L'extraordinaire efficacité avec laquelle nous parvenons, en quelques mots, à partager notre savoir avec d'autres (« Pour la boulangerie, prenez à droite dans la petite rue derrière l'église ») reste inégalée dans le monde animal comme en informatique.
- *L'apprentissage en un essai.* Version extrême de cette efficacité : il nous arrive d'apprendre en un seul essai. Si j'utilise un nouveau verbe, disons « daxer », ne fût-ce qu'une seule fois, cela vous suffit à le connaître. Entendons-nous bien : certains réseaux de neurones aussi sont capables de stocker un épisode spécifique. Mais ce que les machines ne font pas encore bien, et que le cerveau humain réussit à merveille, c'est d'intégrer cette information au sein d'un réseau de connaissances. Instantanément, vous parvenez non seulement à mémoriser le verbe « daxer », mais également à le conjuguer et à l'insérer dans d'autres phrases (Veux-tu daxer avec moi ? Daxeriez-vous si vous le pouviez ? La daxation est un problème ; etc.). Lorsque je vous dis « viens daxer », vous n'apprenez pas juste un mot, mais vous l'insérez dans un vaste système de symboles et de règles : c'est un verbe à l'infinitif, du premier groupe, qui se conjugue (je daxe, tu daxes, etc.). Apprendre, c'est parvenir à insérer des connaissances nouvelles au sein d'un réseau existant.
- *L'apprentissage de règles systématiques, formulées dans un langage de la pensée.* Les règles de grammaire ne sont qu'un exemple d'un talent particulier de notre cerveau – celui de découvrir les règles générales qui se cachent derrière des cas particuliers. En mathématiques, en sciences, en

musique, nous parvenons à extraire des principes très généraux, d'authentiques règles abstraites. Prenez l'arithmétique : notre capacité d'additionner deux nombres est extrêmement générale – une fois que nous avons appris la procédure avec de petits nombres, nous pouvons la systématiser à des nombres arbitrairement grands. Mieux, nous en tirons des inférences d'une extraordinaire généralité. Bon nombre d'enfants, vers 5-6 ans, découvrent que chaque nombre  $n$  possède un successeur  $n + 1$ , et que la suite des nombres entiers est donc infinie – il n'y a pas de plus grand nombre. Je me souviens encore avec émotion du moment où j'ai pris conscience de ce qui était, en réalité, mon premier théorème mathématique. Quel extraordinaire pouvoir d'abstraction ! Comment notre cerveau, avec un nombre fini de neurones, parvient-il à conceptualiser l'infini ?

Aucun réseau de neurones ne sait aujourd'hui représenter des connaissances aussi systématiques que « chaque nombre possède un successeur ». Les valeurs de vérité absolues, vrai ou faux, ne sont pas leur tasse de thé. La systématité<sup>18</sup>, cette capacité de généraliser sur la base d'une règle symbolique plutôt que d'une ressemblance superficielle, échappe encore aux modèles actuels : l'apprentissage que l'on dit profond manque encore d'une compréhension profonde.

Notre cerveau, lui, semble pouvoir se créer des formules dans une sorte de langage mental. S'il parvient à exprimer le concept d'ensemble infini, c'est parce qu'il possède un langage interne doué d'une fonction de négation (infini = non fini). Le philosophe américain Jerry Fodor, dans ses livres et ses articles, a théorisé cette faculté : il postule que notre pensée consiste en symboles qui se combinent selon les règles d'un « langage de la pensée<sup>19</sup> ». Sa particularité est d'être récursif : chaque objet nouvellement créé (disons, le concept d'infini) peut immédiatement être réutilisé dans des combinaisons nouvelles, sans limites. Combien existe-t-il d'infinis ? Telle est la question saugrenue que se posa le mathématicien Georg Cantor, et qui le conduisit à formuler l'une des théories mathématiques les plus innovantes du xx<sup>e</sup> siècle. Cette capacité de faire « un usage infini de moyens finis », selon la jolie formule de Wilhelm von Humboldt (1767-1835), caractérise la pensée humaine.

On parvient aujourd'hui à modéliser l'apprentissage des règles de la numération ou de la géométrie chez l'enfant – mais pour ce faire, il est nécessaire de postuler que notre cerveau manipule des grammaires et sélectionne, avec agilité, la plus courte et la plus probable d'entre elles<sup>20</sup>. L'apprentissage s'apparente à une sorte de programmation : apprendre, c'est sélectionner, parmi toutes les formules disponibles dans le langage de la pensée, la plus simple qui s'ajuste aux données.

Les réseaux de neurones actuels sont incapables de représenter l'éventail de phrases, de formules, de règles et de théories abstraites avec lesquelles le cerveau d'*Homo sapiens* modélise le monde. Ce n'est sans doute pas un hasard : il y a là quelque chose de profondément humain, que l'on ne retrouve pas dans le cerveau des autres espèces animales, et que les neurosciences contemporaines ne sont pas encore parvenues à résoudre – une vraie singularité de notre espèce. Notre cerveau semble être le seul à disposer d'un jeu de symboles qui se combinent selon une syntaxe complexe et arborescente<sup>21</sup>. Mon laboratoire, par exemple, a montré que le cerveau humain ne peut pas s'empêcher d'entendre une série de sons (bip bip bip boup), sans en théoriser immédiatement la structure abstraite (trois sons identiques suivis d'un son différent). Placés dans la même situation, les singes détectent trois sons, entendent que le dernier est différent, mais ne semblent pas intégrer ces connaissances dans une formule unique<sup>22</sup>. Il faut des dizaines de milliers d'essais pour qu'un singe apprenne à renverser l'ordre d'une séquence (passer de ABCD à DCBA), alors que cinq essais suffisent à n'importe quel gamin de 4 ans<sup>23</sup>. Même un bébé de quelques mois est capable d'inférer des règles abstraites et systématiques – une capacité qui échappe totalement à la fois aux réseaux de neurones conventionnels et aux autres espèces de primates.

- *La composition des connaissances.* Une fois que j'ai appris, disons, à additionner deux chiffres, cette faculté fait partie de mon répertoire de talents : elle devient immédiatement disponible pour l'ensemble de mes facultés mentales. Je peux l'utiliser dans des dizaines de contextes distincts, par exemple au restaurant ou sur ma feuille d'impôt. Surtout, je



peux la combiner avec d'autres facultés – je n'ai aucune difficulté, par exemple, à suivre un algorithme qui me demande de prendre un nombre, de lui ajouter 2 et de décider si le résultat est plus grand ou plus petit que 5. Le cerveau humain semble disposer d'une sorte d'ordinateur interne, une véritable machine de Turing capable d'enchaîner les opérations dans un ordre arbitraire<sup>24</sup>.

Il est étonnant de voir que les réseaux de neurones actuels n'ont pas encore cette flexibilité. Ce qu'ils apprennent reste confiné dans des connexions cachées, inaccessibles, dispersées et très difficiles à réutiliser pour d'autres tâches. La capacité de *composer* les connaissances, c'est-à-dire de les recombinaisonner pour résoudre des problèmes nouveaux, leur échappe. L'intelligence artificielle actuelle ne résout que des problèmes extrêmement étroits : le logiciel AlphaGo, bien que champion du monde du jeu de go, est un expert borné, incapable de généraliser ses talents à tout autre jeu un tant soit peu différent (y compris le jeu de go sur un échiquier  $15 \times 15$  plutôt que  $19 \times 19$ ). Pour notre cerveau, par contre, apprendre, c'est expliciter les connaissances en sorte que l'on puisse les recombinaisonner avec d'autres.

Là encore, nous avons affaire à un aspect singulier du cerveau humain, lié au langage et difficile à reproduire dans une machine. René Descartes l'avait constaté dès 1637 dans le célèbre *Discours de la méthode* :

« S'il y avait [des machines] qui eussent la ressemblance de nos corps, et imitassent autant nos actions que moralement il serait possible, [...] nous aurions toujours deux moyens très certains pour reconnaître qu'elles ne seraient point pour cela de vrais hommes. Le premier est que jamais elles ne pourraient user de paroles ou d'autres signes en les composant, comme nous faisons pour déclarer aux autres nos pensées. Car on peut bien concevoir qu'une machine soit tellement faite qu'elle profère des paroles [...], mais non pas qu'elle les arrange diversement pour répondre au sens de tout ce qui se dira en sa présence, ainsi que les hommes les plus hébétés peuvent faire. Et le second est que, bien qu'elles fissent

plusieurs choses aussi bien ou peut-être mieux qu'aucun de nous, elles manqueraient infailliblement en quelques autres, par lesquelles on découvrirait qu'elles n'agiraient pas par connaissance, mais seulement par la disposition de leurs organes. Car, au lieu que la raison est un instrument universel qui peut servir en toutes sortes de rencontres, ces organes ont besoin de quelque particulière disposition pour chaque action particulière. »

La raison, instrument universel... Les capacités que recense Descartes pointent vers un deuxième système d'apprentissage, hiérarchiquement plus élevé que le précédent, et qui repose sur des règles et des symboles. Dans ses premières étapes, notre système visuel ressemble vaguement aux réseaux de neurones actuels : il filtre ses entrées, apprend les combinaisons fréquentes, et cela lui suffit à reconnaître un visage, un mot, ou une configuration du jeu de go. Mais ensuite, le style de traitement change radicalement : l'apprentissage se met à ressembler à un raisonnement, une inférence logique qui tente de capturer les règles d'un domaine. Parvenir à créer des machines qui atteignent ce second niveau d'intelligence est le grand défi de la recherche contemporaine.

## *Apprendre, c'est inférer la grammaire d'un domaine*

C'est une caractéristique de l'espèce humaine : nous essayons en permanence de tirer, d'une situation particulière, des conclusions de haut niveau, qu'en retour nous mettons à l'épreuve sur de nouvelles observations. Hiérarchiser ainsi ses connaissances, en tentant de formuler des lois abstraites qui rendent compte de nos observations, est une méthode d'apprentissage extraordinairement efficace, car les lois les plus abstraites sont précisément celles qui s'appliquent au plus grand nombre d'observations. Trouver la bonne

loi, la règle logique qui rend compte de toutes les données, c'est accélérer massivement l'apprentissage.

Prenons un exemple : imaginons que je vous présente une dizaine d'urnes remplies de billes de différentes couleurs. Je prends une urne au hasard, dans laquelle je n'ai encore jamais puisé, j'y plonge la main, et j'en sors une bille verte. Pouvez-vous en déduire quoi que ce soit sur le contenu de cette urne ? Quelle sera la couleur du prochain tirage ?

La première réponse qui vous vient sans doute à l'esprit est : « Je n'en ai pas la moindre idée – vous ne m'avez donné pratiquement aucune information, comment pourrais-je connaître la couleur des autres billes ? » Oui mais... imaginez à présent que vous ayez effectué quelques tirages dans les autres urnes et qu'à chaque fois vous ayez constaté la règle suivante : dans une urne donnée, toutes les billes sont de la même couleur. Dans ce cas, le problème devient trivial. Même si vous me montrez une urne pour la première fois, il me suffit d'en extraire une seule boule verte pour en déduire que toutes les autres boules seront de cette couleur. Muni de cette règle générale, un seul tirage me suffit. Ainsi, une connaissance de haut niveau, que l'on appelle le niveau « méta », peut guider tout un ensemble d'observations de plus bas niveau. La métarègle qui dit que toutes les billes d'une urne sont de la même couleur, une fois apprise, accélère massivement l'apprentissage. Bien sûr, elle peut s'avérer fautive. Vous serez vigoureusement surpris (je devrais dire « métasurpris ») si la dixième urne que vous explorez contient des billes de toutes les couleurs. Dans ce cas, vous devrez changer de modèle et remettre en question l'hypothèse que toutes les urnes sont semblables. Peut-être hasarderez-vous une hypothèse d'encore plus haut niveau, une méta-méta-hypothèse selon laquelle les urnes sont de deux sortes : unicolores ou multicolores – auquel cas il vous faudra au moins deux tirages par urne avant de conclure quoi que ce soit. Dans tous les cas, le fait de hiérarchiser vos connaissances et de formuler des hypothèses de haut niveau vous aura fait gagner un temps précieux.

Apprendre efficacement, c'est donc hiérarchiser les informations : se forger, dès que possible, des règles générales, qui résument toute une série d'observations. Dès l'enfance, notre cerveau applique à bon escient ce principe

de hiérarchie. Prenez un petit enfant de 2 ou 3 ans qui se promène dans un jardin et à qui ses parents apprennent un mot nouveau, disons « papillon ». Souvent, il lui suffit d'entendre le mot une fois ou deux, et c'est terminé : son sens est mémorisé. Cette vitesse d'apprentissage est stupéfiante. Elle dépasse tout ce que l'intelligence artificielle d'aujourd'hui parvient à imiter. Le problème, en effet, c'est que le mot est accompagné d'une scène complexe, pleine de fleurs, d'arbres, de jouets, de personnes, qui tous sont des candidats potentiels pour le sens du mot papillon. Et je ne parle même pas des significations moins évidentes : chaque instant que nous vivons est plein de bruits, d'odeurs, de mouvements, d'actions, mais aussi d'idées abstraites qui possèdent aussi un nom qu'il a bien fallu apprendre. Ainsi « papillon » pourrait-il vouloir dire « couleur », « ciel », « bouge » ou « symétrie ». Le problème se pose avec la plus grande acuité pour les mots abstraits. Comment un enfant apprend-il les mots « penser », « croire », « non », « liberté » ou « mort » ? Comment comprend-il ce que veut dire « je », alors que, chaque fois qu'il l'a entendu, son interlocuteur parlait de... lui ? !

L'apprentissage rapide des mots abstraits disqualifie les modèles trop simplistes de l'apprentissage. Les réseaux de neurones qui essaient simplement de corréliser les entrées avec les sorties, les images avec les mots, mettent des milliers d'essais à comprendre que le mot « papillon » fait référence à cet objet coloré, là, dans un coin de l'image... et ce principe de corrélation mot-image ne permettra jamais de comprendre des mots dépourvus de référence fixe, comme « nous », « toujours » ou « odeur ».

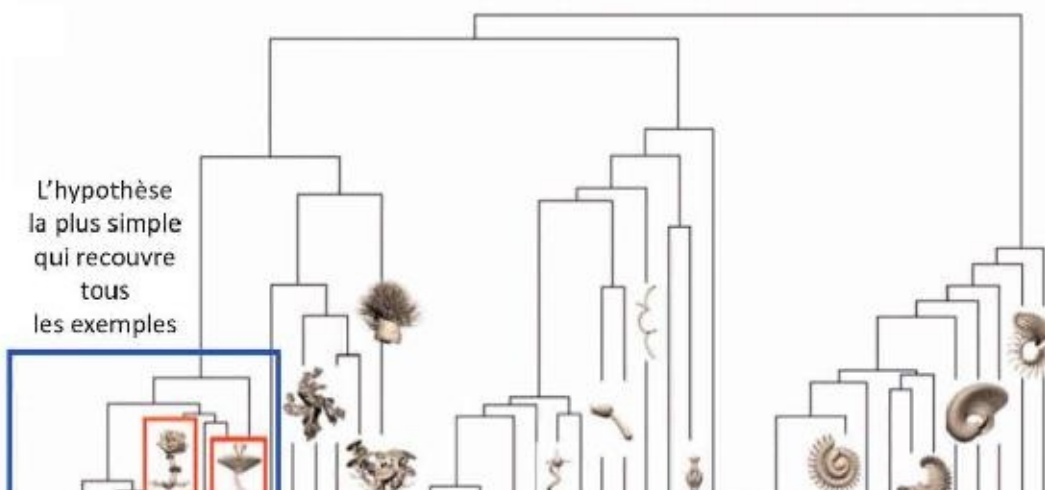
L'apprentissage des mots pose un immense défi aux sciences cognitives. Nous savons cependant qu'une partie de la solution réside dans l'existence de représentations non linguistiques, abstraites, logiques. Avant même d'apprendre les mots, l'enfant possède une sorte de langage de l'esprit avec lequel il peut formuler des hypothèses très abstraites, et les mettre à l'épreuve ; son cerveau n'est pas une ardoise vierge, et ses pensées organisées restreignent fortement l'espace d'apprentissage. De plus, si un enfant apprend aussi vite le sens des mots, bien plus vite qu'un réseau de neurones, c'est qu'il sélectionne parmi l'ensemble des hypothèses possibles en se laissant guider par toute une panoplie

de règles de haut niveau – des métarègles qui accélèrent massivement l'apprentissage, exactement comme dans le problème de l'urne.

L'une de ces règles, c'est de toujours privilégier l'hypothèse la plus petite qui soit compatible avec l'ensemble des données. Pour prendre un exemple, lorsqu'un bébé entend sa maman dire « regarde le chien », rien n'exclut que le mot « chien » fasse référence à ce chien-là en particulier (Médor) – ou, à l'inverse, qu'il signifie n'importe quel quadrupède, mammifère ou même animal. Comment l'enfant découvre-t-il la véritable signification de ce mot : seulement les chiens, mais tous les chiens ? Des expériences suggèrent qu'il raisonne en bon logicien, en testant toutes les hypothèses mais en ne conservant que la plus étroite qui colle avec ce qu'il a entendu. S'il n'entend le mot « chien » que dans un seul contexte, il peut, temporairement, croire que ce mot ne réfère qu'à un animal unique, ce chien-là – mais dès qu'il en voit deux, il peut en déduire que ce mot renvoie à toute une catégorie (figure 5). Un modèle mathématique de ce processus prédit qu'il suffit de trois ou quatre instances pour converger vers la signification d'un mot<sup>25</sup>. C'est ce que font les enfants, dix, cent ou mille fois plus vite que les réseaux de neurones actuels.

D'autres astuces permettent également à l'enfant d'apprendre à parler en un temps record. L'une de ces métarègles exprime un truisme : en général, celui qui parle fait attention à ce dont il parle. Muni de cette règle, le bébé peut considérablement restreindre son espace de recherche : il n'est pas obligé de corréler chaque mot avec tous les objets présents dans la scène visuelle, comme le ferait un ordinateur, jusqu'à ce qu'il obtienne suffisamment de données qui lui prouvent que, chaque fois qu'il entend parler de « papillon », le petit insecte coloré est bien présent. Il lui suffit de suivre le regard de ses parents, ou la direction de leur doigt, afin d'essayer de comprendre de quoi ils parlent : c'est ce que l'on appelle l'attention partagée, un principe fondamental de l'apprentissage du langage.

Voici trois « tufas ». Saurez-vous trouver les autres ?





**Figure 5.** Apprendre, c'est sélectionner le modèle le plus simple qui s'ajuste aux données. Supposez que je vous montre la carte du haut et que je vous dise que les 3 objets entourés de rouge sont des « tufas ». Avec si peu de données, comment faites-vous pour retrouver les autres tufas ? Votre cerveau se fait un modèle de la manière dont ces formes ont été engendrées, un arbre hiérarchique de leurs propriétés, et il sélectionne la plus petite branche de l'arbre compatible avec toutes les données du problème.

L'expérience a été faite : prenez un enfant de 2 ou 3 ans, mettez-le en présence d'un jouet nouveau, et d'un adulte qui dit : « Oh, un wog ! » Il ne faut pas plus d'un essai pour que l'enfant retienne que « wog » est le nom de cet objet. Maintenant, composez la même scène, sauf que l'adulte reste muet, c'est un haut-parleur situé au plafond qui dit : « Oh, un wog ! » L'enfant n'apprend rigoureusement rien, parce qu'il ne parvient pas à comprendre l'intention du locuteur<sup>26</sup>. Un bébé ne retient les mots que s'il comprend l'intention de celui qui parle. C'est cette capacité, remarquablement abstraite, qui lui permet d'apprendre des mots : pour ce faire, il doit littéralement se mettre à la place du locuteur et comprendre à quelles pensées les mots qu'il prononce font référence.

L'enfant utilise encore d'autres métrarègles pour apprendre les mots. Par exemple, il se sert du contexte grammatical. Lorsqu'on lui dit : « Oh, le papillon », la présence de l'article « le » lui garantit que ce mot inconnu est un nom. C'est une métrarègle qu'il lui a fallu apprendre : on ne naît pas avec une connaissance innée des articles de toutes les langues. Cependant, l'expérience montre que cet apprentissage est rapide : dès 12 mois, les enfants ont enregistré les principaux articles et mots grammaticaux de leur langue maternelle, et les utilisent pour guider les apprentissages ultérieurs<sup>27</sup>.

S'ils y parviennent, c'est sans doute parce que ces mots sont très fréquents et apparaissent dans de nombreuses phrases, presque toujours devant un nom. Le raisonnement peut paraître circulaire, mais il ne l'est pas vraiment : bébé commence par apprendre ses premiers noms, souvent des substantifs

ultrafamiliers comme « biberon » et « chien », aux alentours de 6 mois... puis il repère que ces mots sont souvent précédés d'un mot très fréquent, l'article « le »... d'où il déduit que tous ces mots font partie de la même catégorie « nom »... et que celle-ci réfère souvent à des choses... ce qui lui permet enfin, lorsqu'il entend « le papillon », de lui chercher un sens possible parmi les choses qui l'entourent, plutôt que de le traiter comme un verbe ou un adjectif. Ainsi, chaque épisode d'apprentissage renforce la règle, qui elle-même facilite les apprentissages ultérieurs, dans un vaste mouvement qui s'accélère chaque jour. On retrouve ici l'idée de *bootstrapping* : personne ne peut s'élever dans le ciel en se tirant par les lanières de ses bottes, comme le Baron de Münchhausen, mais un algorithme d'apprentissage parvient à décoller progressivement en s'appuyant sur très peu.

Il y a encore au moins une métarègle que l'enfant utilise pour accélérer l'apprentissage des mots. C'est la règle d'exclusivité que l'on pourrait énoncer ainsi : un seul mot pour chaque chose. Elle dit qu'il est peu probable qu'il y ait deux noms pour le même concept. Un mot nouveau fait donc, très probablement, référence à un concept nouveau. Muni de cette règle, dès qu'il entend un mot qu'il ne connaît pas, l'enfant peut restreindre sa recherche de sens aux choses dont il ne connaît pas encore le nom. Et, dès 16 mois, il l'applique à bon escient<sup>28</sup>. Tentez l'expérience : prenez deux bols identiques, l'un d'un bleu très banal, et l'autre d'une couleur inhabituelle, disons vert olive ; et dites à l'enfant : « Donne-moi le bol crapité. » L'enfant va vous donner le bol qui n'est pas bleu (mot qu'il connaît déjà) et, des semaines plus tard, il se souviendra que « crapité » fait référence à cette étrange couleur.

Là encore, la métarègle accélère massivement l'apprentissage. Mais il a sans doute fallu l'apprendre, ou au moins la renforcer. En effet, l'expérience montre que les enfants de famille bilingue l'appliquent beaucoup moins que les autres<sup>29</sup> – ils se rendent bien compte que papa et maman utilisent des mots différents pour dire la même chose. Les enfants monolingues, par contre, s'appuient fortement sur la règle d'exclusivité. Si vous leur dites « donne-moi le glax », alors qu'il n'y a que des objets familiers devant eux, ils vont chercher partout, à la recherche de ce mystérieux objet inconnu auquel vous semblez faire référence.



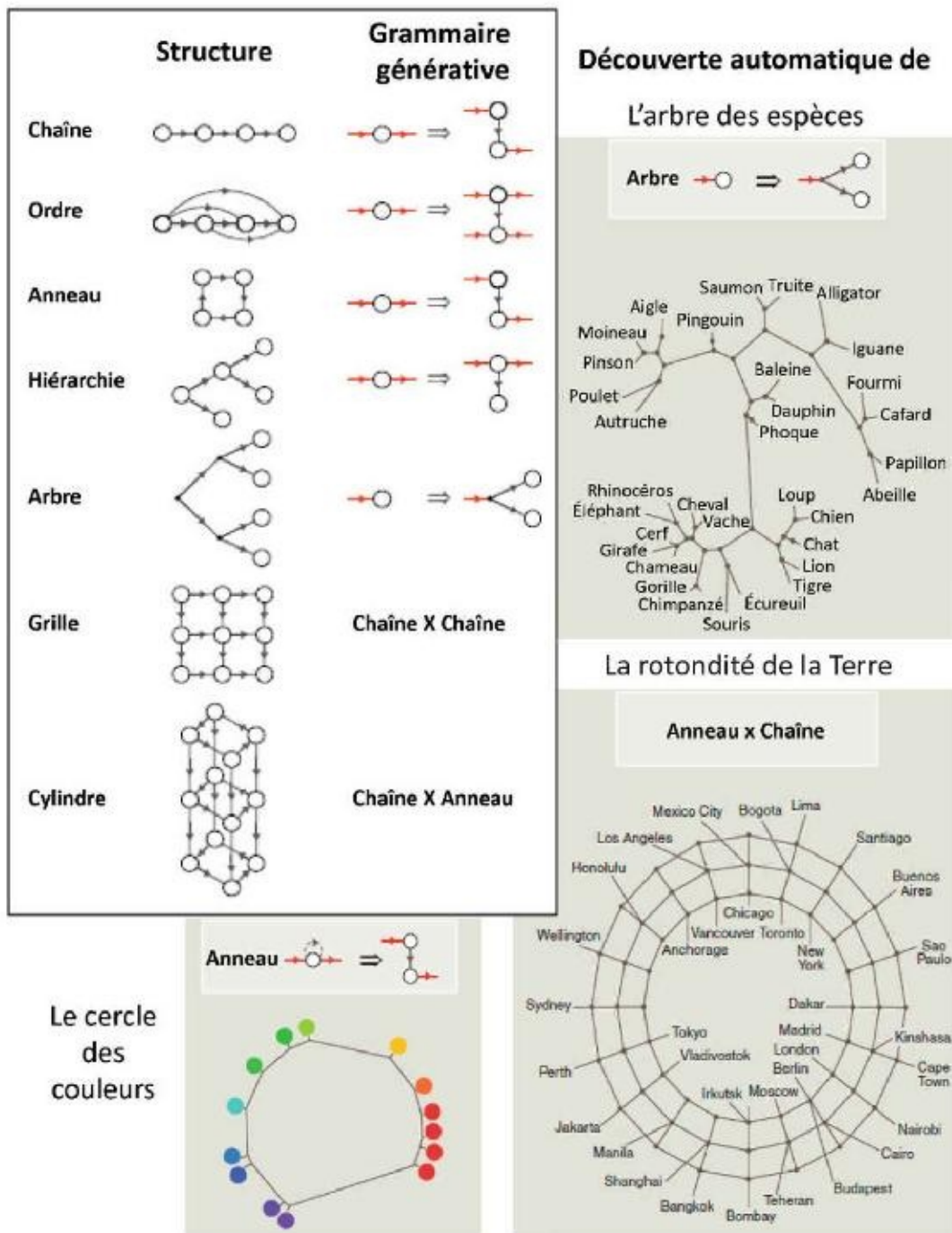
Toutes ces métarègles illustrent ce que l'on appelle la bénédiction de l'abstraction : chaque mot, chaque phrase que l'enfant entend vient renforcer une règle de haut niveau (par exemple, les noms ont tendance à être précédés d'un déterminant « le » ou « la »), qui peut donc être apprise très vite. Grâce à elles, vers 2-3 ans, dans cette période bénie qu'on appelle à juste titre l'explosion lexicale, l'enfant apprend entre 10 et 20 mots nouveaux par jour, sans aucune difficulté, juste sur la base d'indices si ténus qu'ils continuent de faire caler les meilleurs algorithmes de la planète.

La capacité d'utiliser des métarègles semble demander une bonne dose d'intelligence. Est-elle pour autant propre à l'espèce humaine ? Pas entièrement. D'autres animaux sont également capables de faire preuve d'abstraction. Prenez le cas de Rico, un chien de berger que sa maîtresse a dressé à rapporter des objets les plus divers<sup>30</sup>. Il suffit de lui ordonner : « Rico, va chercher le dinosaure », et l'animal passe dans la salle de jeux pour revenir, quelques secondes plus tard, avec un dinosaure en peluche dans la gueule. Les éthologues qui l'ont testé ont démontré que Rico connaît près de 200 mots. Mais, le plus extraordinaire, c'est que lui aussi utilise le principe d'exclusion pour apprendre des mots nouveaux. Si vous lui dites : « Rico, rapporte le sikirid » (un mot nouveau), Rico se met à chercher parmi les objets présents et ne revient qu'avec un objet nouveau, dont il ne connaît pas encore le nom. Lui aussi utilise la métarègle « un seul nom pour chaque chose ».

Comme les mathématiciens et les informaticiens sont des gens astucieux, ils commencent à découvrir quels algorithmes permettent à des machines d'apprendre ce genre de hiérarchie de règles, de métarègles et même de méta-métarègles, jusqu'à un niveau arbitraire. Dans ces algorithmes hiérarchiques, chaque épisode d'apprentissage contraint non seulement les paramètres de bas niveau, mais aussi les connaissances du plus haut niveau, les hyperparamètres abstraits qui, à leur tour, vont biaiser les apprentissages ultérieurs. S'ils ne savent pas encore imiter l'extraordinaire efficacité du cerveau humain, ces systèmes atteignent parfois des performances remarquables. À titre d'exemple, la figure 6 montre l'un de ces algorithmes récents, une sorte de scientifique artificiel qui trouve le meilleur modèle du monde extérieur<sup>31</sup>. Ce système dispose non

seulement de règles abstraites, mais aussi d'une véritable grammaire qui lui permet, en combinant ces règles élémentaires, d'engendrer une infinité de structures géométriques. Pour lui, une ligne, c'est un enchaînement de points reliés de proche en proche, par l'application de la règle « chaque point possède deux voisins, un à gauche, un à droite » – et le système découvre tout seul que c'est la meilleure manière d'organiser les entiers naturels (une ligne qui va de zéro à l'infini) comme les politiciens (depuis l'ultragauche jusqu'à l'extrême droite). Une variante de la même grammaire engendre un arbre binaire où chaque nœud possède un parent et au moins deux enfants. C'est cette structure d'arbre qui est automatiquement sélectionnée lorsqu'on demande au système de se représenter les êtres vivants – la machine se comporte comme un véritable Darwin artificiel.

D'autres combinaisons de règles engendrent des plans (la carte d'une ville), des cylindres, des sphères (notre planète)... L'algorithme existe également dans des versions plus sophistiquées capables d'exprimer des idées encore plus abstraites. Par exemple, les informaticiens américains Noah Goodman et Josh Tenenbaum ont conçu un système capable de découvrir le principe de causalité<sup>32</sup> – l'idée même que certains événements en causent d'autres. Sa formulation est absconse et mathématique : « Dans un graphe directionnel et acyclique liant diverses propriétés entre elles, il existe un sous-ensemble de propriétés dont toutes les autres dépendent. » Peu importe que cette expression soit presque incompréhensible – elle illustre le type de formules internes que cette grammaire mentale est capable d'exprimer et de tester. Le système la met à l'épreuve, et ne la conserve que si les données dont on le nourrit s'y conforment (c'est-à-dire si certaines variables sont des causes et d'autres des conséquences). Une fois de plus, le fait d'être capable de formuler une règle de haut niveau permet d'accélérer massivement l'apprentissage, parce que la règle restreint l'espace des possibles.



**Figure 6.** Apprendre, c'est inférer la grammaire d'un domaine. Au MIT, deux informaticiens ont inventé un algorithme qui découvre, seul, la structure d'un domaine scientifique. Le système dispose d'une grammaire de règles dont les combinaisons engendrent toutes sortes de structures nouvelles : lignes, plans, cercles, cylindres, arbres.... En sélectionnant la structure la plus appropriée aux données, l'algorithme fait des découvertes qui ont pris des années aux scientifiques :

l'arbre des espèces animales (Darwin, 1859), la rotondité de la Terre (Parménide d'Élée, 600 avant J.-C.), ou encore le cercle des couleurs (Newton, 1675).

Selon cette vision, apprendre consiste à sélectionner, parmi toutes les grammaires abstraites, celle qui s'ajuste le mieux aux données. Nous verrons que c'est un excellent modèle de ce que fait l'enfant : scientifique en herbe, il formule des théories et les confronte au monde extérieur. Mais cela suppose que l'enfant soit bien plus qu'un simple réseau de neurones. Dès la naissance, il dispose déjà de deux ingrédients essentiels : d'une part, toute la machinerie qui permet d'engendrer des formules abstraites, ce que l'on pourrait appeler un *langage de la pensée* ; d'autre part, la capacité de sélectionner ces idées à bon escient en fonction de leur adéquation aux données reçues par le cerveau.

Telle est la nouvelle vision du cerveau<sup>33</sup> : un immense modèle génératif, massivement structuré, capable d'imaginer des myriades de règles et de structures hypothétiques, d'halluciner d'immenses espaces de pensées – mais qui se contraint progressivement à ne produire que celles qui s'ajustent à la réalité.

## *Apprendre, c'est raisonner en bon statisticien*

Oui, mais comment sélectionner la meilleure hypothèse ? Sur quel critère sélectionner une idée, un modèle du monde extérieur ? Il se trouve qu'il existe une stratégie idéale pour le faire. C'est cette stratégie que je voudrais maintenant expliciter, parce qu'elle constitue le noyau d'une des théories les plus récentes et les plus productives de l'apprentissage : l'hypothèse que le cerveau se comporte comme un scientifique en herbe, qui calcule en permanence avec les probabilités. Selon cette théorie, apprendre, c'est raisonner en bon statisticien et choisir, parmi plusieurs hypothèses, celle qui a la plus forte probabilité d'être correcte, en tenant compte des données disponibles.

Comment fonctionnent les théories scientifiques ? Lorsqu'un scientifique formule une théorie, il ne se contente pas d'exhiber des formules mathématiques – il formule des prédictions. L'importance de la théorie se juge à la richesse des prédictions originales qui en découlent. C'est la vérification ou la réfutation de ces prédictions qui entraîne la validation ou la chute d'une théorie. Le chercheur applique une logique simple : j'énonce plusieurs théories, je déroule l'écheveau de leurs prédictions, et j'élimine celles dont les prédictions sont invalidées par l'expérience.

Évidemment, tout chercheur sait qu'une seule expérience suffit rarement : il faut souvent accumuler les statistiques pour démêler le vrai du faux et remonter, de fil en aiguille, à la meilleure hypothèse. Ce lent processus d'accumulation du savoir s'apparente à un apprentissage. En science, l'ignorance recule au fur et à mesure qu'une suite de conjectures et réfutations (selon le titre du livre du philosophe Karl Popper) permet de raffiner la théorie. De la même manière, dans chacune de nos têtes, l'ignorance s'efface lorsque notre cerveau parvient, au fil de ses observations, à formuler une théorie de plus en plus juste du monde extérieur.

Notre cerveau agit-il en bon scientifique ? Depuis une trentaine d'années, ce qui n'était qu'une métaphore est devenu un algorithme effectif et une théorie précise du cerveau. Mathématiciens et informaticiens ont formulé une théorie du raisonnement idéal en présence d'incertitudes. Cette théorie sophistiquée porte un nom : on la dit bayésienne, du nom de son tout premier découvreur, le révérend Thomas Bayes (1701-1761), un pasteur presbytérien et mathématicien anglais devenu membre de la Royal Society. Mais peut-être devrait-on l'appeler laplacienne, car c'est au grand mathématicien français Pierre-Simon, marquis de Laplace (1749-1827), que l'on doit sa première formalisation complète. C'est seulement dans les vingt dernières années qu'elle est devenue l'un des piliers de l'apprentissage par ordinateur, car elle seule garantit l'extraction d'un maximum d'informations de chaque donnée, si minuscule soit-elle.

Qu'ont découvert Bayes et Laplace ? Tout simplement la bonne manière de raisonner avec les probabilités afin de remonter des observations à leurs causes. Apprendre, c'est parvenir à tirer un maximum de déductions de chaque

observation, même les plus incertaines – et c’est précisément ce que garantit la règle de Bayes.

Revenons aux fondements de la logique. Depuis l’Antiquité, l’humanité a compris comment raisonner avec des valeurs de vérité, « vrai » ou « faux ». Aristote a introduit les règles de déduction qu’on appelle syllogismes et que nous appliquons tous plus ou moins intuitivement. Par exemple, la règle appelée *modus tollens* (littéralement « procédé qui nie ») dit que, si P implique Q et qu’il s’avère que Q est fausse, alors P doit être fausse. C’est cette règle qu’applique Sherlock Holmes dans un épisode fameux, connu sous le nom du « curieux incident du chien pendant la nuit » :

- Y a-t-il quelque autre point sur lequel vous désireriez attirer mon attention ?, demande l’inspecteur Gregory, de Scotland Yard.
- Le curieux incident du chien pendant cette nuit-là, répond Sherlock.
- Mais... le chien n’a rien fait cette nuit-là !
- C’est justement ce qu’il y a de curieux !

En effet, raisonne Sherlock, si le chien avait repéré un inconnu, *alors* il aurait aboyé. Puisqu’il ne l’a pas fait, c’est que le criminel était une personne familière... un raisonnement qui permet au célèbre détective de restreindre la recherche et de démasquer le coupable.

Quel rapport, me direz-vous, avec l’apprentissage ? Eh bien, apprendre, c’est raisonner tel un détective : il s’agit toujours de remonter aux causes cachées des phénomènes, afin d’en déduire le modèle qui les gouverne. Mais, dans le monde réel, les observations sont rarement vraies ou fausses : elles sont incertaines, probabilistes. Et c’est exactement là que se situent les contributions fondamentales du révérend Bayes et du marquis de Laplace : la théorie bayésienne nous dit comment raisonner avec les probabilités, quelles sortes de syllogismes nous devons appliquer lorsque les données ne sont pas parfaites, mais probabilistes, afin de remonter aux causes probables de nos observations. *La Probabilité, logique de la science* : tel est le titre d’un très beau livre du

statisticien E. T. Jaynes sur la théorie bayésienne<sup>34</sup>. Il y montre que ce que nous appelons probabilité n'est autre que l'expression de notre incertitude. La théorie exprime, avec une précision mathématique, les lois selon lesquelles l'incertitude doit évoluer lorsque nous faisons une observation nouvelle. C'est l'extension parfaite de la logique au brouillard des probabilités et des incertitudes.

Prenons un exemple, proche de celui que le révérend Bayes s'était posé au XVIII<sup>e</sup> siècle. Supposons que j'observe un bateleur jouer à pile ou face avec une pièce de monnaie. Si la pièce n'est pas truquée, elle a autant de chances de tomber sur pile que sur face : fifty-fifty. La théorie classique des probabilités, celle qu'on enseigne à l'école, dit comment, à partir de cette probabilité, prédire les chances d'observer un certain tirage (par exemple 3 piles et 2 faces). La théorie bayésienne, qu'on devrait enseigner, dit bien plus. Elle décrit comment suivre le chemin inverse : si j'observe plusieurs tirages, comment dois-je réviser mon opinion sur la pièce ? L'hypothèse par défaut, c'est que la pièce n'est pas biaisée... sauf si je la vois tomber vingt fois sur pile ! Dans ce cas, il me faut réviser mon *a priori*, car la pièce est certainement truquée. Mon hypothèse de départ est devenue improbable. Mais de combien ? La théorie de Bayes et de Laplace explique précisément comment réviser nos croyances après chaque tirage. Chaque hypothèse est affublée d'un nombre qui correspond à un niveau de confiance. À chaque observation, ce nombre change d'une valeur qui dépend du degré d'improbabilité du tirage observé. Exactement comme en sciences, plus une observation expérimentale est improbable, plus elle viole les prédictions d'une théorie, plus elle permet de la rejeter.

La théorie bayésienne est remarquablement efficace. Elle a été utilisée, par exemple, par Alan Turing pour décrypter le code Enigma – le fameux mécanisme qui servait à crypter les messages allemands au cours de la Seconde Guerre mondiale. La machine Enigma comprenait de très nombreux rouages et rotors, câblés de telle sorte que leurs combinaisons engendraient plus d'un milliard de configurations différentes qui changeaient après chaque lettre. Chaque matin, le responsable du chiffrement plaçait la machine dans la configuration prévue pour cette journée, tapait un texte en entrée, et Enigma crachait une suite infernale de lettres que seul le possesseur de la clé de

chiffrement pouvait décoder. À toute autre personne, le texte semblait parfaitement aléatoire. Mais ici réside le génie de Turing : il découvrit que, si deux machines avaient été initialisées de la même manière, cela introduisait un léger biais dans la distribution des lettres, en sorte que deux messages avaient un peu plus de chances de se ressembler. Ce biais était si faible qu'aucune lettre isolée ne suffisait à conclure ; mais en accumulant les improbabilités, lettre après lettre, Turing pouvait acquérir la certitude que la même configuration avait été utilisée deux fois et, de fil en aiguille, avec l'aide d'un instrument qui préfigurait nos ordinateurs, briser le code Enigma.

Imaginez, à présent, que le même type de raisonnement se produise à l'échelle de tout le cerveau<sup>35</sup>. Chaque région du cerveau formule une ou plusieurs hypothèses et transmet ses prédictions aux autres. Ainsi, chaque module contraint les hypothèses du suivant, en échangeant des messages qui sont autant de prédictions probabilistes sur le monde extérieur. Ces signaux sont dits « descendants » parce qu'ils partent des aires cérébrales de haut niveau, comme le cortex frontal, et qu'ils descendent, de proche en proche, jusqu'aux aires sensorielles de bas niveau. On pense que ces signaux expriment le nuage d'hypothèses que notre cerveau juge plausibles *a priori* et qu'il souhaite mettre à l'épreuve.

Au niveau des aires sensorielles, ces *a priori* viennent se froter aux messages « ascendants » issus du monde extérieur, de la rétine par exemple. C'est à ce moment que le modèle se confronte à la réalité. L'approche bayésienne repose sur le calcul d'un signal d'erreur, la différence entre ce que le modèle prédit et ce qui a été observé. L'algorithme bayésien indique comment se servir de ce signal d'erreur afin d'ajuster le modèle interne du monde. S'il n'y a pas d'erreur, c'est que le modèle est juste. Sinon, le signal d'erreur remonte la chaîne des aires cérébrales et ajuste, un peu partout, les hypothèses du modèle. Plus ou moins rapidement, l'algorithme converge vers un modèle adéquat du monde extérieur.

Selon cette vision du cerveau, notre jugement d'adulte s'appuie sur deux types de connaissances : le savoir inné de notre espèce (que les bayésiens appellent l'*a priori*), la pondération des espaces d'hypothèses que nous héritons



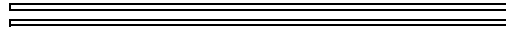
de notre évolution) et l'expérience que nous avons acquise en tant qu'individu (la probabilité *a posteriori* , somme de toutes les inférences que nous avons pu faire au cours de notre vie). Cette division du travail dissout le débat classique entre l'inné et l'acquis. Tout savoir doit s'appuyer sur les deux composantes : un jeu d'hypothèses *a priori* , et un ensemble de règles d'apprentissage qui permettent de les trier.

On peut démontrer mathématiquement que l'approche bayésienne est la meilleure manière d'apprendre. C'est la seule façon d'extraire la substance moelle d'un épisode d'apprentissage, d'en tirer le maximum d'informations. Même quelques bribes d'information, telles les coïncidences suspectes que Turing avait repérées dans le code d'Enigma, permettent d'apprendre. À partir du moment où le système les traite en bon statisticien, en accumulant patiemment les probabilités, il finira fatalement par obtenir une masse suffisante de données pour réfuter certaines théories et en valider d'autres.

Le cerveau fonctionne-t-il ainsi ? Est-il capable de générer, dès la naissance, de vastes espaces d'hypothèses *a priori* parmi lesquels l'apprentissage vient piocher ? Procède-t-il par élimination, en formulant des hypothèses sophistiquées et en les sélectionnant au fur et à mesure que les données observées leur sont favorables ou contraires ? Le bébé agit-il, dès le berceau, comme le plus malin des statisticiens ? Parvient-il à extraire le maximum d'informations de chaque épisode d'apprentissage ? C'est ce que nous allons voir à présent.

DEUXIÈME PARTIE

# COMMENT NOTRE CERVEAU APPREND



« Au-dedans de son âme, chacun possède la puissance du savoir, ainsi que l'organe au moyen duquel chacun acquiert l'instruction. »

PLATON , *La République* (~380 av. J.-C.).

Le débat entre inné et acquis fait rage depuis des millénaires. Le bébé ressemble-t-il à une page blanche, une ardoise vierge, une bouteille vide que l'expérience doit remplir ? Dès le IV<sup>e</sup> siècle avant notre ère, Platon balayait déjà du revers de la main l'idée que notre cerveau vienne au monde dépourvu de toute connaissance. Bien au contraire, affirme Platon, nous sommes les dépositaires, dès la naissance, de deux structures massives : « la puissance du savoir » et « l'organe au moyen duquel chacun acquiert l'instruction ».

Convergence remarquable : nous venons de voir, au chapitre précédent, que c'est aussi la conclusion de la toute récente science des données. L'apprentissage est infiniment plus efficace si l'on dispose, d'une part, d'un vaste espace d'hypothèses, un ensemble de modèles mentaux pourvus de myriades de réglages parmi lesquels choisir ; et d'autre part, d'algorithmes sophistiqués capables d'ajuster leurs paramètres en fonction des données reçues du monde extérieur. Comme le disait l'un de mes amis, dans le débat sur l'inné et l'acquis, on a sous-estimé les deux ! Apprendre exige une double architecture : une immense quantité de modèles génératifs internes, et des algorithmes efficaces pour les ajuster à la réalité.

Les réseaux de neurones artificiels le font à leur manière, en confiant la représentation des modèles mentaux à des millions de connexions ajustables. Cependant, ces systèmes, s'ils capturent assez bien la reconnaissance rapide et inconsciente des images ou de la parole, ne savent pas encore bien représenter des hypothèses plus abstraites, comme les règles de grammaire ou la logique des opérations mathématiques.

Le cerveau humain semble fonctionner sur un mode différent : notre savoir croît par la combinaison progressive de symboles. Selon l'hypothèse du cerveau bayésien, nous venons au monde avec une vaste combinatoire de pensées potentielles. Ce langage de la pensée, doté de primitives abstraites et de règles de grammaire, préexiste à tout apprentissage. Il engendre un immense espace d'hypothèses qu'il ne reste plus qu'à mettre à l'épreuve. Pour ce faire, notre cerveau agit comme un scientifique : en collectant des données statistiques qu'il sélectionne parmi un vaste choix de modèles génératifs possibles du monde extérieur.

Cette vision de l'apprentissage peut sembler contre-intuitive. Elle suggère que le cerveau de chaque bébé humain contient déjà, en puissance, toutes les langues du monde, tous les objets, tous les visages, tous les outils qu'il pourra rencontrer, et également tous les mots, tous les faits, tous les événements dont il pourra se souvenir. La combinatoire du cerveau doit être telle que tous ces objets de pensée sont déjà là, en puissance, avec leurs probabilités *a priori* – et la capacité de les mettre à jour lorsque l'expérience exige qu'elles soient révisées. Est-ce ainsi que fonctionne un bébé ?

# CHAPITRE 3

## Le savoir invisible : les étonnantes intuitions des bébés

---

Quoi de plus démuni, en apparence, qu'un nouveau-né ? Quoi de plus raisonnable que de penser, avec Locke, que cette « ardoise vierge » doit encore recevoir la marque de l'environnement ? Rousseau le martelait déjà avec force dans *Émile ou De l'éducation* (1762) : « Nous naissons capables d'apprendre, mais ne sachant rien, ne connaissant rien. » Près de deux siècles plus tard, l'hypothèse est reprise par Alan Turing, père de l'informatique contemporaine : « Il est vraisemblable que le cerveau de l'enfant soit comme un cahier qu'on vient d'acheter chez le papetier. Guère de mécanismes, beaucoup de pages vierges. »

Nous savons aujourd'hui que rien n'est plus faux. Les apparences sont trompeuses. En dépit de son immaturité, le petit de l'espèce humaine possède déjà un vaste savoir, hérité de son histoire évolutive. Mais cela ne se voit pas dans son comportement. Il a fallu des avancées méthodologiques importantes en sciences cognitives pour mettre en évidence l'extraordinaire base de données

avec laquelle chaque bébé vient au monde. Objets, nombres, probabilités, visages, langage, *etc.* : l'étendue des capacités des bébés est loin d'être négligeable.

## *Le concept d'objet*

Nous avons tous l'intuition que le monde est fait d'objets rigides. En réalité, il est constitué d'atomes, mais à l'échelle où nous vivons, ces atomes s'agglomèrent et se comportent souvent comme des entités cohérentes qui bougent ensemble, s'entrechoquent... ce que nous appelons des « objets ». L'existence des objets est une propriété fondamentale de notre environnement. Avons-nous besoin de l'apprendre ? Non. Des millions d'années d'évolution semblent avoir intégré cette connaissance au plus profond de nos cerveaux. Dès quelques mois de vie, un bébé sait déjà que le monde est constitué d'objets, qu'ils bougent d'un seul bloc, qu'ils occupent de l'espace, qu'ils ne peuvent pas disparaître sans raison, et qu'ils ne peuvent pas être deux à occuper la même place<sup>36</sup>. Leur cerveau connaît déjà les lois de la physique qui énoncent que la trajectoire d'un objet, dans l'espace comme dans le temps, doit être continue et sans rupture.

Comment le sait-on ? Parce que les bébés manifestent leur surprise dans des situations expérimentales qui, par malice, violent les lois de la physique des objets. Dans les laboratoires de sciences cognitives, les expérimentateurs se transforment en prestidigitateurs ([figure 7](#)). Sur de petits théâtres spécialement conçus pour les bébés, ils leur jouent toutes sortes de tours : des objets apparaissent, disparaissent, se multiplient, se dissolvent... L'expérience montre que même un bébé de quelques semaines est sensible à la magie : il possède déjà des intuitions profondes du monde physique et, comme nous tous, il est abasourdi lorsque celles-ci sont réfutées.



**Figure 7.** Loin d'être des ardoises vierges de toute connaissance, les bébés possèdent un vaste savoir en physique, en arithmétique, en probabilité ou en géométrie. Au laboratoire, la mesure de leur surprise, lorsqu'ils sont soumis à des situations qui violent les règles de ces domaines, révèle la sophistication de leurs intuitions du monde.

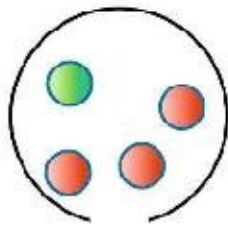
En filmant un enfant de très près, afin de déterminer où s'oriente son regard et surtout pendant combien de temps il scrute la scène, on peut mesurer finement son degré d'étonnement, et en déduire quels étaient ses attentes, son savoir implicite.

Cachons un objet derrière un livre, puis, d'un coup sec, écrasons-le à plat, comme si l'objet qu'il masquait n'existait plus (une petite trappe, prévue à cet effet, l'a escamoté) : les bébés sont stupéfaits ! Ils ne parviennent pas à comprendre qu'une chose puisse s'évanouir sans laisser de trace. Ils manifestent également leur étonnement lorsqu'un objet disparaît derrière un écran et réapparaît derrière un autre, sans qu'on l'ait vu franchir l'espace vide entre les deux écrans. Ils sont épatés qu'un petit train descende une pente en semblant traverser une paroi rigide sans être arrêté par elle. Et ils s'attendent à ce qu'un objet forme un tout cohérent : s'ils voient les deux extrémités d'un bâton bouger de façon cohérente de chaque côté d'un écran, ils s'attendent à voir un seul bâton, et sont saisis de stupeur lorsque l'écran s'abaisse et leur laisse voir deux bâtons distincts ([figure 8](#)).

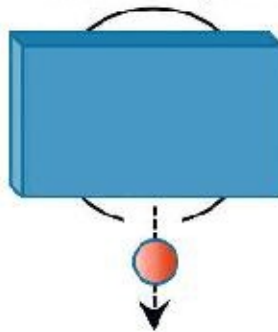
Les bébés possèdent donc un vaste savoir sur le monde, mais ils ne connaissent pas tout dès le départ, loin s'en faut. Il faut quelques mois pour qu'un bébé comprenne comment deux objets se soutiennent l'un l'autre<sup>37</sup>. Au départ, il ignore qu'un objet tombe quand on le lâche. Ce n'est que très progressivement qu'il prend conscience de tous les facteurs qui font qu'un objet tombe ou ne tombe pas. Lorsqu'on place un jouet au bord d'une table par exemple, le bébé commence par découvrir que, pour ne pas tomber, l'objet doit être en contact avec la table ; puis qu'il faut qu'il y soit bien posé, ni dessous ni contre, mais sur le dessus. Il lui faudra encore quelques mois pour s'apercevoir que cette règle ne suffit pas : encore faut-il que le centre de gravité de l'objet reste au-dessus de la table.



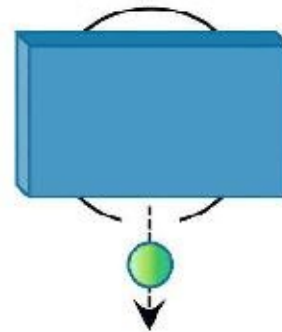
### Intuition des nombres et des probabilités



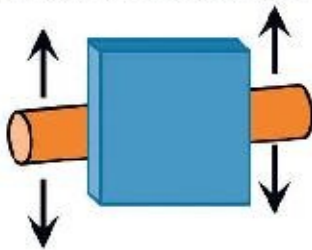
Pas de surprise



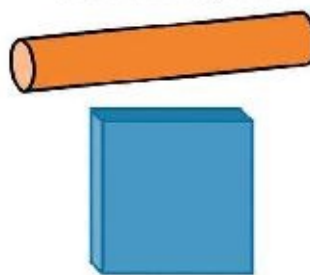
Surprise !



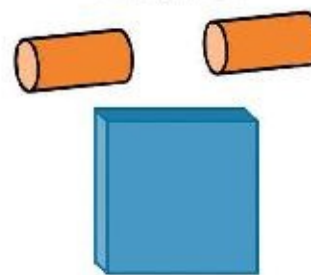
### Intuition des objets



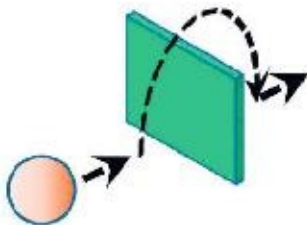
Pas de surprise



Surprise !



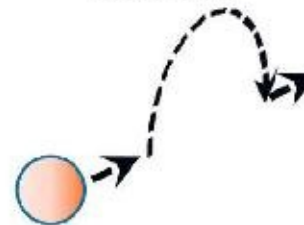
### Intuition de la psychologie



Pas de surprise



Surprise !



**Figure 8.** Pour évaluer les intuitions des bébés, on détermine s'ils regardent plus longtemps une scène surprenante. Lorsqu'une urne contient une majorité d'objets rouges, bébé est surpris qu'il en sorte un objet vert (intuition des nombres et des probabilités). Si deux bouts d'un bâton bougent de façon cohérente, bébé est abasourdi qu'elles s'avèrent appartenir à deux objets différents (intuition des objets). Et si une balle saute spontanément par-dessus un obstacle, bébé en déduit qu'il s'agit d'un être animé qui s'échappera par la droite, et il est interloqué de voir que la balle continue de sauter alors que l'obstacle a disparu (intuition de la psychologie).

Pensez-y lorsque vous désespérerez de voir votre bébé faire tomber, pour la dixième fois, sa cuillère de la table : il est en train d'expérimenter ! Comme tout scientifique, il lui faut toute une série d'expériences pour rejeter l'une après l'autre toutes les théories incorrectes selon lesquelles : (1) les objets tiennent tout seuls en l'air ; (2) il suffit qu'ils en touchent un autre pour ne pas tomber ; (3) il suffit qu'ils soient au-dessus d'un autre ; (4) la majorité de leur volume doit être au-dessus d'un autre ; et ainsi de suite.

Cette attitude expérimentale se prolonge à l'âge adulte. Nous adorons tous les gadgets qui semblent violer les lois habituelles de la physique (ballons d'hélium, mobiles en équilibre, culbuto avec un centre de gravité excentré...) et nous nous délectons des spectacles de magie (lapins qui s'évanouissent dans un chapeau, femme coupée en deux...). Si ces situations nous divertissent, c'est qu'elles violent les intuitions que nous possédons tous depuis la naissance et que nous avons raffinées dès la première année de notre vie. Josh Tenenbaum, spécialiste de l'intelligence artificielle au MIT, fait l'hypothèse que les bébés ont dans la tête un moteur physique, un modèle mental du comportement des objets semblable à celui que les logiciels de jeux vidéo utilisent pour simuler des mondes virtuels. C'est en faisant tourner ce monde dans leur tête qu'ils découvrent, très tôt, ce qui est physiquement possible ou probable.

## *Le sens du nombre*

Prenons un second exemple : l'arithmétique. Quoi de plus évident qu'un bébé ignore tout des mathématiques ? Et pourtant, dès les années 1980, des expériences démontrent le contraire<sup>38</sup> : un bébé de quelques mois à qui l'on présente régulièrement des images sur lesquelles figurent deux objets finit par se lasser... jusqu'à ce qu'on lui montre une carte avec trois objets : son attention renouvelée, il fixe plus longtemps du regard cette situation nouvelle. En manipulant la nature, la taille et la densité des objets, on peut prouver que c'est bien le nombre abstrait, le cardinal de l'ensemble, et non pas tel ou tel autre

paramètre physique, que perçoit l'enfant. La meilleure preuve est qu'on peut familiariser un bébé de quelques heures de vie avec des sons, puis passer à des images : s'il entend « tu tu tu tu », soit quatre sons, il s'intéresse davantage à une image qui comprend également quatre objets qu'à une image qui en comprend douze – et *vice versa* <sup>39</sup> . Autrement dit, il dispose, dès la naissance, de la capacité de reconnaître un nombre approximatif, d'une façon intuitive, sans avoir besoin de compter, et indépendamment du fait que l'information provienne de l'ouïe ou de la vision.

Supposez à présent qu'un enfant voie un premier objet disparaître derrière un écran, puis un second. L'écran s'abaisse ensuite – et s'il voit les deux objets attendus, l'enfant ne leur jette qu'un bref regard. Mais si, par contre, il ne découvre qu'un seul objet ou, au contraire, trois objets, sa surprise s'exprime en regardant beaucoup longtemps la scène inattendue <sup>40</sup> . Ce comportement de surprise en réaction à la violation d'un calcul mental montre que, dès quelques mois de vie, l'enfant comprend que  $1 + 1 = 2$ . Il dispose d'un modèle interne de la scène cachée, et sait le manipuler en lui ajoutant ou en lui retranchant des objets. Et cela marche pour  $1 + 1$ ,  $2 - 1$ , mais aussi pour  $5 + 5$  et  $10 - 5$ ... À condition que l'erreur soit suffisamment grande, un bébé de 9 mois détecte que l'opération  $5 + 5 = 5$  est fautive <sup>41</sup> .

S'agit-il vraiment d'une compétence innée ? Peut-être les premiers mois de vie suffisent-ils à l'enfant pour apprendre le comportement des ensembles d'objets ? S'il ne fait guère de doute que l'enfant raffine, au fil des mois, la précision avec laquelle il perçoit le nombre <sup>42</sup> , les données montrent tout aussi clairement que le point de départ n'est pas une ardoise vierge. Les nouveau-nés perçoivent les nombres dès quelques heures de vie – et il en va de même des singes, des pigeons, des corbeaux, des salamandres, des poussins et même des poissons. Chez le poussin, les scientifiques peuvent contrôler toutes les entrées sensorielles depuis l'éclosion de l'œuf : le petit poussin n'a pas vu un seul objet, et pourtant il comprend l'organisation des nombres <sup>43</sup> – ce qui montre bien que l'arithmétique fait partie des compétences innées que l'évolution confère à de nombreuses espèces, dont la nôtre. Un circuit cérébral des nombres a été identifié chez le singe, et même chez le corbeau : il contient des neurones

sensibles à des nombres spécifiques d'objets. Certains neurones préfèrent voir 1 objet, d'autres 2, 3, 5 ou même 30 – et ces cellules sont présentes même chez les animaux qui n'ont pas reçu d'entraînement spécifique<sup>44</sup>.

Au passage, ces résultats battent en brèche l'une des principales théories du développement de l'enfant, celle du grand psychologue suisse Jean Piaget. Celui-ci estimait que les bébés ignorent la « permanence de l'objet » – le fait qu'un objet continue d'exister lorsqu'on ne le voit plus – et que cette capacité n'apparaît que vers la fin de la première année. Il pensait également que le concept de nombre était hors de portée des enfants avant quelques années de vie et qu'ils devaient lentement apprendre à l'abstraire. C'est exactement l'inverse : les concepts d'objet et de nombre sont des primitives de la pensée, ils font partie du noyau de connaissances avec lesquelles nous venons au monde et qui, par leurs combinaisons, nous permettent de formuler des pensées plus complexes<sup>45</sup>.

Le sens du nombre n'est qu'un exemple de ce que j'appellerai le savoir invisible – ces intuitions précoces que nous possédons dès la naissance et qui guident tous nos apprentissages. Voici encore d'autres exemples des compétences que les chercheurs ont démontrées chez les bébés de quelques jours ou quelques mois.

## *L'intuition des probabilités*

Du nombre à la probabilité, il n'y a qu'un pas... que les chercheurs ont franchi récemment en se demandant si les bébés de quelques mois pouvaient anticiper sur le résultat d'un tirage de Loto ! Je m'explique : on leur présente une urne emplie de balles qui bougent au hasard. Il y a quatre balles : trois rouges et une verte (figure 8). À un moment donné, un panneau vient cacher l'urne... et il en sort une balle, soit verte, soit rouge. Eh bien, la surprise de l'enfant est directement liée au caractère improbable de ce qu'il voit : si c'est une balle rouge qui sort, c'est-à-dire l'événement le plus probable (puisque l'urne contient une majorité de balles de cette couleur), le bébé n'y jette qu'un bref regard...

alors que si c'est l'événement improbable qui se produit, la sortie d'une balle verte qui n'avait qu'une chance sur quatre de se produire, il la regarde bien plus longtemps. Et si l'on introduit une cloison qui bloque les balles, si l'on éloigne plus ou moins les balles de la sortie, ou si l'on fait varier le délai avant l'apparition des objets, on s'aperçoit que l'enfant intègre tous ces paramètres à son calcul de probabilités. Dans sa tête tourne une simulation probabiliste fondée sur le nombre et la nature des objets en jeu, et la durée de son regard reflète l'improbabilité de la situation observée.

Toutes ces compétences dépassent les réseaux de neurones actuels. Habitons-nous à cette idée : la réaction de surprise ne va pas de soi, ce n'est pas une réponse triviale de notre cerveau. Lorsque nous sommes surpris, c'est signe que nous pensions que cet événement avait une chance infime de se produire – et donc que notre cerveau a fait un calcul de probabilité. Or même le cerveau du bébé présente des réactions de surprise très élaborées. *Ergo* , il effectue des calculs probabilistes. C'est, actuellement, l'une des théories les plus en vogue sur le fonctionnement du cerveau : l'idée que celui-ci est un calculateur probabiliste qui manipule des distributions de probabilités et les utilise pour anticiper sur la suite probable des événements en cours. Et ce calculateur, même un bébé en est équipé.

Une série d'expériences récentes démontre qu'effectivement, le bébé possède déjà tous les mécanismes pour réaliser des inférences probabilistes complexes. Vous souvenez-vous de la théorie mathématique du révérend Bayes, cette théorie des probabilités qui permet de remonter d'une observation à ses causes probables ? Eh bien, même un bébé de quelques mois semble déjà raisonner selon la règle de Bayes<sup>46</sup> . En effet, non seulement il sait passer de l'urne aux probabilités (un raisonnement en marche avant), mais également des observations à l'urne correspondante (un raisonnement en sens inverse). Donnez-lui par exemple une urne, mais sans lui montrer son contenu. Faites entrer une personne qui, les yeux bandés, tire au hasard une série de boules. Les boules apparaissent les unes après les autres, et il se trouve que la majorité est rouge. Le bébé en déduit que l'urne doit contenir un surcroît de boules rouges, et si vous lui révélez que son contenu comprend une majorité de boules vertes, il

est surpris et regarde longtemps. Sa logique est limpide : si l'urne est surtout composée de boules vertes, comment expliquer que le tirage aléatoire soit principalement composé de boules rouges ?

Cela n'a l'air de rien, mais cela implique une extraordinaire capacité de raisonnement implicite, inconscient mais très sophistiqué, qui fonctionne dans les deux sens : étant donné un échantillon, deviner les caractéristiques de l'ensemble d'où il provient ou, au contraire, étant donné une population, deviner à quoi devrait ressembler un échantillon aléatoire.

Il existe, dès la naissance, une logique intuitive. De nombreuses variantes de cette expérience montrent à quel point l'enfant se comporte comme un scientifique en herbe capable de raisonner en bon statisticien, d'éliminer les hypothèses les moins probables et de rechercher les causes cachées des phénomènes<sup>47</sup>. Par exemple, la psychologue américaine Fay Xu a montré que, si l'enfant voit une personne, les yeux ouverts, tirer une majorité de boules rouges d'une urne qui en comprend surtout des jaunes, il est surpris, mais il en déduit que cette personne préfère les boules rouges<sup>48</sup>. Et s'il voit qu'un tirage se comporte de façon régulière avec, disons, une alternance parfaite d'une boule jaune, une rouge, une jaune, une rouge, et ainsi de suite, il en déduit que c'est un humain qui effectue le tirage, et non pas une machine<sup>49</sup>.

« La probabilité est le langage de la science », nous dit Jaynes. Les expériences les plus récentes démontrent que ce langage sans mots, cette capacité de manipuler les probabilités et de les combiner pour en tirer des conclusions raffinées, même un bébé de quelques mois la maîtrise déjà. Non seulement il n'a pas besoin de l'apprendre, mais au contraire, cette connaissance sert de fondement aux autres apprentissages : le bébé s'appuie dessus pour apprendre, parce que la logique et le sens des probabilités lui permettent de tirer les conséquences logiques des observations qu'il effectue. Il expérimente en permanence, et son cerveau de scientifique en herbe accumule les conclusions de ces recherches.

## *La connaissance des animaux et des personnes*

Si les bébés ont une bonne idée du comportement des objets, ils savent également qu'il existe une autre catégorie d'entités au comportement bien différent : les animaux et les personnes. Dans la première année de vie, ils comprennent que ces choses-là se conduisent différemment des autres : elles sont autonomes, animées d'un mouvement propre. Elles n'ont donc pas besoin d'attendre, pour se déplacer, qu'un autre objet les frappe, comme une boule de billard. Leur déplacement est motivé de l'intérieur, et non pas causé de l'extérieur.

Les bébés ne sont donc pas surpris de voir les animaux se déplacer seuls. En fait, pour eux, tout objet qui se meut de lui-même, même s'il a la forme d'un triangle ou d'un carré, se voit immédiatement affublé de l'étiquette « animal » et, dès lors, tout change. Un petit enfant sait que les êtres animés ne se déplacent pas en ligne droite, comme les autres objets, mais que leurs déplacements sont régis par des intentions et des croyances.

Prenons un exemple : faisons voir à un bébé une sphère qui se déplace en ligne droite, saute par-dessus un mur, puis continue son mouvement vers la droite (figure 8 ). Progressivement, l'enfant s'en lasse. S'est-il simplement habitué à cette trajectoire étrange ? Non, il a compris bien plus : il en a déduit que cet objet animé est doté d'une intention, celle d'aller à droite – et qu'il est particulièrement motivé, puisqu'il n'hésite pas à sauter un mur pour y parvenir. Maintenant, enlevons le mur. Dans ce cas, le bébé n'est pas surpris de voir la sphère aller droit au but, sans sauter. Par contre, il écarquille les yeux si celle-ci continue de sauter en l'air, sans raison particulière puisqu'il n'y a plus aucun mur ! En l'absence du mur, la même trajectoire qu'auparavant engendre un effet de surprise, parce que l'enfant ne comprend pas à quelle étrange intention elle pourrait bien correspondre<sup>50</sup> . D'autres expériences montrent que l'enfant infère les intentions et les préférences des personnes. Il comprend notamment que, plus le mur est haut, plus la motivation de la personne doit être grande pour qu'elle le

saute. De leurs observations, les bébés déduisent non seulement les buts et les intentions des personnes qui les entourent, mais aussi leurs croyances, leurs capacités et leurs préférences<sup>51</sup>.

Ce sens des êtres vivants ne s'arrête pas là. Très vite, vers 10 mois, les bébés attribuent une personnalité à chaque personne : s'ils voient quelqu'un jeter un enfant à terre, par exemple, ils en déduisent que c'est un individu malintentionné et se détournent de lui, au profit d'une autre personne qui, elle, a aidé un enfant à se relever<sup>52</sup>. Bien avant d'être capables de prononcer les mots « méchant » et « gentil », ils sont capables de formuler ces concepts dans leur langage de la pensée. Ce jugement est subtil : même un bébé de 9 mois distingue quelqu'un qui fait du mal volontairement d'un autre qui le fait par accident ; ou encore quelqu'un qui refuse volontairement d'aider et d'un autre qui n'en a pas la possibilité<sup>53</sup>. Comme nous le verrons plus loin, cette compétence sociale joue un rôle fondamental dans l'apprentissage. En effet, même un enfant de 1 an comprend si quelqu'un cherche à lui enseigner quelque chose. Il fait la différence entre une action banale et une action faite dans le but d'enseigner ; il possède donc, selon le psychologue hongrois Gyorgy Gergely, un sens inné de la pédagogie.

## *La perception des visages*

L'une des manifestations les plus précoces de cette compétence sociale, c'est la perception des visages. Chez l'adulte, un rien suffit à déclencher la perception d'un visage : une caricature, un smiley, un masque... Certains amateurs détectent même le visage du Christ parmi des taches dans la neige ou dans le brûlé d'un toast ! Étonnamment, cette hypersensibilité est présente dès la naissance : un bébé de quelques heures se tourne plus vite vers un smiley que vers une image similaire dont les éléments ont été chamboulés (même si l'expérimentateur s'assure que le nouveau-né n'a jamais eu la possibilité de voir le moindre visage). Une équipe a même réussi la prouesse de présenter de la



lumière à travers la paroi de l'utérus<sup>54</sup> . Résultat : trois points disposés en forme de visage (••) attirent le fœtus plus que trois points disposés en pyramide (•••). Un biais pour les visages semble déjà présent *in utero* !

Nombreux sont les chercheurs qui pensent que cette attraction quasi magnétique vers les visages joue un rôle essentiel dans le développement précoce de l'attachement – d'autant plus que l'un des symptômes les plus précoces de l'autisme consiste précisément à ne pas regarder les personnes dans les yeux. En attirant le regard vers les visages, un biais inné nous forcerait, dès quelques mois de vie, à apprendre à les reconnaître – et de fait, dès cet âge, une région du cortex visuel de l'hémisphère droit commence à répondre aux visages plus qu'à d'autres images<sup>55</sup> . La spécialisation pour les visages représente l'un des meilleurs exemples de la collaboration harmonieuse entre inné et acquis. Le bébé présente, dans ce domaine, des compétences strictement innées, mais aussi un extraordinaire instinct d'apprendre, et c'est la combinaison des deux qui lui permet, en un peu moins d'un an, d'aller au-delà de la simple réponse béate à deux yeux et une bouche, et de commencer à préférer les visages de son espèce et de son environnement immédiat<sup>56</sup> .

## *Le don des langues*

Ce n'est d'ailleurs pas seulement dans la modalité visuelle que s'exprime la compétence sociale des petits enfants, mais également par le biais de l'audition. En effet, dès la naissance, un bébé préfère écouter sa langue maternelle qu'une langue étrangère<sup>57</sup> – un résultat proprement extraordinaire, qui implique qu'un apprentissage a eu lieu pendant la grossesse. Effectivement, dès le troisième trimestre de grossesse, le fœtus entend déjà. La mélodie du langage, filtrée par la paroi utérine, passe jusqu'à lui, et il commence à la retenir. « Lorsque tes paroles de salutation sont parvenues à mes oreilles, l'enfant a tressailli d'allégresse en moi », dit Elizabeth enceinte à Marie qui lui rend visite<sup>58</sup> . L'évangéliste ne s'y est pas trompé : dès cette période, le cerveau en formation, même s'il n'est peut-

être pas encore conscient<sup>59</sup>, parvient déjà à reconnaître et à mémoriser certaines informations auditives.

Cette capacité innée est évidemment bien plus facile à étudier chez le bébé prématuré. Sorti du ventre de sa mère, il devient possible de l'équiper de mini-capteurs d'électroencéphalographie et de débit sanguin cérébral. On découvre alors que même un prématuré de six mois et demi, donc né deux mois et demi avant terme, répond déjà au langage parlé : son cerveau détecte aussi bien le changement d'une syllabe que celui de la voix<sup>60</sup>.

On a longtemps pensé que l'acquisition du langage ne démarrait pas avant 1 ou 2 ans. Pourquoi ? Parce que, comme son nom latin l'indique, l'enfant nouveau-né est *in-fans* : il ne parle pas et ne laisse donc rien paraître de ses talents d'analyste. Et pourtant, sur le plan de la compréhension du langage, son cerveau est un véritable génie des statistiques. Pour le montrer, il a fallu déployer toute la panoplie des méthodes des sciences cognitives : préférence pour tel ou tel stimulus, réponse au changement, enregistrements cérébraux. On s'est alors aperçu que sa compétence était considérable. Dès la naissance, bébé fait la différence entre la plupart des voyelles et des consonnes de toutes les langues du monde. Il les perçoit déjà comme des catégories : entre les syllabes *ba*, *da* et *ga*, même si le son varie continûment, son cerveau les traite déjà comme des catégories distinctes, séparées par des frontières qui se situent au même endroit que chez l'adulte.

Sous l'influence de l'environnement, ces compétences précoces, innées, vont se spécialiser tout au long de la première année de vie. Très vite, bébé repère que certains sons ne sont pas utilisés dans sa langue maternelle : en Angleterre pas besoin de faire la différence entre « u » et « ou » ; au Japon, la distinction entre un « r » et un « l » est inutile. En quelques mois (six pour les voyelles, douze pour les consonnes), le cerveau du bébé fait ainsi le tri entre ses hypothèses de départ et ne conserve que les phonèmes utiles aux langues qui l'entourent.

Mais ce n'est pas tout : il apprend également ses premiers mots. Comment les repère-t-il ? D'abord, il détecte les variations de prosodie, qui font que la voix du locuteur monte, descend, s'arrête... et marque ainsi les frontières des mots et des phrases. Un autre mécanisme consiste à identifier quels groupes de

sons se suivent fréquemment. Là encore, le bébé se comporte comme un statisticien en herbe. Il s'aperçoit, par exemple, que la syllabe *bi* est souvent suivie de *be*, elle-même suivie de *ron*. Un rapide calcul de probabilités lui montre que cela ne peut pas être le fait du hasard : ces syllabes se suivent avec une probabilité trop élevée, elles doivent former un mot, « bi-be-ron » – et c'est ainsi que ce mot s'ajoute au vocabulaire de l'enfant et pourra ensuite être mis en relation avec un objet précis que l'adulte est en train de nommer<sup>61</sup>. Dès 6 mois, l'enfant connaît déjà les mots qui reviennent régulièrement dans son environnement, tels que bébé, papa, maman, biberon, pied, boire, couche, etc. Ces mots vont s'imprimer dans sa mémoire à tel point qu'à l'âge adulte, ils conserveront un statut particulier, et continueront d'être traités plus efficacement que d'autres mots de sens, de sonorité et de fréquence comparables, mais acquis plus tardivement.

Très vite, la même analyse statistique permet au bébé d'identifier que certains mots reviennent plus fréquemment que d'autres : ce sont les petits mots grammaticaux comme les déterminants (le, la, des, du...) et les pronoms (je, tu, il, on...). Dès la fin de sa première année, un bébé en connaît déjà un bon paquet, et il s'en sert pour repérer les autres mots. S'il entend, par exemple, sa maman dire : « Je prépare des gnocchis », il sait en détacher les petits mots connus, « je » et « des », et, par élimination, découvrir que « prépare » et « gnocchis » sont aussi des mots. Très vite, il comprendra qu'après « le » vient souvent un nom, et après « je » un verbe – en sorte que vers 20 mois, il sera choqué si on lui dit des fausses phrases comme « je biberon » ou « le finit<sup>62</sup> ».

Bien sûr, l'analyse probabiliste ne marche pas à tous les coups, et c'est ainsi que le vocabulaire s'enrichit également de faux mots (« Un navion ; regarde le navion ! »). Néanmoins, par approximations successives, l'enfant parvient à dépasser n'importe quel algorithme actuel d'intelligence artificielle. Lorsqu'il souffle sa première bougie, il a déjà posé les bases des principales règles de sa langue maternelle, et ce à tous les niveaux, depuis les sons élémentaires (les phonèmes) jusqu'à la mélodie (la prosodie), au vocabulaire (le lexique mental) et aux règles de grammaire (la syntaxe).

Aucune autre espèce de primate ne présente de telles capacités. L'expérience a été tentée plusieurs fois : plusieurs couples de scientifiques ont adopté un bébé chimpanzé, l'ont traité comme un membre de leur famille, en lui parlant en anglais ou en langue des signes, ou encore avec des symboles visuels... pour s'apercevoir, quelques années plus tard, qu'aucun de ces animaux ne maîtrisait un langage digne de ce nom : tout au plus connaissaient-ils quelques centaines de mots<sup>63</sup>. Le linguiste Noam Chomsky a donc raison de postuler que notre espèce possède un *language acquisition device*, un système spécialisé pour l'acquisition du langage qui se déclenche automatiquement dans les premières années de la vie. Comme le disait déjà Darwin dans *L'Ascendance de l'homme*, « le langage n'est certainement pas un instinct dans le sens propre du mot, car tout langage doit être appris », mais c'est « une tendance instinctive à acquérir un art ». Ce qui est inné en nous, c'est l'instinct d'apprendre une langue, quelle qu'elle soit – un instinct tellement irrépressible que le langage apparaît spontanément, en quelques générations, chez les humains qui en sont dépourvus. Dans les communautés de sourds, la langue des signes apparaît spontanément et, dès la seconde génération, elle possède déjà des caractéristiques linguistiques universelles<sup>64</sup>.

# CHAPITRE 4

## Naissance d'un cerveau

---

« L'enfant naît avec un cerveau inachevé et non pas, comme le postulat de l'ancienne pédagogie l'affirmait, avec un cerveau inoccupé. »

Gaston BACHELARD ,  
*La Philosophie du non* (1940).

L'existence des compétences précoces du bébé réfute l'hypothèse que son cerveau n'est qu'une ardoise vierge, une table rase, une page blanche uniquement capable d'absorber l'empreinte de son environnement. Elle conduit à une prédiction simple : si l'on examine le cerveau d'un nouveau-né, on doit y observer, dès la naissance et même avant, des structures cérébrales bien organisées, correspondant à chacun des grands domaines du savoir.

Cette idée a longtemps été contestée. Jusqu'à il y a une vingtaine d'années, le cerveau du nouveau-né demeurait une *terra incognita* . L'imagerie cérébrale venait d'être inventée et elle n'avait pas encore été appliquée au cerveau en développement. La vision théorique prédominante était celle de l'empirisme, l'idée que le cerveau naît vierge de toute connaissance et subit l'empreinte de son environnement. C'est seulement avec l'avènement de méthodes raffinées

d'imagerie par résonance magnétique que l'on a enfin pu visualiser l'organisation précoce du cerveau humain et découvrir que, conformément à nos attentes, dès la naissance, pratiquement tous les circuits du cerveau adulte sont déjà présents dans celui du bébé.

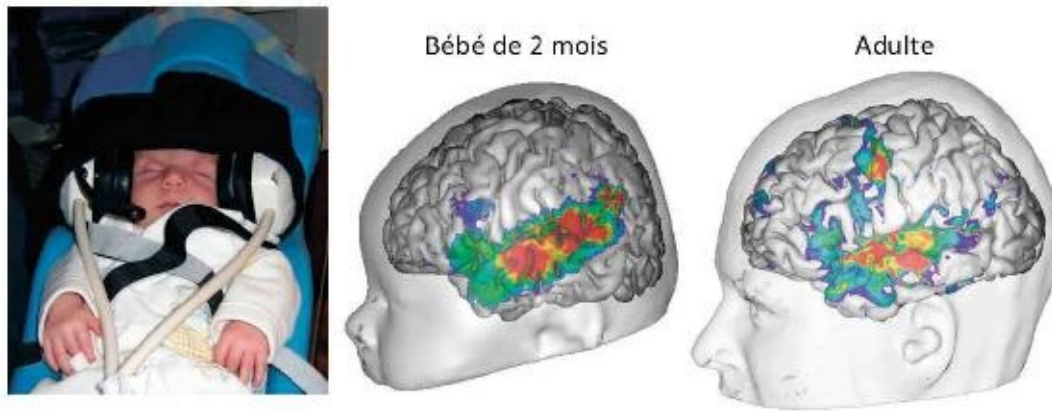
## *Le cerveau du bébé, déjà bien structuré*

Mon épouse Ghislaine Dehaene-Lambertz et moi, avec notre collègue neurologue Lucie Hertz-Pannier, avons été parmi les premiers à employer l'IRM fonctionnelle chez le bébé de 2 mois<sup>65</sup>. Nous nous appuyions, bien sûr, sur l'expérience des neuropédiatres. Avec quinze ans de recul, les médecins considéraient l'IRM comme un test anodin qu'ils n'hésitaient pas à prescrire à tout âge, y compris chez le prématuré. Cependant, leurs examens, purement anatomiques, visaient seulement à détecter des lésions précoces. Personne n'avait utilisé l'IRM fonctionnelle chez le bébé normal, dépourvu de toute pathologie, afin de voir si, dès cet âge, les circuits du cerveau pouvaient s'activer sélectivement. Pour y parvenir, nous avons dû résoudre toute une série de difficultés : protéger l'enfant du bruit de la machine par un casque antibruit, l'immobiliser en le langeant serré dans un berceau à la forme de l'IRM, le rassurer en l'habituant à cette situation inhabituelle, le surveiller en permanence au pied de la machine...

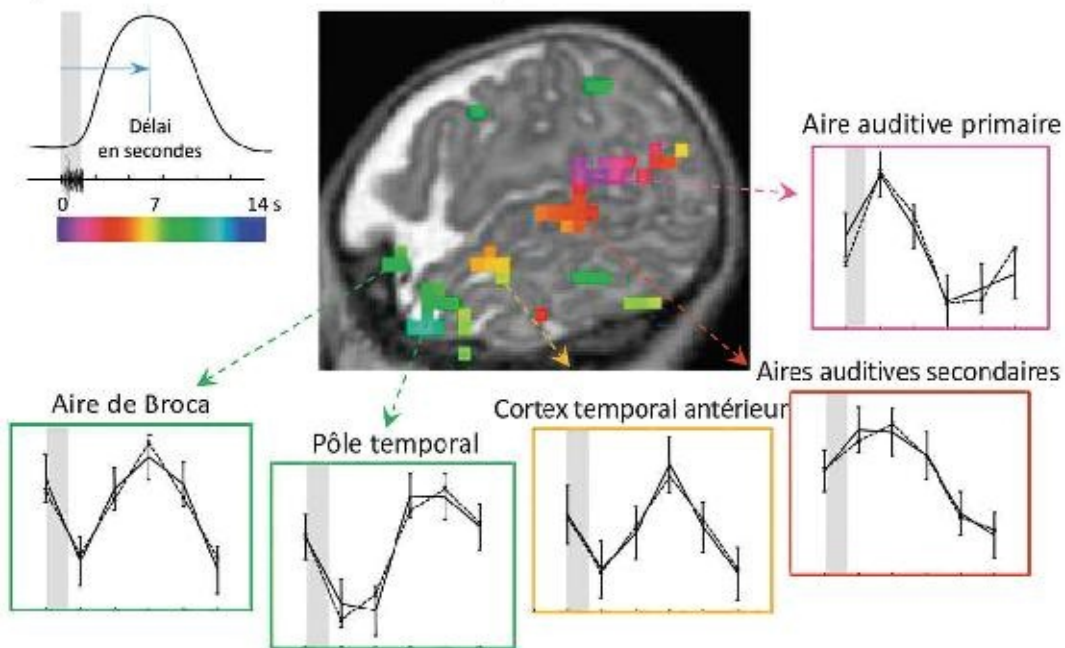
Le jeu en valait la chandelle, car les résultats furent spectaculaires. Nous avons choisi d'étudier le langage, car nous savions que les bébés l'apprennent très vite, dès la première année de vie. Et effectivement, nous avons observé qu'à 2 mois, lorsqu'un bébé entend des phrases de sa langue maternelle, il active déjà les mêmes régions du cerveau que celles d'un adulte ([figure 9](#)).

La première région à s'activer, c'est le cortex auditif primaire, le point d'entrée de toutes les informations auditives dans le cortex. Celui-ci s'allume dès

que la phrase démarre. Cela vous paraît peut-être évident, mais chez le très petit enfant, cela n'allait pas de soi. Jusqu'à récemment, certains chercheurs estimaient que les aires sensorielles sont tellement désorganisées à la naissance que les sens en sont chamboulés. Selon eux, pendant plusieurs semaines, le cerveau du bébé mélange l'audition, la vision et le toucher, et il lui faut du temps pour apprendre à séparer ces modalités sensorielles<sup>66</sup>. Nous savons aujourd'hui que c'est faux – dès la naissance, l'audition active les aires auditives, la vision les aires visuelles, le toucher les aires tactiles, sans que nous ayons besoin de l'apprendre. Cette organisation en territoires distincts pour chacune des modalités sensorielles nous est donnée par les gènes, elle se met en place chez tous les mammifères, et son origine se perd dans la nuit de notre évolution (figure 10)<sup>67</sup>.



Réponse du cerveau du bébé à une seule phrase



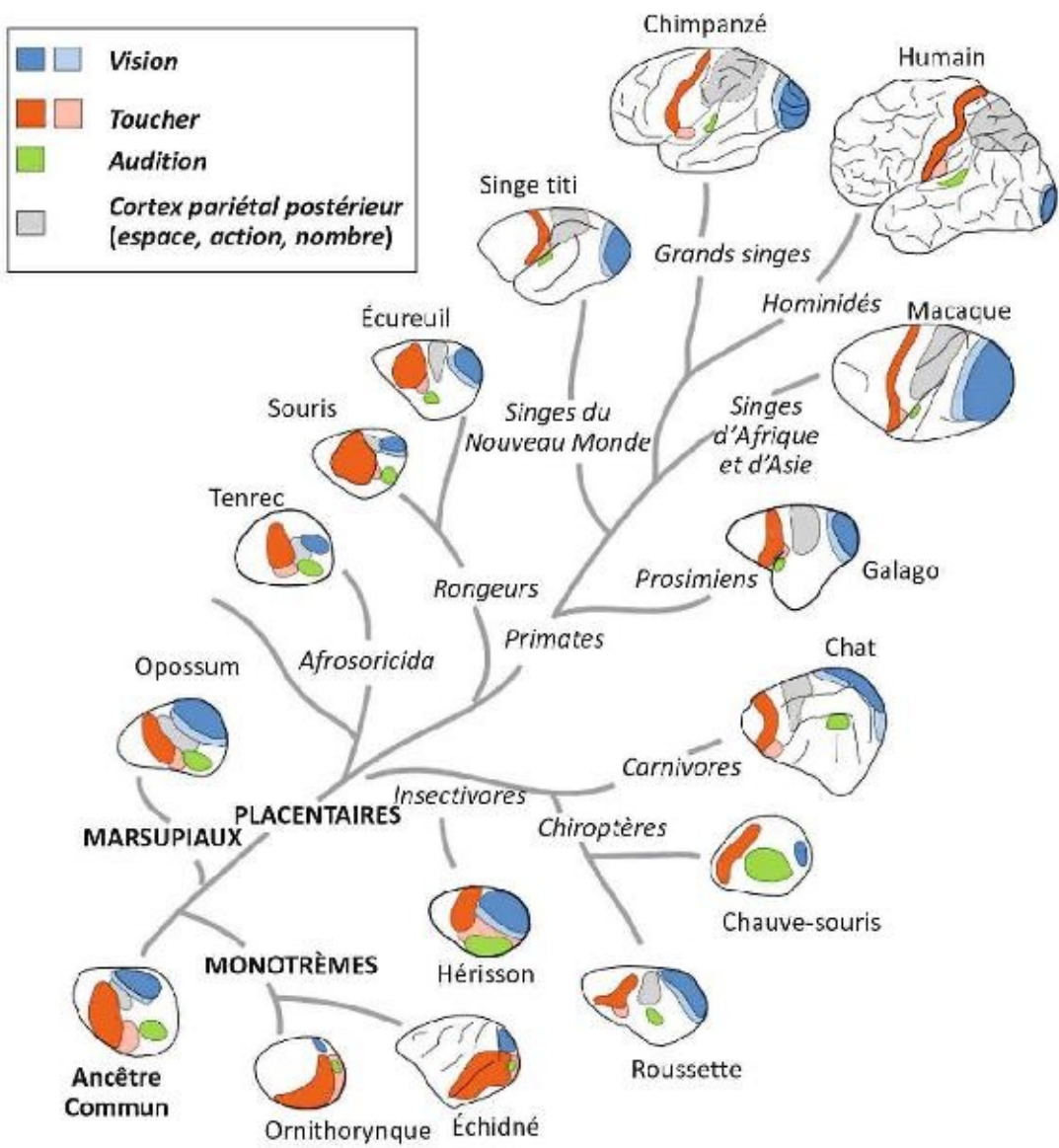
**Figure 9.** Dès la naissance, le cerveau humain applique au domaine du langage des opérations spécifiques. Lorsqu'un bébé passe une IRM et qu'on lui fait écouter des phrases de sa langue maternelle, on constate qu'un réseau bien précis de régions cérébrales s'active déjà – le même que chez l'adulte. L'activité démarre dans l'aire auditive primaire, puis s'étend progressivement aux aires temporales et frontales, dans le même ordre que chez l'adulte. Ces données réfutent l'idée d'un cerveau initialement désorganisé, simple cire vierge soumise à l'empreinte de l'environnement.

Mais revenons au bébé qui écoute une phrase. L'activité de son cerveau se propage rapidement bien au-delà de la région auditive primaire. Après une



fraction de seconde, d'autres aires s'allument, dans un ordre fixe : d'abord les régions auditives secondaires, voisines du cortex sensoriel primaire ; puis tout un chapelet de régions qui s'éloignent progressivement vers l'avant du lobe temporal ; et, enfin, l'aire de Broca, à la base du lobe frontal gauche, en même temps que la pointe du lobe temporal. Cette organisation sophistiquée, latéralisée à l'hémisphère gauche, est étonnamment semblable à celle d'un adulte. Dès l'âge de 2 mois, le bébé active la même hiérarchie d'aires cérébrales phonologiques, lexicales, syntaxiques et sémantiques qu'un adulte. Chez lui aussi, plus on avance dans la hiérarchie du cortex, plus les réponses sont lentes, et plus elles intègrent des informations de niveau de plus en plus élevé (figure 9)<sup>68</sup>.

Bien sûr, un petit bébé ne comprend pas encore les phrases qu'on lui raconte, il n'a pas encore découvert les mots ni les règles de grammaire de sa langue maternelle. Toutefois, dans son cerveau, l'information linguistique est déjà canalisée dans des circuits hautement spécialisés, les mêmes que chez l'adulte. Si un bébé apprend aussi vite à comprendre et à parler, alors que tous les autres primates en sont incapables, c'est sans doute parce que son hémisphère gauche abrite déjà une hiérarchie de régions spécialisées dans la détection de régularités statistiques à tous les niveaux : son, mot, phrase et texte.



**Figure 10.** Le cerveau humain hérite son architecture de l'évolution des mammifères. De nombreuses régions spécialisées (ici les principales aires sensorielles) partagent leur plan de base avec d'autres espèces. Elles s'établissent *in utero*, sous l'influence de nombreux gènes, et sont déjà actives au cours du troisième trimestre de la grossesse. Le cerveau humain se caractérise par des aires sensorielles proportionnellement plus petites, tandis qu'on observe une énorme expansion des régions cognitives du cortex pariétal postérieur (en gris), temporal et surtout préfrontal. Ce sont ces régions qui abritent notre langage de la pensée et nous permettent d'augmenter nos connaissances tout au long de la vie.

## *Les autoroutes du langage*

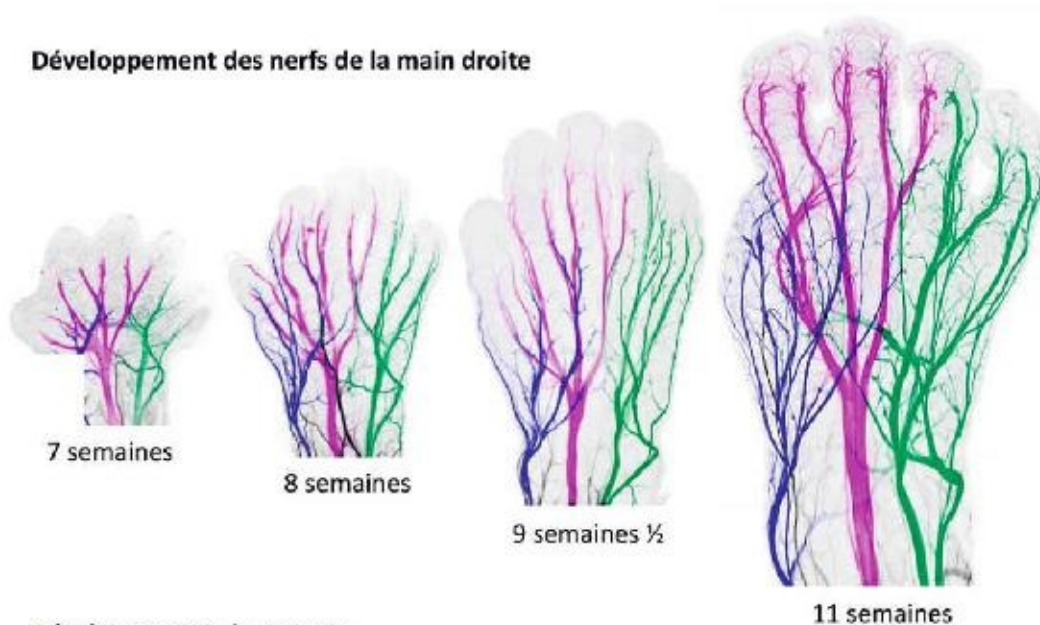
Puisque toutes ces aires cérébrales s'activent en chaîne, c'est qu'elles doivent être connectées entre elles. Chez l'adulte, on commence à comprendre quels faisceaux d'axones interconnectent les régions du langage. Il existe, en particulier, un gros câble formé de millions de fibres nerveuses, qu'on appelle le faisceau arqué, et qui relie les aires temporales et pariétales avec les aires frontales et tout particulièrement la fameuse « aire de Broca ». C'est un marqueur de l'évolution du langage. Chez 96 % des humains, il est bien plus gros dans l'hémisphère gauche, qui se consacre au langage. Cette asymétrie n'existe que dans l'espèce humaine, pas chez d'autres singes, pas même chez le chimpanzé.

Là encore, cette caractéristique anatomique n'est pas le résultat de l'apprentissage : elle est présente dès le départ. En fait, lorsqu'on examine les connexions du cerveau d'un nouveau-né, on découvre que non seulement le faisceau arqué, mais tous les grands faisceaux de connexions du cortex et des régions sous-corticales sont en place dès la naissance ([figure 11](#))<sup>69</sup>.

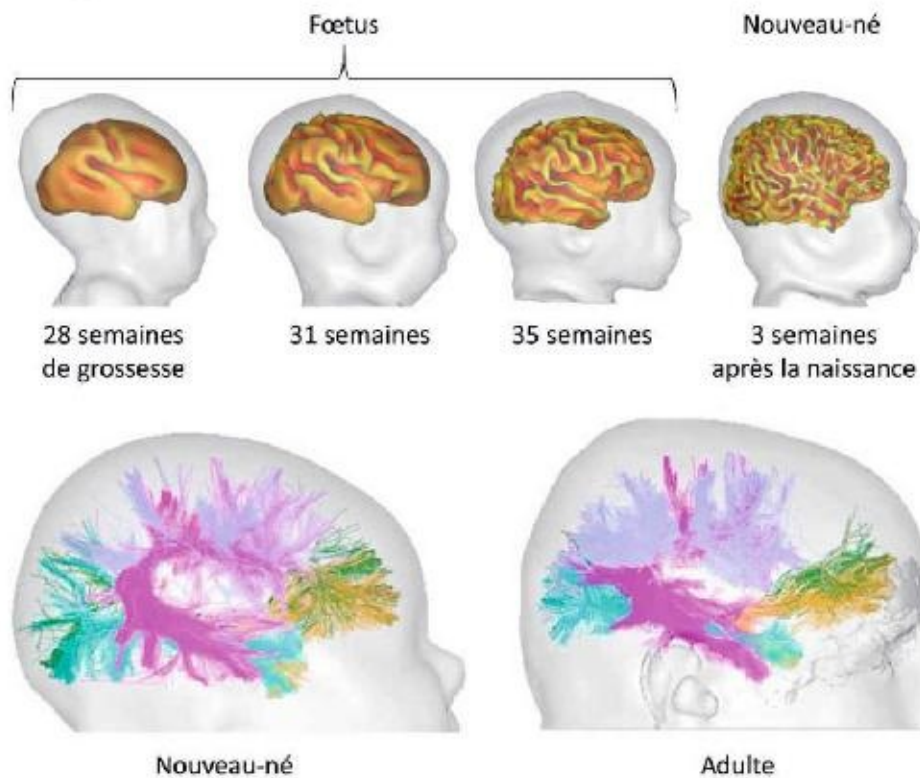
Ces « autoroutes du cerveau » se construisent au cours du troisième trimestre de grossesse. Chaque neurone excitateur en formation envoie son axone explorer les régions qui l'entourent, tel un Christophe Colomb du cerveau, parfois jusqu'à plusieurs centimètres de distance. Cette exploration est guidée et canalisée par des messages chimiques, des molécules dont la concentration varie d'une région à l'autre et qui agissent donc comme des panneaux de signalisation. La tête de

l'axone renifle littéralement cet environnement chimique d'origine génétique, et en déduit la direction qu'elle doit suivre. Ainsi se met en place, sans intervention du monde extérieur, un réseau croisé de connexions nerveuses propre à l'espèce humaine. Comme nous le verrons dans un instant, celui-ci sera raffiné ensuite par l'apprentissage – mais l'échafaudage initial est inné, il se construit *in utero* .

### Développement des nerfs de la main droite



### Développement du cerveau

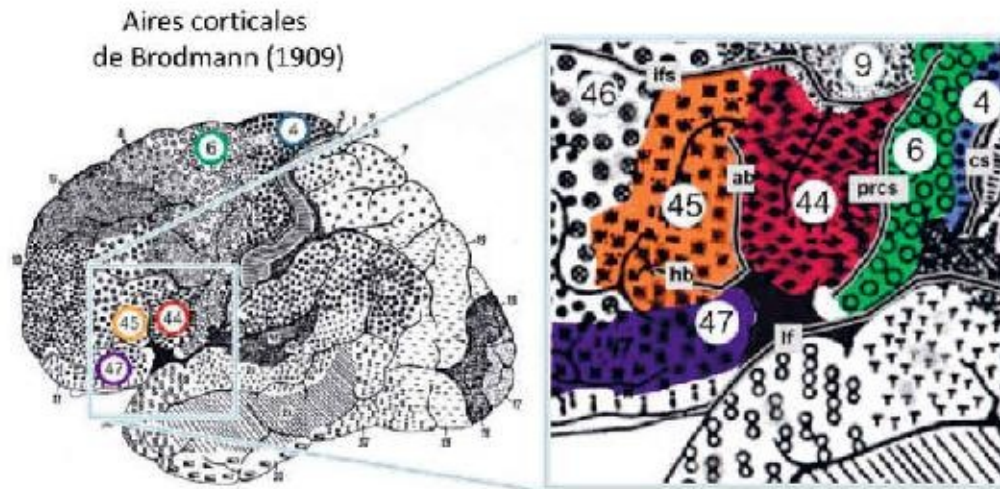


**Figure 11.** Dans les premières semaines de grossesse, l'organisme s'auto-organise sur une base génétique. Aucun apprentissage n'est nécessaire pour que les cinq doigts se forment et reçoivent une innervation spécifique. De même, l'architecture fondamentale du cerveau se met en place sans apprentissage. Dès la naissance, le cortex est déjà organisé, plissé et connecté d'une manière propre à tous les êtres humains. Les détails, cependant, sont libres de varier en fonction de

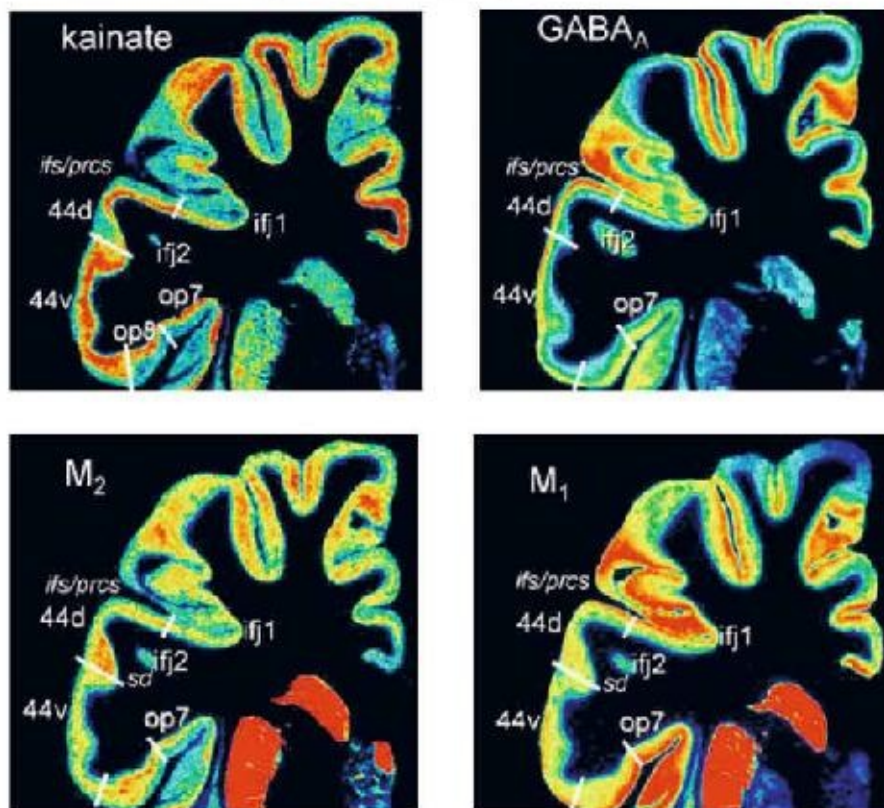
l'environnement. Dès le troisième trimestre de gestation, le cerveau du fœtus commence à s'adapter aux informations qu'il reçoit du monde extérieur.

Devrions-nous en être surpris ? Il y a une vingtaine d'années seulement, de nombreux chercheurs considéraient comme extrêmement improbable que le cerveau soit autre chose qu'une masse désorganisée de connexions aléatoires<sup>70</sup>. Ils ne pouvaient imaginer que notre ADN, qui ne contient qu'un nombre restreint de gènes, puisse abriter le plan détaillé de circuits aussi spécialisés que ceux de la vision, du langage ou de la motricité. Mais c'est mal raisonner : notre génome contient tous les détails du plan de notre corps, il sait fabriquer un cœur et ses quatre ventricules, deux yeux, vingt-quatre vertèbres, deux oreilles internes chacune avec ses trois canaux parfaitement perpendiculaires, ou encore dix doigts, leurs phalanges, et jusqu'à leur innervation d'une extrême précision... Alors pourquoi pas un cerveau doté de multiples organes internes ?

Les prouesses de l'imagerie contemporaine révèlent comment, dès les deux premiers mois de grossesse, alors que les doigts de la main ne sont encore que des bourgeons, on les voit déjà se faire envahir par trois nerfs, le radial, le médian et l'ulnaire, qui visent chacun des cibles spécifiques (figure 11)<sup>71</sup>. La même mécanique de haute précision existe dans le cerveau : de la même manière que le bourgeon de la main se scinde en cinq doigts, le cortex se subdivise en plusieurs dizaines de régions hautement spécialisées, séparées par des frontières très nettes (figure 12)<sup>72</sup>. Dès les premiers mois de grossesse, de nombreux gènes s'expriment sélectivement en différents points du cortex<sup>73</sup>. Vers vingt-huit semaines le cerveau commence à se plisser, et on voit apparaître les principaux sillons qui caractérisent le cerveau humain. Chez un fœtus de trente-cinq semaines, tous les principaux plis du cortex sont déjà bien formés, et l'on voit déjà pointer l'asymétrie caractéristique de la région temporale qui abrite les aires du langage<sup>74</sup>.



Frontières entre régions définies par quatre molécules différentes



**Figure 12.** Le cortex humain est subdivisé en régions spécialisées. Dès 1909, le neurologue allemand Korbinian Brodmann constate que la taille et la répartition des neurones varient d'une région à l'autre. Ainsi, dans la région de Broca, impliquée dans le langage, il isole trois aires (numérotées 44, 45 et 47). Ces distinctions sont aujourd'hui confirmées et raffinées par l'imagerie moléculaire. Le cortex est tapissé d'aires distinctes, dont les frontières sont marquées par des varia-

tions considérables de densité des récepteurs de neurotransmetteurs. Au cours de la grossesse, certains gènes s'expriment sélectivement dans les différentes régions du cortex et contribuent à le subdiviser en organes spécialisés.

## *L'auto-organisation du cortex*

Le plissement du cortex et la croissance de ses connexions vont se poursuivre tout au long de la grossesse. Apparaissent successivement des plissements secondaires et tertiaires : des plis sur des plis. Leur épigénèse devient progressivement de plus en plus dépendante de l'activité du système nerveux. Selon le retour que le cerveau reçoit des organes des sens, certains circuits seront stabilisés, tandis que d'autres, inutiles, vont dégénérer. Ainsi le plissement du cortex moteur n'est-il pas exactement le même chez les gauchers et chez les droitiers, tandis que les gauchers contrariés, qu'on a forcés à écrire avec la main droite, montrent une sorte de compromis : la forme de leur cortex moteur est typique d'un gaucher, mais sa taille présente l'asymétrie typique d'un droitier<sup>75</sup>. Comme le concluent les auteurs de cette étude, « la morphologie du cortex d'un adulte reflète l'accumulation de biais innés et d'expériences précoces au cours du développement ».

On pense aujourd'hui que l'apparition des plis du cortex, qui permet à chacun d'abriter des fonctions différentes, s'apparente à celle des taches sur la peau d'une girafe ou des rayures sur celle d'un zèbre : c'est un processus d'auto-organisation qui dépend à la fois des gènes et de l'environnement chimique des cellules, mais qui exige extrêmement peu d'information génétique pour se mettre en place<sup>76</sup>. Les zébrures, les taches et les cellules hexagonales apparaissent spontanément dans toutes sortes de systèmes biologiques ou physiques, à n'importe quelle échelle, depuis les empreintes digitales jusqu'au bouillonnement de l'eau dans une casserole, aux colonnes de basalte des volcans, aux plissements dans le sable d'une plage, aux dunes du désert ou aux nuages pommelés d'un ciel d'été.

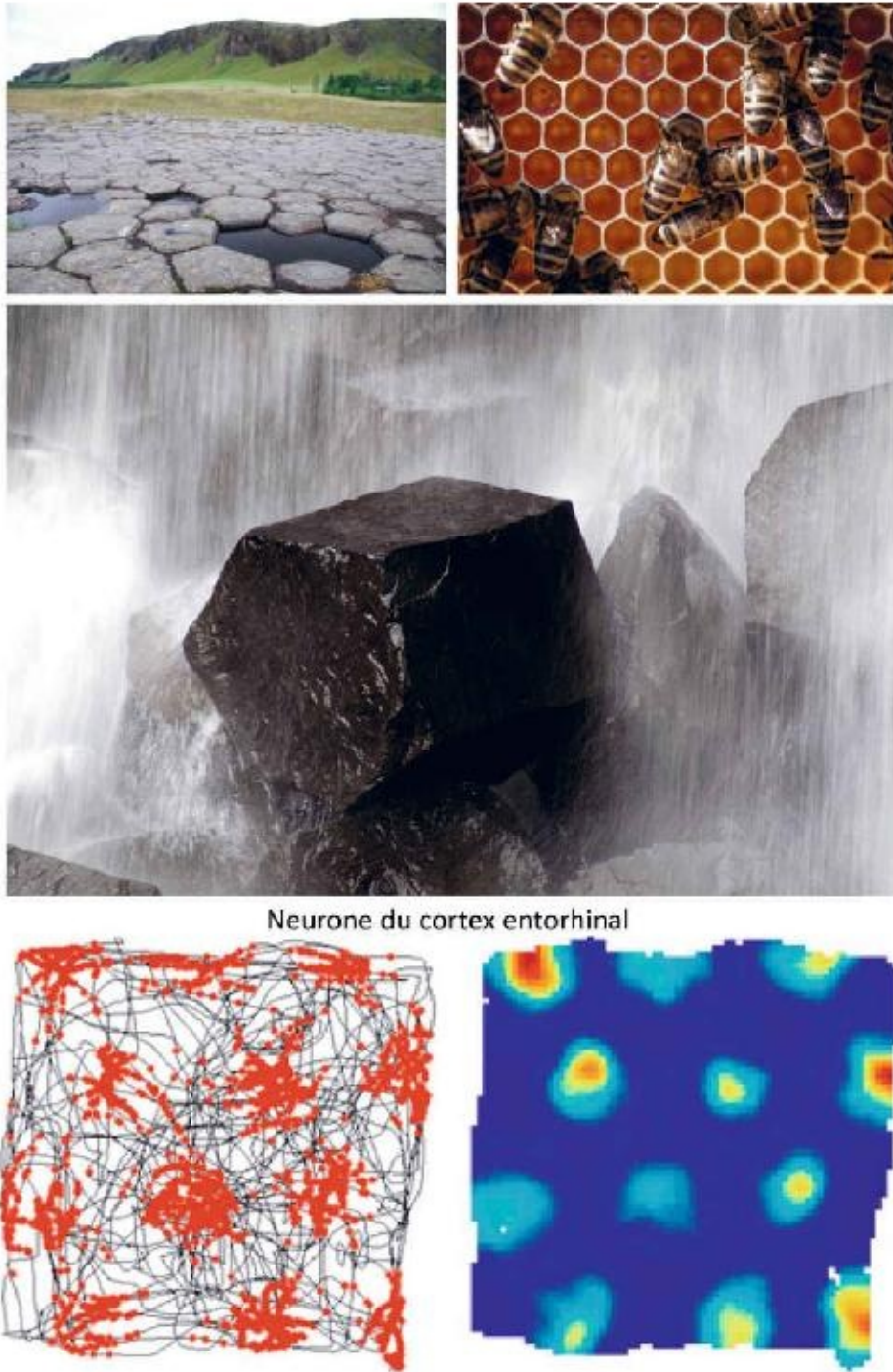


Le mathématicien Alan Turing a été le premier à l'expliquer : il suffit qu'il existe une amplification locale et une inhibition à plus grande distance. Avec le vent, le sable commence à s'accumuler en une petite bosse qui, elle-même, accroche de mieux en mieux les grains de sable, tandis que derrière elle le vent tourbillonne et arrache le sable – et en quelques heures, une dune se forme. Dès qu'il y a excitation locale et inhibition à grande distance, on voit apparaître une région dense (la dune) entourée d'une région moins dense (le creux), lui-même suivi d'une autre dune, à l'infini. Suivant les paramètres émergent spontanément des motifs en taches, en rayures ou en hexagones.

Ce mécanisme est omniprésent dans l'auto-organisation du cerveau en développement : le cortex est plein de colonnes, de rayures et de frontières qui constituent autant de modules spécialisés pour le traitement d'une information particulière. Dans le cortex visuel, par exemple, apparaît une alternance de bandes spécialisées pour les informations issues de l'œil gauche et de l'œil droit : les colonnes de dominance oculaire. Mais le même mécanisme se produit à plus haut niveau, pas forcément à la surface du cortex, mais dans un espace plus abstrait. L'un des exemples les plus spectaculaires, qui a valu le prix Nobel à ses découvreurs, est l'existence de neurones de grille (*grid cells*) dont les décharges forment des hexagones presque parfaits (figure 13). Ces neurones ont été découverts en pratiquant des enregistrements dans le cerveau d'un rat, dans une région bien précise du cerveau, le cortex entorhinal, alors que l'animal se déplace dans une grande pièce<sup>77</sup>.

On savait déjà qu'il existait, dans une région voisine (l'hippocampe), des « cellules de lieu » qui cartographiaient l'espace, c'est-à-dire qu'elles ne déchargeaient que si l'animal se trouvait à un endroit bien précis de la pièce. La découverte nouvelle fut de constater que les cellules de grille, elles, ne répondent pas seulement à un endroit de la pièce, mais à toute une série de positions différentes, qui se répartissent selon un réseau de triangles équilatéraux. Ceux-ci se regroupent pour former des hexagones, un peu comme les taches sur la peau d'une girafe ou les colonnes de basalte dans une roche volcanique ! Alors que l'animal se promène au petit bonheur, chaque cellule de grille lui indique, par ses décharges, à quel endroit il se trouve par rapport à un réseau de triangles qui

pavent tout l'espace. Le comité Nobel a appelé ce système le « GPS du cerveau », et c'est vrai qu'il fonctionne comme un système de coordonnées spatiales, une grille qui quadrille la pièce... ou plutôt qui la triangule, et qui permet à l'animal de savoir en permanence où il est (car ces neurones déchargent même dans le noir).



Neurone du cortex entorhinal

**Figure 13.** Lorsqu'un système physique s'auto-organise, qu'il s'agisse de lave ou de cire d'abeille, il n'est pas rare que s'y forment des hexagones. Le système nerveux ne fait pas exception : dans une région du cortex entorhinal, véritable GPS du cerveau, on voit apparaître des cellules de grille qui recouvrent l'espace d'un pavage hexagonal. Lorsqu'un rat explore une salle d'un mètre carré, les décharges neuronales ne surviennent que lorsque l'animal se trouve aux som-

mets d'un réseau d'hexagones. De telles cellules de grille apparaissent spontanément, un jour après que la souris a commencé à bouger : le sens de l'espace s'appuie sur un GPS pratiquement inné.

Mais pourquoi les neurones produisent-ils des triangles et des hexagones, plutôt que les rectangles de nos cartes géographiques ? Depuis Descartes, mathématiciens et cartographes s'appuient sur deux axes perpendiculaires (abscisse et ordonnée, x et y, longitude et latitude), alors pourquoi notre cerveau agit-il autrement ? Sans doute parce que les neurones s'auto-organisent au cours du développement et que les formations en hexagone apparaissent le plus naturellement du monde dès qu'un système physique se stabilise (figure 13). De fait, plusieurs modèles mathématiques inspirés des idées de Turing rendent bien compte de l'émergence des cellules de grille : cette organisation en triangles et en hexagones est un attracteur spontané de la dynamique du cortex entorhinal<sup>78</sup>.

Cette théorie d'auto-organisation des cartes du cerveau commence à être mise à l'épreuve. Des expériences extraordinaires ont montré que le GPS cérébral émerge effectivement très précocement au cours du développement. Deux groupes indépendants de chercheurs ont réussi l'exploit d'implanter des électrodes chez le tout petit rat, à peine né, avant même qu'il ne commence à marcher<sup>79</sup>. Ils ont ainsi pu examiner si l'on y trouvait déjà, dans le cortex entorhinal, des cellules de grille, et également des cellules de lieu (celles qui répondent à un endroit seulement), et encore des cellules de direction de la tête – un troisième type de cellule, dont je n'ai pas encore parlé, qui fonctionne comme le compas d'un navire : chaque neurone décharge lorsque l'animal s'oriente dans une certaine direction, par exemple nord-ouest ou plein sud. Les résultats montrent que tout ce système est pratiquement inné : les cellules de direction de la tête sont présentes dès qu'on commence à enregistrer, et les cellules de lieu et de grille émergent à peine un ou deux jours après que le raton a commencé à se déplacer. Ce n'est guère étonnant : dès qu'il se meut, il est crucial que le petit rat sache en permanence où il est, afin de retrouver le foyer où l'attend sa mère. L'évolution a donc trouvé le moyen de doter son cerveau d'un compas, d'une carte et d'une mémoire des lieux.

Qu'en est-il chez l'homme ? Nous savons, par des moyens indirects, que le cerveau de l'adulte contient aussi une carte neuronale avec une symétrie hexagonale, exactement au même endroit que chez le rat<sup>80</sup>. Nous savons aussi que l'enfant, même très petit, possède un sens de l'espace. Même un très jeune enfant aveugle parvient à s'orienter dans une pièce : si on l'emmène du point A au point B, puis au point C, il sait revenir en ligne droite de C vers A. Le petit de l'espèce humaine possède donc, comme le raton, un module mental pour la navigation spatiale<sup>81</sup>. Nous ne sommes pas encore parvenus à voir directement cette carte dans le cerveau du bébé, car il reste extraordinairement difficile d'obtenir des images du cerveau en activité à ce très jeune âge (essayez de faire une IRM à un bébé qui marche à quatre pattes...). Mais nous sommes à peu près certains de l'y trouver, dès que des méthodes d'enregistrement mobile de l'activité cérébrale seront disponibles.

Je pourrais multiplier les exemples d'autres modules spécialisés dans le cerveau du bébé. Nous savons, par exemple, que son cortex visuel contient dès quelques mois de vie (quoique pas dès la naissance) une région qui répond préférentiellement aux visages, plus qu'aux images de maisons<sup>82</sup>. C'est certainement en partie le résultat d'un apprentissage, mais étroitement canalisé, guidé, contraint par la connectivité cérébrale d'atterrir toujours au même endroit, dans un module spécifique du cortex. Nous savons également que le cortex pariétal du bébé répond déjà au nombre d'objets<sup>83</sup>, au même endroit que celui qui s'active lorsqu'un adulte effectue des calculs ou qu'un singe mémorise un nombre d'objets. Chez le singe, le neuroscientifique allemand Andreas Nieder a réussi le tour de force de montrer que cette région contient des neurones sensibles au nombre d'objets : il existe des neurones spécialisés pour un objet, d'autres pour deux objets, trois objets et ainsi de suite... Et ces neurones sont présents même si le singe n'a jamais été entraîné à effectuer une tâche numérique ; on pense donc qu'ils émergent de façon innée. Là encore, un modèle mathématique d'auto-organisation, cette fois-ci fondé sur la propagation d'une sorte d'onde à la surface du cortex, explique tous les détails de cette organisation précoce : une sorte de ligne numérique ; un espace neuronal dans

lequel les nombres 1, 2, 3, 4, *etc.* occupent des positions successives, émerge spontanément au sein d'un réseau de neurones aléatoires<sup>84</sup>.

L'idée d'auto-organisation diffère radicalement de l'approche actuelle de l'intelligence artificielle, fondée sur le *big data*. Contrairement aux réseaux de neurones actuels, notre cerveau ne semble pas avoir besoin de millions de données pour s'organiser. Bien au contraire, ses noyaux de connaissances se développent spontanément, par simulation interne. Rares sont les informaticiens qui, tel Josh Tenenbaum, professeur au MIT, envisagent de copier cette auto-organisation dans une intelligence artificielle. Tenenbaum travaille à la simulation d'un bébé virtuel qui viendrait au monde avec la capacité de générer, de lui-même, des millions d'images réalistes. Celles-ci lui serviraient alors de base d'apprentissage, sans qu'il ait besoin d'aucune autre donnée. Selon cette vision radicale, avant même la naissance, les circuits neuronaux s'entraînaient à l'aide d'une base de données qu'ils engendrent eux-mêmes<sup>85</sup> (encore un exemple du *bootstrapping* du Baron de Münchhausen). La majorité de leur apprentissage se produit en interne, sans aucune interaction avec l'extérieur, et il ne leur reste plus ensuite qu'à pratiquer quelques ajustements en fonction des données supplémentaires issues du monde réel.

La vision qui émerge de toutes ces recherches, c'est celle du pouvoir des gènes et de l'auto-organisation dans la mise en place du cerveau humain. Dès la naissance, le cortex du bébé est plissé presque comme celui d'un adulte, subdivisé en aires sensorielles et cognitives spécialisées qui sont déjà interconnectées par des faisceaux de fibres précis et reproductibles. C'est une collection de modules qui, chacun, projette sur le monde extérieur un type de représentation particulier. Le cortex entorhinal, avec ses cellules de grille, trace des plans en deux dimensions (parfaits pour coder l'espace et y naviguer). Comme nous le verrons plus loin, d'autres régions, comme le cortex pariétal, tracent des lignes (excellentes pour coder les quantités linéaires comme le nombre, la taille, ou encore le temps qui passe), tandis que la région de Broca projette des arbres enchâssés les uns dans les autres (idéal pour coder les structures du langage). Nous héritons, de notre évolution, un jeu de primitives

parmi lesquelles nous pourrons ensuite sélectionner celles qui sont les plus aptes à représenter les situations et les concepts que nous devons apprendre.

# CHAPITRE 5



# La part de l'acquis

---

Je viens d'insister sur la part d'inné dans notre cerveau, trop souvent négligée. Mais, bien sûr, cette organisation précoce ne reste pas inchangée : elle se modifie et s'enrichit sous l'effet de l'expérience. C'est l'autre face de la médaille : comment l'apprentissage modifie-t-il les circuits du cerveau de l'enfant ? Pour le comprendre, il nous faut remonter un siècle en arrière, jusqu'aux découvertes fondamentales du grand anatomiste espagnol Ramón y Cajal.

## *Qu'est-ce que la plasticité cérébrale ?*

« Tout le monde sait que les talents d'un pianiste [...] exigent des années de gymnastique musculaire et mentale. Pour comprendre ce phénomène important, il nous faut admettre qu'en plus du renforcement des voies organiques préétablies, des voies nouvelles se créent par ramification et croissance progressive des arbres dendritiques et axonaux. »

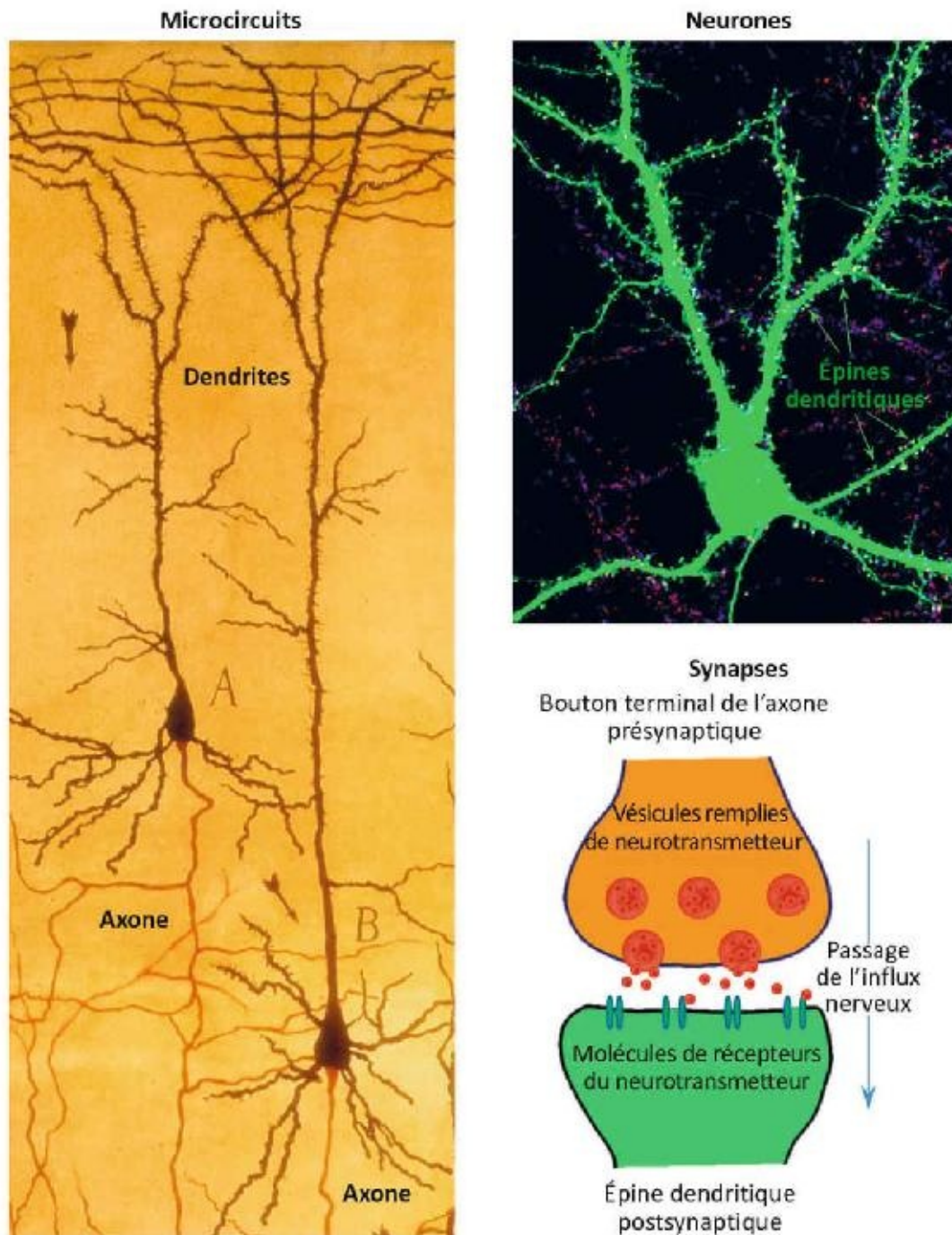
Santiago RAMÓN Y CAJAL ,  
prix Nobel de médecine (1904).

Santiago Ramón y Cajal est l'un des héros des neurosciences. Muni de son microscope, il a été le premier à cartographier la micro-organisation du cerveau. Dessinateur de génie, il en a produit des images à la fois réalistes et simplifiées

qui comptent parmi les œuvres majeures de l'illustration scientifique. Mais, surtout, il a su passer de l'observation à l'interprétation, et de l'anatomie à la fonction, avec une sûreté de jugement impressionnante. Car son microscope ne lui donnait à voir que la forme des neurones et de leurs circuits, mais il a su en tirer des déductions hardies et souvent justes sur la manière dont ils fonctionnent.

La plus grande découverte de Ramón y Cajal, c'est d'avoir compris que le cerveau est constitué de cellules nerveuses disjointes (les neurones), et non pas d'un réseau continu, un *reticulum*, comme on le pensait avant lui. Il s'est également rendu compte que ces neurones avaient une forme très particulière : contrairement aux autres cellules, comme les globules rouges par exemple, qui sont compactes et à peu près rondes, les cellules nerveuses présentent des branchements d'une incroyable complexité. Chaque neurone déploie un immense arbre constitué de plusieurs milliers de branches de plus en plus petites, les dendrites (*dendron* signifie « arbre » en grec). Les populations de neurones forment donc une forêt inextricable d'arborisations neuronales.

Cette complexité n'a pas rebuté le neurobiologiste espagnol. Dans un diagramme demeuré célèbre, partant d'un merveilleux dessin de l'anatomie du cortex, Ramón y Cajal y a ajouté une chose toute simple, et cependant lumineuse et d'une grande portée théorique : des flèches ! Celles-ci indiquent la direction dans laquelle circule l'influx nerveux : depuis les dendrites vers le corps cellulaire du neurone et, enfin, le long de l'axone ([figure 14](#)). C'était une spéculation audacieuse, mais qui s'est avérée juste. Ramón y Cajal avait compris que la forme des neurones correspond à leur fonction.



**Figure 14.** Microcircuits, neurones et synapses sont les supports matériels de la plasticité du cerveau : ils s'ajustent à chaque fois que nous apprenons. Le grand anatomiste Ramón y Cajal a, le premier, tracé les circuits du cortex et compris que les neurones sont des cellules disjointes dont les arborescences, appelées dendrites (en haut à gauche, en noir), recueillent les informations en provenance d'autres neurones et envoient un message le long de leur axone (en brun). Aujourd'hui, les neurones peuvent être visualisés avec une précision inégalée. On y discerne les épines

dendritiques, ces petits champignons qui abritent les synapses, les points de connexion entre deux neurones. Avec l'apprentissage, tous ces éléments changent : la présence, le nombre et la force des synapses, la taille des épines dendritiques, le nombre de branches des dendrites et des axones, et même l'isolation des axones par une gaine de myéline, qui détermine la vitesse de transmission de l'influx nerveux.

À l'aide de son arbre dendritique, un neurone collecte les informations en provenance des autres cellules, et il les rassemble pour n'envoyer qu'un seul message. Celui-ci est diffusé par l'axone, un tronc qui quitte le corps cellulaire et s'étend en direction d'autres circuits parfois distants de plusieurs centimètres.

Et Ramón y Cajal comprit une troisième chose : que les neurones communiquent entre eux par le biais des synapses. Il fut le premier à comprendre que les neurones constituent des cellules distinctes les unes des autres – mais son microscope lui a aussi révélé qu'elles se touchent presque en certains points. Ce sont ces zones de jonction qu'on appelle les synapses (la découverte est de Ramón y Cajal, mais le nom lui-même est dû au grand physiologiste britannique Charles Sherrington). Chaque synapse est donc le point de rencontre de deux neurones ou, plus exactement, de l'axone d'un neurone et de la dendrite d'un autre neurone. Le neurone qu'on appelle présynaptique envoie son axone très loin, jusqu'à venir rencontrer une dendrite d'un second neurone, postsynaptique, et s'y connecter.

Que se passe-t-il dans une synapse ? Un second prix Nobel, le neurophysiologiste Thomas Südhof, y a consacré toutes ses recherches, et il en a conclu que les synapses sont les unités de calcul du système nerveux – de véritables nanoprocesseurs. Souvenez-vous que notre cerveau en comprend environ 1 000 milliards de milliards : c'est une machinerie d'une complexité inégalée. Je ne peux qu'en résumer ici les aspects les plus simples ([figure 14](#)). Le message qui voyage dans l'axone est électrique, mais la plupart des synapses le transforment en un message chimique. Au niveau de la synapse, l'axone se termine par un bouton qui contient des vésicules, des sortes de poches remplies d'une molécule qu'on appelle le neurotransmetteur (du glutamate, par exemple). Lorsque le signal électrique parvient jusqu'au bouton terminal, ces poches

s'ouvrent, et la molécule se déverse dans l'espace synaptique qui sépare les deux neurones. Et c'est pour cela qu'on l'appelle un neurotransmetteur : cette molécule transmet le message d'un neurone à l'autre. En effet, la molécule libérée s'attache à la membrane du second neurone en des points particuliers qu'on appelle des récepteurs. Le neurotransmetteur agit sur le récepteur comme une clé dans une serrure : il ouvre littéralement un canal dans la membrane postsynaptique. S'y déversent des ions, des atomes chargés positivement ou négativement. Le flux de ces ions engendre donc un courant électrique au sein de la cellule cible. La boucle est bouclée : le message est passé d'électrique à chimique, puis de chimique à électrique, et il a franchi l'espace entre les deux neurones.

Quel rapport, me direz-vous, avec l'apprentissage ? Eh bien nos synapses se modifient en permanence, et ces modifications reflètent ce que nous apprenons<sup>86</sup>. Chaque synapse est une petite usine chimique, et de très nombreux éléments de cette usine peuvent se modifier : le nombre de vésicules, leur taille, le nombre de récepteurs, leur efficacité, la taille et même la forme de la synapse... tous ces paramètres affectent la force avec laquelle le message électrique présynaptique sera transmis au neurone postsynaptique. Et ces changements ne surviennent pas au hasard : ils stabilisent l'activité des neurones en renforçant leur capacité de s'activer l'un l'autre s'ils l'ont déjà fait par le passé.

La règle fondamentale est si simple qu'elle a été devinée dès 1949 par le psychologue Donald Hebb. On peut la résumer par une formule simple : « *Neurons that fire together wire together* » ou, en français, « se coactiver, c'est se connecter ». Autrement dit, lorsque deux neurones s'activent au même moment, leurs interconnexions se renforcent. Plus exactement, si le neurone émetteur, donc présynaptique, décharge et que, quelques millisecondes plus tard, le neurone récepteur, postsynaptique, entre également en activité, alors la synapse se renforce : à l'avenir, la transmission entre ces deux neurones sera encore plus efficace. Si, par contre, la synapse échoue à se faire entendre, en sorte que le neurone postsynaptique ne décharge pas, alors la synapse s'affaiblit.

On comprend pourquoi ce phénomène stabilise l'activité nerveuse : il renforce les circuits qui ont bien fonctionné par le passé, pour que la même

activité ait plus de chances de se reproduire. Grâce à la plasticité synaptique, de vastes tapisseries neuronales, chacune composées de millions de neurones, se succèdent dans un ordre toujours plus précis et reproductible. La souris qui parcourt chaque couloir de son labyrinthe dans l'ordre exact, la violoniste qui fait jaillir de ses doigts une fontaine de notes, ou l'enfant qui récite une poésie réveillent une partition neuronale dont chaque note, chaque transition est enregistrée dans quelques centaines de millions de synapses.

Bien entendu, le cerveau n'enregistre pas, sans discrimination, tous les événements de notre vie. Seuls les épisodes qu'il juge les plus importants s'impriment dans nos synapses. C'est pourquoi la plasticité synaptique est modulée par les grands réseaux de neurotransmetteurs, particulièrement l'acétylcholine, la dopamine et la sérotonine, qui déterminent les épisodes importants à apprendre. La dopamine, par exemple, est le neurotransmetteur associé à toutes les récompenses : nourriture, sexe, drogue... et, au cas où vous vous poseriez la question : oui, même le rock'n' roll<sup>87</sup> ! Les décharges du circuit de la dopamine ponctuent tout ce que nous aimons, tout ce à quoi nous sommes « accros », et signalent au reste du cerveau que l'épisode en cours est positif, inattendu, qu'il a tourné mieux que prévu. L'acétylcholine, elle, s'attache plus généralement à tous les moments importants. Leurs effets sont massifs. Si vous vous souvenez si bien de ce que vous faisiez le 11 septembre 2001, lorsque vous avez appris l'attentat du World Trade Center, c'est qu'un ouragan de neurotransmetteurs a balayé vos circuits cérébraux ce jour-là, notamment dans l'amygdale, siège des émotions fortes, et que vos synapses en ont été massivement altérées.

La capacité des synapses à se modifier en fonction de l'activité de leurs neurones pré-et postsynaptiques a été découverte dans ces conditions artificielles : il fallait tétaniser les neurones en les stimulant à un rythme effréné avec un fort courant électrique pour que la force de leurs synapses change. Ce mécanisme ressemblait certes à une mémoire, car la synapse restait modifiée pendant plusieurs heures, un phénomène intitulé « potentiation à long terme<sup>88</sup> ». Mais était-il vraiment utilisé par l'organisme pour stocker des informations ? Les premières preuves sont venues d'un animal marin, la limace de mer (*Aplysia*

*californica* ), aux neurones gigantesques. Dans ses ganglions nerveux – car cet invertébré primitif n’a pas vraiment de cerveau –, un troisième prix Nobel, Eric Kandel, a identifié toute une cascade de modifications synaptiques et moléculaires lorsque l’animal est conditionné un peu comme le chien de Pavlov<sup>89</sup> .

Ensuite, au fur et à mesure que les techniques d’enregistrement et de visualisation des synapses progressaient, les preuves que la plasticité synaptique est responsable de l’apprentissage se sont accumulées. Les modifications synaptiques se produisent précisément dans les circuits que l’animal utilise pour apprendre. Lorsqu’une souris apprend à éviter un lieu parce qu’elle y a reçu une petite décharge électrique, on voit changer les synapses de l’hippocampe, une région responsable du sens de l’espace et de la mémoire épisodique<sup>90</sup> : les connexions entre l’hippocampe et l’amygdale enregistrent l’épisode traumatisant. Mais si la souris est terrifiée par un son, ce sont les synapses qui relient l’amygdale au cortex auditif qui se modifient<sup>91</sup> . Et ces changements ne sont pas simplement concomitants à l’apprentissage : ils semblent jouer un rôle causal car si, dans les minutes qui suivent un événement, on interfère avec les mécanismes moléculaires qui permettent aux synapses de se modifier, l’animal ne le retient pas en mémoire<sup>92</sup> .

## *Le portrait d’un souvenir*

Qu’est-ce donc qu’un souvenir ? Et quelle est sa base matérielle dans le cerveau ? La plupart des chercheurs s’accordent sur le scénario suivant, qui distingue des périodes d’encodage et de rappel<sup>93</sup> .

Commençons par l’encodage. Chacune de nos perceptions, de nos actions ou de nos pensées repose sur l’entrée en activité d’un sous-ensemble bien précis de neurones (tandis que les autres restent inactifs, voire s’inhibent). C’est l’identité de ces neurones actifs, distribués dans de nombreuses régions du cerveau, qui définit le contenu de la pensée. Lorsque je vois, disons, Emmanuel Macron à

l'Élysée, certains neurones répondent à son visage (dans la région temporale inférieure), d'autres à sa voix (dans la région temporale supérieure), d'autres à l'agencement du bureau présidentiel (dans la région parahippocampique), et ainsi de suite : rien n'est codé par un neurone unique, mais toujours par des groupes de neurones délimités et interconnectés. Si je rencontre ma fromagère au marché, c'est l'entrée en activité de groupes de neurones légèrement différents qui me permet, en principe, de ne pas la confondre avec le président.

Supposons maintenant que ces épisodes soient jugés suffisamment marquants par les systèmes émotionnels pour valoir la peine d'être stockés en mémoire. Comment s'y enregistrent-ils ? Les neurones qui forment une assemblée fortement activée se modifient : ils expriment de nouveaux gènes, modifient la force de leurs connexions, parfois en forment de nouvelles. Ces changements forment la base physique de l'apprentissage : collectivement, ils constituent le substrat de la mémoire.

Les neurones peuvent à présent s'éteindre : le souvenir, lui, restera dormant, inconscient mais bien réel, inscrits dans l'anatomie même de nos circuits neuronaux. Plus tard, grâce à ces changements synaptiques, il suffira d'un indice (disons, la photo de l'Élysée) pour que le cerveau parvienne à recréer, de proche en proche, une configuration d'activité proche de celle de départ, et se souvienne du visage d'Emmanuel Macron. C'est le rappel : selon cette théorie, se rappeler, c'est rejouer, dans son cerveau, les décharges neuronales associées à une expérience passée.

La mémoire n'est donc pas le fait d'une seule région du cerveau – elle est partout, car chaque circuit est capable de se modifier en réponse à une configuration fréquente d'activité neuronale. Mais tous les circuits ne jouent pas le même rôle. Même si la terminologie reste floue et continue d'évoluer, les chercheurs distinguent au moins quatre sortes de mémoire.

- La mémoire de travail, ou mémoire à court terme, conserve, pendant un très court moment, une représentation mentale à l'état actif. Elle repose principalement sur la décharge vigoureuse de nombreux neurones dans le cortex pariétal et préfrontal, qui soutiennent à leur tour les neurones



d'autres régions plus périphériques<sup>94</sup>. C'est elle qui nous permet de garder un numéro de téléphone en tête : pendant le temps qu'il nous faut pour le taper dans notre smartphone, les neurones se soutiennent l'un l'autre et conservent ainsi l'information à l'état actif. Même si cette mémoire implique probablement de brefs changements synaptiques<sup>95</sup>, elle ne dure pas plus de quelques secondes : dès que nous sommes distraits, l'assemblée de neurones actifs s'évanouit.

- La mémoire épisodique : c'est l'hippocampe, une structure située dans la profondeur des hémisphères cérébraux, en deçà du cortex, qui enregistre les épisodes forts ou moins forts de notre vie quotidienne. Les neurones de l'hippocampe semblent mémoriser le contexte de chaque événement : ils codent où, quand, comment, avec qui les choses se sont produites. Ils enregistrent l'épisode en question par des changements synaptiques, en sorte que nous pouvons nous les rappeler plus tard. Le célèbre patient H. M., dont les hippocampes des deux hémisphères avaient été détruits par une intervention chirurgicale, n'enregistrait plus rien : il vivait dans un éternel présent, incapable d'ajouter le moindre enregistrement nouveau à sa biographie mentale. Des données récentes suggèrent que l'hippocampe intervient dans toutes sortes d'apprentissages rapides, dès lors que l'information apprise est unique : les neurones de l'hippocampe attribueraient une séquence de décharges spécifiques à chaque événement ou découverte digne d'intérêt<sup>96</sup>.
- La mémoire sémantique : les souvenirs ne restent pas toujours dans l'hippocampe. Pendant la nuit, le cerveau les rejoue et les déplace ailleurs, dans le cortex. Là, ils se transforment en connaissances permanentes : notre cerveau généralise, il extrait l'information présente dans les épisodes que nous avons vécus et l'intègre dans notre vaste bibliothèque de connaissances du monde. Au bout de quelques jours, nous nous souvenons du nom du président de la République sans conserver le moindre souvenir du lieu ou du moment où nous l'avons entendu pour la première fois : d'épisodique, la mémoire est devenue sémantique. L'épisode unique est

devenu connaissance durable, et le code neural s'est déplacé de l'hippocampe vers le cortex<sup>97</sup> .

- La mémoire procédurale : lorsque nous répétons maintes et maintes fois la même activité (lacer nos chaussures, réciter un poème, calculer, jongler, jouer du violon, faire du vélo...), les neurones du cortex et des circuits sous-corticaux finissent par se modifier afin que l'information circule mieux. Plus rapide, plus reproductible, élaguée de toute activité parasite, l'activité nerveuse s'enchaîne avec la précision d'un mécanisme d'horlogerie. C'est la mémoire procédurale : l'enregistrement, inconscient, des configurations d'activités routinières que nous utilisons régulièrement. Ici, l'hippocampe ne joue plus aucun rôle, car c'est la répétition du geste qui transfère la mémoire vers un support implicite. C'est pourquoi le patient H. M., même dépourvu de tout souvenir conscient, pouvait encore apprendre des procédures nouvelles. Les chercheurs lui ont même appris à écrire à l'envers, en regardant sa main dans un miroir : n'ayant aucun souvenir de s'être entraîné, il était tout surpris de se découvrir aussi compétent !

## *Vraies synapses et faux souvenirs*

Aujourd'hui, la maîtrise des chercheurs sur les circuits de la mémoire est telle que nous ne sommes plus très loin du rêve de Michel Gondry dans le film *Eternal Sunshine of the Spotless Mind* : effacer sélectivement du canevas de notre mémoire les souvenirs qui nous empoisonnent la vie, comme le stress post-traumatique des soldats de la guerre du Vietnam ou, au contraire, y peindre la toile illusoire d'un faux souvenir.

Les deux manipulations ont déjà été réalisées chez la souris. L'équipe d'un quatrième prix Nobel, le professeur Susumu Tonegawa, lui a d'abord fait vivre un bref traumatisme : placée dans une pièce, la souris reçoit de petits chocs

électriques déplaisants. Il est facile de s'assurer que cet épisode est bien imprimé en mémoire : la souris évite le lieu où il s'est produit.

On peut alors jouer avec ce souvenir. L'enregistrer, tout d'abord. Grâce à un microscope sophistiqué à deux photons, on parvient à visualiser quels neurones sont activés, et l'on constate que dans l'hippocampe des neurones différents s'activent pour la pièce A, associée au choc électrique, et pour la pièce B, où rien ne s'est produit. Le manipuler, ensuite : pendant que l'animal est dans la pièce A, on lui envoie des chocs tout en activant, dans son cerveau, la population de neurones qui code pour la pièce B. Revenue dans la pièce B, la souris s'alarme et se paralyse de frayeur : tout se passe comme si le mauvais souvenir était à présent attaché à la pièce B, où rien ne s'est pourtant jamais produit<sup>98</sup>.

L'équipe de Susumu Tonegawa s'est ensuite attachée à transformer le mauvais souvenir en un bon : en réactivant les neurones tandis que les souris étaient mises en présence de partenaires de l'autre sexe – un bon moment garanti –, les chercheurs sont parvenus à effacer l'association avec le choc électrique. Les souris, loin d'éviter la pièce maudite, se sont mises à l'explorer comme si elles y recherchaient les partenaires dont elles se souvenaient<sup>99</sup>.

Peut-on, enfin, effacer un souvenir traumatique ? Oui, il suffit de le réévoquer tout en affaiblissant, au même moment, les synapses excitées par ce rappel en mémoire. Dans les jours qui suivent, la souris ne montre plus la moindre réminiscence de ce mauvais moment<sup>100</sup>.

Dans la même veine, le chercheur français Karim Benchenane a réussi le tour de force d'imprimer une nouvelle mémoire dans le cerveau de la souris pendant son sommeil. Il s'est servi du fait que, lorsque nous dormons, notre hippocampe réactive une trace des souvenirs de la veille, et notamment des lieux où nous sommes passés (nous y reviendrons au [chapitre 10](#)). Chaque fois que le cerveau de la souris endormie réévoquait un endroit particulier de son enclos, le chercheur lui donnait une petite injection de dopamine, le neurotransmetteur de la récompense. Résultat : dès le réveil, la souris se précipitait vers cet emplacement ! Ce lieu anodin avait acquis, dans sa mémoire, la douceur de la Provence ou du Paris de nos premières amours.

Plus proches de nous, des expériences ont commencé à imiter, chez l'animal, les effets des apprentissages scolaires. Que se passe-t-il lorsqu'un singe apprend des lettres, des chiffres, ou l'usage d'un outil<sup>101</sup> ? Le chercheur japonais Atsushi Iriki a montré qu'un singe pouvait apprendre à se servir d'un râteau pour rapprocher de lui des morceaux de nourriture trop éloignés pour être saisis de la main. Après quelques milliers d'essais, l'animal devient aussi expert qu'un croupier : il ne lui faut plus que quelques dixièmes de seconde pour ramener l'objet à lui en tirant le râteau d'un geste vif. Le singe parvient même à utiliser un premier râteau, de taille moyenne, pour en tirer un second, plus long, qui lui permette enfin d'atteindre un point éloigné ! Cet apprentissage déclenche toute une cascade de changements cérébraux. La consommation d'énergie augmente dans une région bien précise du cortex, la région pariétale antérieure – la même aire que celle qu'un être humain utilise pour contrôler les mouvements de la main, écrire, attraper un objet ou utiliser un marteau ou une pince. De nouveaux gènes s'y expriment, les synapses y bourgeonnent, les arborescences dendritiques et axonales s'y multiplient – et tous ces branchements supplémentaires se traduisent, chez l'animal expert, par une augmentation de 23 % de l'épaisseur du cortex. On voit aussi les faisceaux de connexions s'altérer : les axones en provenance d'une région lointaine, à la jonction avec le cortex temporal, s'épanouissent jusqu'à grandir de plusieurs millimètres et envahir une partie de la région pariétale antérieure qui, auparavant, n'était pas innervée par ces neurones.

On voit bien, dans ces exemples, à quel point les effets de la plasticité cérébrale s'étendent dans le temps et dans l'espace. Retraçons-en ensemble les grandes lignes. Un ensemble de neurones, qui code pour un événement ou un concept que nous souhaitons mémoriser, s'active dans notre cerveau. Comment ce souvenir s'enregistre-t-il ? Au commencement est la synapse : microscopique point de contact entre deux neurones, c'est elle qui se renforce lorsque les neurones qu'elle relie s'activent de conserve – c'est la fameuse règle de Hebb : se coactiver, c'est se connecter. Une synapse qui se renforce est comme une usine qui augmente sa productivité : elle recrute plus de neurotransmetteurs, du

côté présynaptique, et plus de molécules de récepteurs, du côté postsynaptique. Elle augmente également sa surface, afin de les accueillir.

Avec un surcroît d'apprentissage, c'est la forme même du neurone qui change. Une sorte de champignon, qu'on appelle une épine dendritique, se forme à l'endroit où la dendrite accueille la synapse. Le cas échéant, une seconde synapse vient doubler la première. D'autres synapses distantes, mais qui atterrissent sur le même neurone, se renforcent également<sup>102</sup>.

Ainsi, lorsque l'apprentissage se prolonge, l'anatomie même du cerveau finit par s'altérer. Avec les avancées de la microscopie, et notamment la révolution apportée par les nouveaux microscopes à deux photons, fondés sur le laser et la physique quantique, on peut directement voir pousser les boutons synaptiques et, de l'autre côté de la synapse, les bourgeons axonaux, exactement comme un arbre au printemps, à chaque épisode d'apprentissage. Accumulés, ces changements peuvent être substantiels, de l'ordre du millimètre, et ils commencent à devenir détectables chez l'homme grâce à l'imagerie par résonance magnétique. Apprendre à jouer de la musique<sup>103</sup>, à lire<sup>104</sup>, à jongler<sup>105</sup> ou même à conduire un taxi dans une grande ville<sup>106</sup> se traduit par des changements détectables de l'épaisseur du cortex ou de la force des connexions qui les relie : les autoroutes du cerveau s'améliorent au fur et à mesure qu'on les utilise.

Építome de l'apprentissage, la synapse n'en est pas pour autant le seul mécanisme. Lorsque nous apprenons, le bouillonnement synaptique force les neurones à engendrer de nouveaux branchements, tant du côté de l'axone que des dendrites. À distance de la synapse, les axones s'entourent d'une gaine d'isolant, la myéline, qui ressemble à l'adhésif qu'on enroule autour des fils électriques pour les isoler. Plus un axone est utilisé, plus cette gaine fait de tours. Elle isole ainsi de mieux en mieux l'axone, ce qui lui permet de transmettre les informations à une vitesse supérieure.

Les neurones ne sont donc pas seuls. Après un long apprentissage, tout leur environnement change également : leur entourage de cellules gliales, qui les nourrissent et les soignent, et même le réseau vasculaire de veines et d'artères

qui leur apportent oxygène, glucose et nutriments. À ce stade, c'est tout un circuit qui s'est transformé.

Certains chercheurs contestent d'ailleurs le dogme qui fait de la synapse l'acteur indispensable de tous les apprentissages. Des données récentes suggèrent que les cellules de Purkinje, qui sont des neurones spéciaux du cervelet, peuvent apprendre à produire un délai d'une fraction de seconde, sans que cet apprentissage n'exige la moindre modification de leurs synapses<sup>107</sup>. La dimension de temps, spécialité du cervelet, semble reposer sur un mécanisme interne à la cellule, et non synaptique. Chaque neurone garderait trace d'un ou plusieurs intervalles de temps, peut-être même par le biais de modifications chimiques stables de l'ADN.

Nous ne comprenons d'ailleurs pas comment ces changements implémentent les apprentissages les plus élaborés, fondés sur le « langage de la pensée », cette combinatoire de concepts propre à l'espèce humaine. Comme nous l'avons vu, les réseaux de neurones conventionnels expliquent comment des millions de synapses, en se modifiant, nous permettent d'apprendre à reconnaître un chiffre, un objet ou un visage. Par contre, il n'existe aucun modèle vraiment satisfaisant des réseaux de neurones qui sous-tendent l'acquisition du langage ou des règles mathématiques. Passer de la synapse aux règles symboliques que nous apprenons en cours de maths reste, à l'heure actuelle, une gageure. Gardons l'esprit ouvert, car nous sommes très loin d'avoir parfaitement compris le code par lequel notre cerveau parvient à enregistrer nos acquis et nos souvenirs.

## *La nutrition, élément clé de l'apprentissage*

Ce qui est certain, c'est qu'avec l'apprentissage des changements massifs se produisent : non seulement les neurones se modifient, avec leurs échafaudages de dendrites et d'axones, mais leur entourage de cellules gliales change

également. Toutes ces transformations exigent du temps. Chaque épisode d'apprentissage met en œuvre une cascade de modifications biologiques qui dure plusieurs jours. De nombreux gènes spécialisés dans la plasticité doivent s'exprimer afin que les cellules produisent l'échafaudage de protéines et de membranes qui vont consolider les synapses, les dendrites et les axones en construction. Et cela absorbe également de l'énergie, beaucoup d'énergie – la consommation du cerveau du jeune enfant représente jusqu'à 50 % du bilan énergétique du corps. Glucose, oxygène, vitamines, fer, iode, acides gras, nutriments les plus divers : tous ces ingrédients sont essentiels au cocktail d'une enfance réussie. Le cerveau ne se nourrit pas que de stimulation intellectuelle. Faire et défaire quelques millions de synapses par seconde exige une alimentation équilibrée, une bonne oxygénation et de l'exercice physique<sup>108</sup>.

Un triste épisode illustre l'extrême sensibilité du cerveau en développement. En novembre 2003, un désastre méconnu a frappé les enfants d'Israël<sup>109</sup>. Du jour au lendemain, les hôpitaux pédiatriques de tout le pays ont soudain vu affluer des dizaines de bébés atteints de graves symptômes neurologiques : léthargie, vomissements, troubles du regard et de la vigilance allant jusqu'au coma et, pour deux d'entre eux, au décès. Une course contre la montre s'engagea : quelle était la source de cette maladie jusqu'alors inconnue ? L'enquête finit par montrer que tous étaient nourris au biberon avec le même lait en poudre à base de soja.

L'analyse de ce lait confirma les pires craintes. D'après l'étiquette, le lait, conformément aux normes, aurait dû comprendre 385 milligrammes de thiamine, plus connue sous le nom de vitamine B1. En réalité, il n'en contenait pas la moindre trace. Contacté, le fabricant admit qu'il avait altéré la composition de son lait au début de l'année 2003 : pour des raisons économiques, il avait cessé d'y ajouter de la thiamine. Or celle-ci est un nutriment essentiel du cerveau. Pire, l'organisme ne stocke pas la thiamine, et tout déséquilibre alimentaire créé rapidement une grave carence.

Les neurologues savaient déjà que la carence en thiamine, chez l'adulte, entraîne des troubles neurologiques sévères, un tableau qu'on appelle syndrome de Wernicke-Korsakoff et que l'on rencontre la plupart du temps chez les grands

alcooliques. À la phase aiguë, cette déficience induit une encéphalopathie de Wernicke, maladie neurologique parfois mortelle. Confusion mentale, troubles des mouvements des yeux, incapacité de coordonner les mouvements, déficit de vigilance pouvant aller jusqu'au coma : ses symptômes ressemblent en tout point à ceux des bébés d'Israël.

Lorsque cette vitamine essentielle fut ajoutée à l'alimentation des enfants, leur état s'améliora en quelques jours, et ils purent rentrer chez eux. On estime entre 600 et 1 000 le nombre de bébés israéliens qui, au cours de leurs premiers mois de vie, furent privés de thiamine pendant environ deux à trois semaines. On pensait alors que la restauration d'une alimentation équilibrée les avait sauvés. Ce qu'on ignorait, c'est qu'ils présenteraient, des années plus tard, des déficits importants dans le domaine du langage. La psychologue israélienne Naama Friedmann a testé une soixantaine d'entre eux à l'âge de 6 ou 7 ans. Malgré les années écoulées, la majorité souffrait encore de déficits majeurs de compréhension et de production du langage. Leur grammaire était anormale. Le simple fait de nommer une image, comme celle d'un mouton, posait des difficultés à certains d'entre eux. Par contre, leur traitement conceptuel semblait intact : ils savaient associer, par exemple, l'image d'une pelote de laine avec celle d'un mouton plutôt que d'un lion. Et sur tous les autres plans, et notamment celui de l'intelligence (le fameux QI), ils ne semblaient pas affectés.

Cette histoire étonnante, heureusement exceptionnelle, illustre les limites de la plasticité cérébrale. L'apprentissage du langage repose évidemment sur l'immense plasticité du cerveau : n'importe quel bébé est capable d'apprendre toutes les langues du monde, depuis les tons du chinois jusqu'aux clics des Bantous d'Afrique du Sud. Mais cette plasticité n'est pas infinie, et quelques semaines de privations peuvent entraîner un déficit permanent. Et comme le cerveau est organisé de façon modulaire, il se peut très bien que ce déficit soit restreint à un domaine étroit, tel que l'apprentissage du langage. La littérature pédiatrique regorge d'exemples, le plus connu d'entre eux étant le syndrome d'alcoolisation fœtale, causé par l'exposition du fœtus à l'alcool ingurgité par la mère. L'alcool est un puissant tératogène, un véritable poison pour le cerveau en développement, à éviter absolument tout au long de la grossesse. Pour que



s'épanouissent les arbres dendritiques, il faut apporter au cerveau tous les nutriments dont il a besoin.

## *Pouvoirs et limites de la plasticité synaptique*

Dans un cerveau bien nourri, jusqu'où la plasticité cérébrale peut-elle aller ? Peut-elle recâbler notre cerveau ? Notre anatomie cérébrale est-elle susceptible de varier considérablement en fonction des événements que nous avons vécus ? La réponse est non : la plasticité est une variable d'ajustement, certes fondamentale dans nos apprentissages, mais encadrée, confinée par toutes sortes de contraintes génétiques qui font de nous ce que nous sommes, le croisement d'un génome fixe et d'expériences uniques.

Il est temps que je vous reparle de Nico. Souvenez-vous : il s'agit de cet artiste que je vous ai présenté en introduction et qui peint de splendides tableaux à l'aide de son seul hémisphère gauche (voir la [figure 1](#) ). En effet, à l'âge de 3 ans et 7 mois, Nico a perdu son hémisphère droit. Il a subi une intervention chirurgicale appelée hémisphérectomie : l'ablation quasi complète d'un hémisphère, en l'occurrence le droit, afin de mettre fin à une épilepsie dévastatrice.

Soutenu par sa famille, ses médecins et par un chercheur de l'École d'éducation de Harvard, Antonio Battro, Nico est néanmoins parvenu à suivre une scolarité élémentaire à Buenos Aires, puis il est allé au lycée à Madrid jusqu'à 18 ans. Son langage oral et écrit, sa mémoire, ses compétences spatiales sont excellentes. Il a obtenu un diplôme de secrétariat informatique. Surtout, il présente des talents remarquables pour le dessin.

Plasticité cérébrale ? Sans aucun doute, puisque Nico maîtrise de nombreuses fonctions qui, chez une personne normale, sont traditionnellement associées à l'hémisphère droit : la capacité de porter attention à la globalité

d'une figure, de recopier un dessin en disposant convenablement les objets dans l'espace ou encore de comprendre l'ironie et les sous-entendus d'une conversation et de deviner les pensées de ses interlocuteurs. Mais plasticité limitée, canalisée, confinée à des circuits précis et qui sont les mêmes que ceux de tous les enfants.

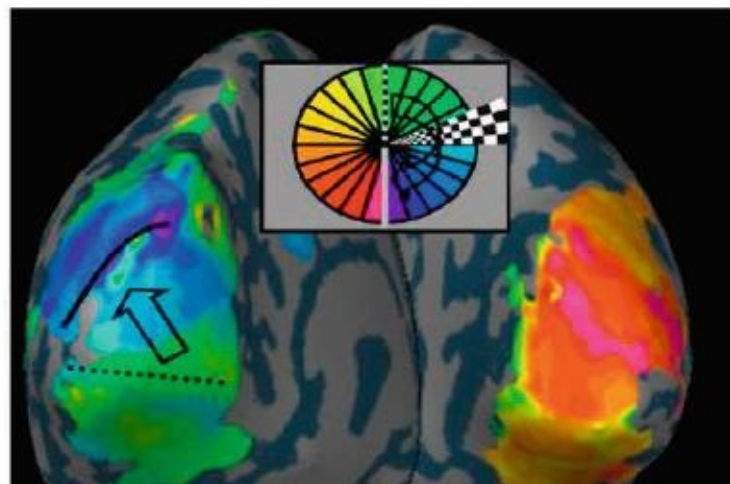
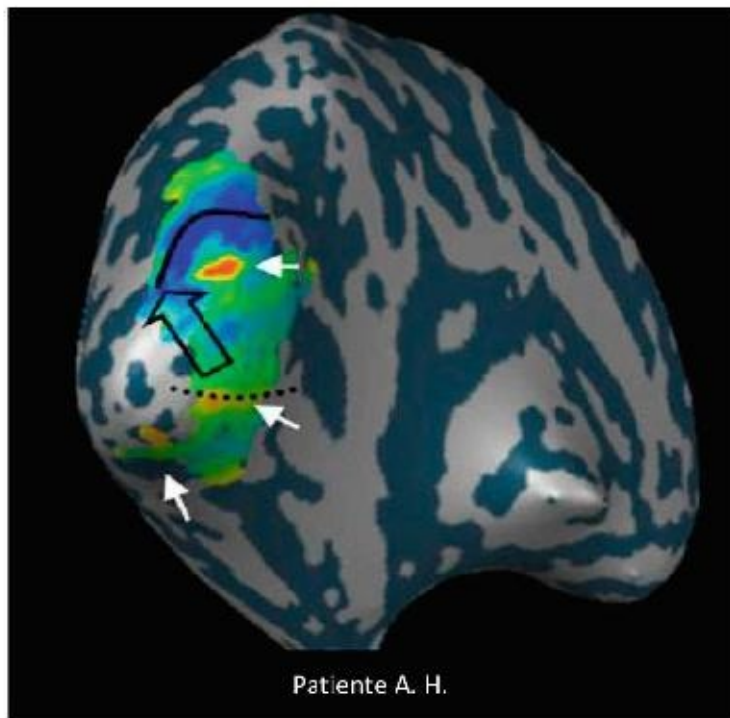
Lorsque nous avons scanné Nico, nous avons constaté qu'il était parvenu à caser tous ses apprentissages dans l'hémisphère gauche, sans que son organisation habituelle en soit bouleversée : toutes les fonctions avaient atterri dans des sites symétriques de leur position habituelle ! Ainsi, la région corticale qui répond aux visages et qui se loge habituellement dans le lobe temporal droit se situait dans l'hémisphère gauche chez Nico – mais à un endroit bien précis, symétrique exact de son site habituel, et souvent activé (faiblement) par les visages chez tous les enfants. Le cerveau de Nico s'était réorganisé, mais sous la contrainte d'une organisation préexistante. Les grands faisceaux de connexions qui, dès la naissance et même *in utero*, parcourent déjà le cerveau du bébé avaient contraint ses apprentissages à rester dans les limites étroites d'une carte corticale universelle.

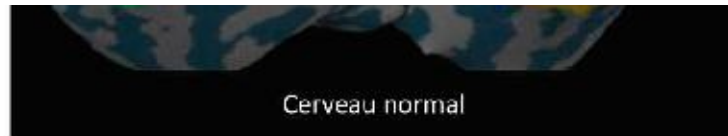
Le pouvoir mais aussi les limites de la plasticité ne sont jamais aussi évidents que lorsqu'on considère les capacités visuelles. Sans surprise, Nico est hémi-anopside, ce qui signifie que sa vision est coupée en deux : une moitié droite où il voit parfaitement, et une moitié gauche où il est totalement aveugle. En effet, avec le croisement des voies visuelles, les entrées visuelles en provenance de la moitié gauche de la rétine atterrissent normalement dans l'hémisphère droit. Chez Nico, elles tombent dans un vide et ne sont pas traitées. Vingt ans de vie ne lui ont pas permis de compenser ce problème fondamental de câblage : la plasticité des connexions visuelles, trop modeste et qui se fige rapidement au cours de la petite enfance, n'a pas su l'empêcher de devenir aveugle pour la moitié gauche du monde.

Maintenant, considérez une petite fille de 10 ans que nous connaissons seulement par ses initiales<sup>110</sup> : A. H. Cette enfant, comme Nico, ne possède qu'un hémisphère gauche mais, contrairement à lui, elle a souffert d'une malformation embryonnaire qui a causé l'arrêt complet du développement de

l'hémisphère droit avant sept semaines de gestation. Autrement dit, tout son cerveau s'est développé sans que l'hémisphère droit ne soit présent. La plasticité a-t-elle bouleversé son cerveau ? Non, mais elle lui a permis d'aller un peu plus loin que Nico. En effet, A. H. voit un peu dans la partie gauche du champ visuel, celle qui aurait dû projeter dans l'hémisphère droit. Sa vision est loin d'être parfaite, mais elle distingue la lumière et le mouvement dans une région proche du centre du regard. L'imagerie cérébrale montre que son cerveau s'est en partie réorganisé (figure 15). Au niveau de l'occiput, dans le cortex occipital qui abrite la vision, on trouve une carte strictement normale de la partie droite du monde – mais également de toutes petites zones qui répondent à la partie gauche. Il semblerait que des axones en provenance de cette moitié de la rétine qui, normalement, aurait dû être aveugle, se soient réorientés en direction de l'autre côté du cerveau. C'est un cas extrême de réorganisation prénatale – et même ainsi, la réorganisation n'est que partielle, très insuffisante pour restaurer une vision normale. Les contraintes génétiques dominent, et la plasticité agit uniquement au sein d'étroites contraintes.

Les scientifiques ont voulu voir jusqu'où on pouvait repousser ces limites génétiques, et une expérience est demeurée célèbre : celle de Mriganka Sur, chercheur au MIT, qui est parvenu à transformer un cortex auditif en cortex visuel<sup>111</sup> ! Chez le furet, *in utero*, il a sectionné les entrées qui, depuis la cochlée et après quelques relais dans le tronc cérébral, se dirigent normalement vers une région bien précise du thalamus et, de là, vers le cortex auditif. Le furet est donc sourd – mais il se produit alors une curieuse réorientation, et les fibres visuelles se mettent à envahir le circuit auditif déconnecté, comme si elles venaient remplacer les entrées manquantes. Résultat : tout un secteur du cortex qui devrait être auditif se comporte à présent comme une aire visuelle. On y trouve toute une carte de neurones sensibles à la lumière et à l'orientation des barres, comme dans n'importe quel cortex visuel. Les synapses s'y adaptent et se connectent à présent à des neurones qui étaient destinés à l'audition, mais qui se sont recyclés en spécialistes de la vision.





**Figure 15.** La plasticité synaptique permet au cerveau de se réorganiser partiellement lorsqu'il subit des atteintes graves. La patiente A. H. (en haut) est née avec un seul hémisphère cérébral : dès sept semaines de grossesse, son hémisphère droit a cessé de se développer. Dans un cerveau normal (en bas), les aires visuelles précoces de l'hémisphère gauche ne représentent que la moitié droite du monde (coloriée en bleu et en vert dans le disque central). Cependant, chez A. H., de toutes petites régions se sont réorientées et se sont mises à répondre à la moitié gauche du monde (en rouge, indiquées par des flèches blanches). Ainsi A. H. n'est-elle pas totalement aveugle du côté gauche, comme le serait un adulte qui perdrait son hémisphère droit. On voit néanmoins que cette réorganisation reste modeste : la part de la génétique domine sur celle de l'expérience.

Faut-il conclure de ces données que la plasticité cérébrale est « massive » et que les entrées « organisent le cortex », comme le voudraient les plus ardents défenseurs de la métaphore de l'ardoise vierge<sup>112</sup> ? Ce n'est pas du tout la conclusion qu'en tire Mriganka Sur. Il insiste au contraire sur le fait qu'il s'agit d'une situation pathologique, et que la réorganisation est loin d'être parfaite : dans le cortex auditif, les cartes visuelles ne sont pas aussi bien différenciées qu'elles devraient l'être. Le cortex visuel est génétiquement préparé à prendre en charge la vision. Au cours du développement normal, chaque région corticale se spécialise précocement sous l'influence de très nombreux gènes de développement. Les axones trouvent leur chemin en suivant des pistes chimiques prédéterminées qui tracent des proto-cartes dans le cerveau en gestation. Ce n'est qu'au bout de la route qu'ils subissent l'influence croissante de l'activité neuronale entrante et s'y adaptent. Le canevas est fixe, seul le petit point change.

Encore faut-il bien comprendre que, lorsque les synapses se modifient sous l'influence de l'activité nerveuse, cela ne signifie pas nécessairement que le cerveau reçoit l'empreinte de l'environnement. *In utero*, avant même de recevoir la moindre entrée sensorielle, le cerveau, les muscles et même la rétine sont déjà parcourus d'une activité spontanée (c'est pour cela que le fœtus bouge). Les neurones sont des cellules excitables : ils déchargent spontanément

et leurs potentiels d'actions s'auto-organisent en ondes massives qui parcourent les tissus. Chaque rétine est ainsi parcourue de vagues aléatoires et, lorsqu'elles parviennent au cortex, ces entrées aident à organiser les cartes visuelles<sup>113</sup>. La plasticité synaptique agit ainsi sans forcément nécessiter d'interactions avec le monde extérieur. Au cours du troisième trimestre de gestation, la frontière entre l'inné et l'acquis se brouille ainsi progressivement, à mesure que le cerveau, déjà bien formé, commence à s'ajuster au monde intérieur et extérieur.

À la naissance, le cortex continue d'être parcouru de décharges aléatoires qui se déroulent sans lien avec les entrées sensorielles. Très progressivement, cette activité endogène, préorganisée, évolue sous l'influence des organes des sens. On peut donner à cette évolution une interprétation précise dans le cadre du modèle du cerveau statisticien<sup>114</sup> : l'activité initiale représente l'*a priori*, les attentes du cerveau, ses hypothèses préalables à toute interaction. Et ces hypothèses s'ajustent progressivement aux données de l'environnement, en sorte qu'après quelques mois de vie, l'activité spontanée a été modifiée pour correspondre de plus en plus étroitement aux statistiques du monde. Le modèle interne du cerveau s'affine à mesure qu'il compile des statistiques sur ses entrées sensorielles. Le résultat final est un compromis, une sélection du meilleur modèle parmi ceux que l'organisation préalable rend disponibles.

## *Qu'est-ce qu'une période sensible ?*

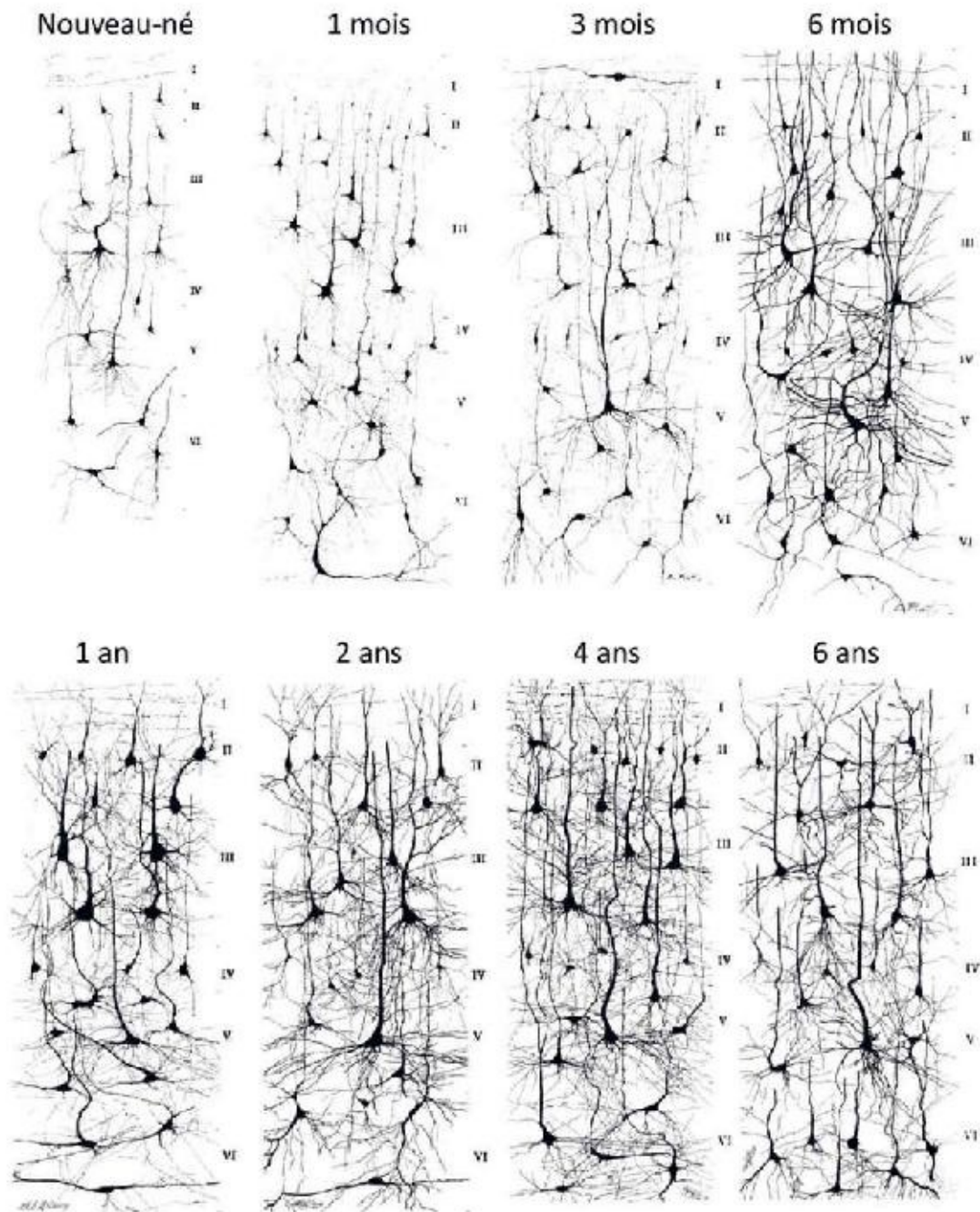
Nous venons de voir que la plasticité cérébrale est à la fois immense et limitée. Tous les grands faisceaux de connexions peuvent et doivent se modifier, mais tous sont déjà en place dès la naissance. Sur le plan cérébral, tous nos apprentissages résultent de petits ajustements, principalement au niveau de microcircuits, souvent à l'échelle de quelques millimètres. Si les neurones et leurs branches terminales bourgeonnent, les grands circuits, eux, restent fermement enracinés dans leur enveloppe génétique. Ils peuvent se myéliniser, s'entourer de cette fameuse gaine isolante qui les accélère et facilite ainsi la

transmission d'informations d'une région à l'autre, mais ils ne peuvent pas se réorienter à volonté.

Cette contrainte spatiale se double d'une contrainte temporelle : dans de nombreuses régions du cerveau, la plasticité n'est présente que pendant un intervalle de temps limité, qu'on appelle une période sensible. Elle atteint un pic, souvent dans la petite enfance, puis elle se referme progressivement. Avec l'âge, l'apprentissage, sans être complètement bloqué, devient de plus en plus difficile à mesure que la plasticité diminue <sup>115</sup> .

Si j'affirme souvent que le bébé est une véritable machine à apprendre, c'est parce que dans les premières années de sa vie, son cerveau est le siège d'un bouillonnement de plasticité synaptique. Les dendrites des neurones pyramidaux du cortex se multiplient à grande vitesse. À la naissance, le cortex ressemble à un bois après le passage d'un ouragan, couvert seulement de quelques troncs d'arbres dénudés. Dans les six premiers mois de vie, véritable printemps du cerveau, branchements et ramifications se multiplient jusqu'à former un fourré inextricable (figure 16) <sup>116</sup> .

Cette complexification pourrait laisser penser que le cerveau, ardoise vierge, reçoit l'empreinte de l'environnement et croît à mesure qu'il enregistre de plus en plus de données. La réalité est bien plus complexe : dans le cerveau immature, les synapses sont engendrées en surnombre, puis elles sont conservées ou éliminées selon leur utilité pour l'organisme. Au cours de la petite enfance, la densité de synapses augmente jusqu'à atteindre le double de celle d'un adulte, et ce n'est qu'ensuite qu'elle décroît lentement. En chaque région du cortex, il nous faut imaginer d'incessantes vagues de surproduction puis de rétraction des synapses inutiles ou, au contraire, de multiplication des synapses et des branches dendritiques et axonales qui ont fait preuve de leur utilité.



**Figure 16.** Dans les premières années de la vie, les arborescences des neurones croissent avec exubérance jusqu'à former un fouillis inextricable. Dans le cerveau d'un enfant de 2 ans, le nombre de synapses est près du double de celui d'un adulte. Mais les arbres dendritiques ne font pas que grandir : ils se font et se défont sous l'influence de l'activité neuronale. Les synapses utiles sont conservées et se multiplient, tandis que les autres sont éliminées.



Pensez-y lorsque vous regardez un tout petit enfant : à chaque seconde qui s'écoule, ce sont plusieurs millions de synapses qui se font et se défont. Cette effervescence explique, en grande partie, l'existence de périodes sensibles. Dans la petite enfance, toute la frondaison dendritique et synaptique peut encore se modifier pour mieux apprendre ; plus le cerveau mûrit, plus l'apprentissage se confine à des modifications marginales.

Ce qui est également remarquable, c'est que les vagues de surproduction synaptique ne surviennent pas partout au même moment<sup>117</sup>. Le cortex visuel primaire, comme toutes les régions sensorielles, atteint beaucoup plus vite sa maturité que les aires corticales de plus haut niveau. Les régions les plus élevées dans la hiérarchie, comme le cortex préfrontal, sont les dernières à se stabiliser : elles continuent de se modifier au cours de l'adolescence et au-delà. Dans l'espèce humaine, le pic de surproduction synaptique se termine vers 2 ans dans le cortex visuel, 3 ou 4 ans dans le cortex auditif et entre 5 à 10 ans dans le cortex préfrontal<sup>118</sup>. La myélinisation, cet enroulement d'un isolant autour des axones, suit la même progression<sup>119</sup>. Dans les premiers mois de vie, les aires sensorielles sont les premières à en bénéficier : leur traitement de l'information s'accélère, et la vitesse de transmission d'une information depuis la rétine jusqu'aux aires visuelles passe d'un quart à un dixième de seconde en quelques semaines<sup>120</sup>. Ce n'est que bien plus lentement que l'isolation atteint les faisceaux qui connectent le cortex frontal, siège de la pensée complexe. Pendant des années, le bébé est un être hybride : ses circuits sensoriels et moteurs mûrissent rapidement, tandis que sa réflexion de plus haut niveau est considérablement plus lente : pendant la première année de sa vie, il lui faut jusqu'à trois ou quatre fois plus de temps qu'un adulte pour prendre conscience d'une information élémentaire, comme la présence d'un visage<sup>121</sup>.

Ces vagues successives de surproduction synaptique et de myélinisation modulent la plasticité cérébrale. Ainsi, une série de périodes sensibles s'ouvrent et se ferment à des moments différents selon les régions cérébrales. Les aires visuelles et auditives sont parmi les plus promptes à perdre leur capacité d'apprendre. L'exemple le mieux étudié, chez l'homme comme chez l'animal, est celui de la vision binoculaire<sup>122</sup>. Nous avons besoin de fusionner les

informations issues de nos deux yeux pour percevoir convenablement la troisième dimension, et cet apprentissage exige que nos deux yeux envoient au cortex des informations de haute qualité pendant une période assez brève : quelques mois chez le chat, quelques années chez l'homme. Si, pendant cette période, un œil reste fermé, renvoie une image floue, ou se décale par la faute d'un strabisme prononcé, le circuit cortical de fusion des deux yeux ne se met pas en place, et cette perte est permanente : le strabisme de l'enfant doit être corrigé dans les premières années de vie, idéalement avant 3 ans, sous peine d'entraîner un déficit définitif au niveau du cortex.

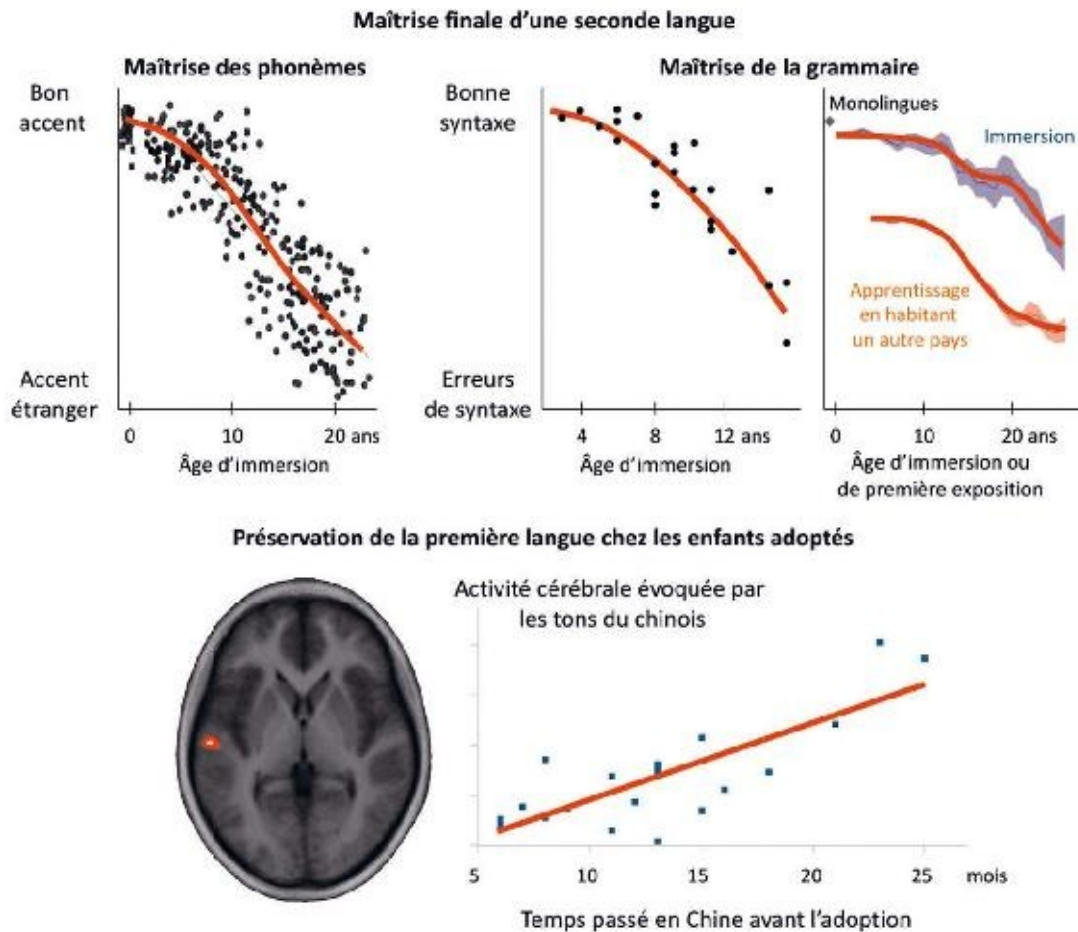
Autre exemple de période sensible : celle qui nous permet de maîtriser les sons de notre langue maternelle. Les bébés sont des champions de l'apprentissage des langues : à la naissance, ils entendent tous les phonèmes de toutes les langues. Où qu'ils naissent, quel que soit leur bagage génétique, il suffit de les immerger dans un bain de langage (monolingue, bilingue ou même trilingue) pour qu'en quelques années ils deviennent des spécialistes de la phonologie de la ou des langues dans lesquelles ils ont baigné. Adultes, nous avons perdu cette faculté d'apprentissage : comme nous l'avons vu, un Japonais peut passer toute sa vie en France sans jamais parvenir à distinguer les sons *R* des sons *L*. Toute sa vie, il confondra glace et grâce, pli et prix, élection et érection...

N'éprouvez aucun sentiment de supériorité car, francophone de naissance, vous ne distinguez ni les *T* dentaux et rétroflexes que n'importe quel locuteur de l'hindi entend comme une évidence, ni les voyelles courtes et longues du finlandais ou du japonais, ni les quatre sortes de tons du chinois. La recherche montre que nous perdons ces capacités vers la fin de la première année de vie<sup>123</sup>. Bébés, nous compilons inconsciemment des statistiques sur ce que nous entendons, et notre cerveau s'ajuste à la distribution des phonèmes qu'emploie notre entourage. Vers 12 mois, quelque chose se fige dans notre cerveau, et nous perdons cette capacité d'apprendre. Sauf extraordinaire, nous ne pourrons plus jamais prétendre être japonais, finlandais ou hindi – la phonologie, c'est (presque) fini. Il faut des efforts immenses pour qu'un adulte recouvre un semblant de discrimination des sons d'une langue étrangère. Ce n'est qu'au prix

d'une rééducation intense, qui commence par amplifier les différences entre *R* et *L* pour les rendre audibles, puis les réduit progressivement, qu'un adulte japonais parvient à recouvrer partiellement la discrimination de ces deux sons<sup>124</sup> .

C'est pour cela que l'on parle de période sensible plutôt que de période critique : l'apprentissage se réduit comme peau de chagrin sans toutefois tomber à zéro. Encore varie-t-il fortement d'une personne à l'autre : à l'âge adulte, ce sont la taille, la forme et le nombre des connexions de notre cortex auditif qui déterminent notre capacité d'apprendre la phonologie des langues étrangères<sup>125</sup> . Pour la plupart d'entre nous, tenter de produire une langue étrangère à l'âge adulte, c'est à coup sûr avoir l'air de l'inspecteur Clouseau, incarné à l'écran par Peter Sellers qui, déguisé en employé du téléphone, déclame avec un épouvantable accent français : « Aïe âme frome ze téléfeune compagnie, zaire ize somessing ze matteure wize youre feune. »

La maîtrise de la phonologie d'une langue étrangère est l'une des premières compétences à décliner avec l'âge : dès ses plus tendres années, un enfant est déjà moins compétent qu'un bébé de quelques mois. La capacité d'apprendre la grammaire d'une langue étrangère, elle, reste ouverte un peu plus longtemps, mais elle finit également par s'effondrer aux alentours de la puberté. On le sait par l'étude des migrants : un enfant qui arrive dans un pays étranger peut devenir excellent dans sa nouvelle langue, mais il présente toujours un petit accent étranger et d'occasionnelles erreurs de syntaxe qui trahissent son origine – et cet écart s'accroît massivement chez les jeunes qui émigrent à l'adolescence ou à l'âge adulte<sup>126</sup> (figure 17).



**Figure 17.** L'apprentissage d'une langue étrangère révèle la diminution de la plasticité cérébrale avec l'âge. Plus on apprend tardivement une langue, moins on parvient à la produire correctement, sans accent étranger ni erreurs de grammaire (en haut). Inversement, plus un enfant passe de temps dans son pays d'origine avant d'être adopté, plus son cerveau conserve une trace de sa langue d'origine (en bas). Celle-ci reste partiellement imprimée dans des circuits neuronaux inconscients.

La modélisation de ces données suggère que la capacité d'apprentissage d'une langue étrangère décline faiblement au cours de l'enfance, mais périclité brutalement vers l'âge de 17 ans<sup>127</sup>. Et comme il faut du temps pour apprendre une langue, les chercheurs recommandent de débiter bien avant l'âge de 10 ans – si possible par l'immersion dans le pays, car rien ne vaut les interactions sociales : la réussite est bien meilleure si l'on est plongé dans une culture étrangère que si l'on se contente d'apprendre à l'école ou par le biais de la

télévision. Il faut faire vite : la plasticité cérébrale pour les langues semble se figer à la fin de la puberté, même si d'autres facteurs liés à la motivation ou à la socialisation ne sont pas totalement exclus.

Encore s'agit-il de la deuxième langue. Si son apprentissage décline relativement lentement, à l'échelle d'une dizaine d'années, c'est sans doute parce qu'elle s'appuie, au moins en partie, sur un cerveau déjà modelé par sa première langue. Que se passerait-il si un enfant était privé de toute exposition au langage pendant les premières années de vie ? Le mythe veut que la question ait été posée pour la première fois par le pharaon Psammétique I<sup>er</sup>. Il aurait confié deux enfants à un berger, avec l'interdiction formelle de leur parler – et les bébés auraient tout de même parlé... en phrygien ! Selon la légende, l'expérience aurait été répétée par l'empereur Frédéric II au XIII<sup>e</sup> siècle, par James IV, roi d'Écosse au XV<sup>e</sup> siècle, ou encore par Jalâluddin Muhammad Akbar, chef de l'Empire moghol au XVI<sup>e</sup> siècle – et les enfants, privés de parole, en seraient morts ! (Les psychanalystes lacaniens raffolent de cette histoire à dormir debout.)

Hélas, point n'est besoin de répandre de telles fables, car cette triste expérience survient régulièrement dans tous les pays du monde : chaque jour, des enfants sourds viennent au monde et, si on ne les aide pas, ils restent prisonniers dans leur bulle. On sait aujourd'hui qu'il est indispensable, dès la première année de vie, de leur donner une langue : soit la langue des signes, ce qui est le plus naturel (les langues signées sont de vraies langues, et les enfants qui les parlent se développent tout à fait normalement), soit une langue parlée, lorsque ces enfants peuvent bénéficier d'un implant cochléaire qui restaure en partie l'audition. Là encore, la recherche montre qu'il faut agir très vite<sup>128</sup> : lorsque les enfants sont implantés après l'âge de 8 mois, ils montrent des déficits permanents dans le domaine de la syntaxe. Ils ne parviennent jamais à comprendre, avec fluidité, des phrases dont les éléments sont déplacés, un phénomène qu'on appelle mouvement syntaxique. Dans la phrase « montre-moi la fille que la grand-mère peigne », il n'est pas évident de comprendre que le groupe nominal « la fille », bien qu'il vienne au début de la phrase, n'est pas le sujet du verbe « peigner », mais son objet. Les enfants sourds, lorsqu'ils sont

appareillés trop tardivement, sont incapables de comprendre de telles phrases et de choisir entre une image où la grand-mère peigne la fille et une autre où la fille peigne la grand-mère.

La petite enfance est donc une phase clé pour le développement du mouvement syntaxique : en l'absence d'interactions linguistiques, la plasticité cérébrale pour cet aspect de la syntaxe se ferme vers la fin de la première année de vie. Souvenez-vous des enfants d'Israël : en 2003, quelques semaines de privation de thiamine dans leurs premiers mois de vie avaient suffi à leur faire perdre définitivement le sens de la syntaxe. Ces résultats convergent avec d'autres études menées chez les enfants sauvages, comme le célèbre Victor de l'Aveyron, ou maltraités, comme la petite fille américaine ironiquement appelée Genie et élevée (ou plutôt rabaissée) dans un placard pendant plus de treize ans, sans qu'on lui parle. Revenus au monde après tant d'années, Victor et Genie se mirent à parler et à apprendre du vocabulaire, mais leur grammaire restait rudimentaire des années plus tard.

L'apprentissage du langage fournit donc d'excellents exemples de périodes sensibles chez l'homme, au niveau phonologique autant que syntaxique. C'est également une bonne illustration de la modularité du cerveau : alors que la grammaire et les sons du langage se figent, d'autres fonctions comme l'apprentissage du sens des mots restent ouvertes tout au long de la vie. C'est grâce à cette plasticité résiduelle que nous pouvons, à tout âge, apprendre ce qu'est un fax, un iPad, un CD-ROM, un trouducteur (néologisme de Ferdinand Céline), un interpote (ami trouvé sur Internet), un phonard (personne qui abuse de son portable) ou un soukamain (sac de femme en désordre ; ces derniers néologismes font partie du loufoque *Dictionnaire des mots qui devraient exister* de Jean-Loup Chiflet).

*Il faut qu'une synapse soit ouverte ou fermée*

Pourquoi la plasticité synaptique se ferme-t-elle ? Quels mécanismes biologiques la bloquent ? L'ouverture et la fermeture des périodes sensibles sont de grands sujets de recherche des neurosciences contemporaines<sup>129</sup> . La fermeture de la période sensible semble liée à l'équilibre entre l'excitation et l'inhibition. Chez l'enfant, les neurones excitateurs sont immédiatement efficaces, tandis que les neurones inhibiteurs se développent plus graduellement. Certains, qui contiennent une protéine appelée parvalbumine, finissent par s'entourer d'une matrice dure, une sorte de filet rigide qui empêche les synapses de bouger et de grandir. Engoncés dans leur filet, les circuits neuronaux ne se modifient plus. Libérez les neurones de ce carcan, par exemple en appliquant un agent pharmacologique tel que la fluoxétine (plus connue sous le nom de Prozac), et la plasticité synaptique revient. C'est une immense source d'espoir pour le traitement des accidents vasculaires cérébraux, où le patient doit réapprendre à se servir des régions préservées.

D'autres facteurs sont également en jeu. Il existe par exemple une protéine appelée Lynx1 qui, lorsqu'elle est présente dans un neurone, inhibe les effets massifs de l'acétylcholine sur la plasticité synaptique. L'acétylcholine, qui signale normalement les moments dignes d'intérêt et module massivement la plasticité, perd son effet sur les circuits adultes envahis par Lynx1. On peut tenter de restaurer la plasticité, par exemple en appliquant un courant qui dépolarise les neurones et les rapproche de leur seuil de décharge<sup>130</sup> . Cette thérapie naissante apporte, là encore, un espoir aux patients figés dans une dépression profonde, et que l'application d'un petit courant électrique à travers le scalp suffit parfois à remettre sur le bon chemin.

On peut se demander pourquoi le système nerveux s'obstine à bloquer ainsi sa propre plasticité. Sans doute est-il avantageux, à un certain moment, de cesser de modifier la structure du cerveau. La simulation des réseaux de neurones montre que, très vite, les neurones de bas niveau acquièrent des champs récepteurs simples et reproductibles, comme des détecteurs de contours dans le cortex visuel. Aucun intérêt de continuer à les mettre à jour, alors que ce type de détecteur est pratiquement optimal : autant s'économiser le coût énergétique lié à la croissance de boutons synaptiques et de bourgeons axonaux. De plus, changer

l'organisation des aires sensorielles primaires, la fondation sur laquelle repose la vision, risque d'ébranler tout l'édifice des aires de niveau supérieur et de leur faire perdre leurs apprentissages. Mieux vaut donc, au bout d'un moment, ne plus y toucher – et c'est sans doute pour cela que l'évolution a bricolé divers mécanismes de fermeture de la période sensible, tout en les déclenchant plus tôt dans les aires sensorielles que dans les régions corticales de plus haut niveau.

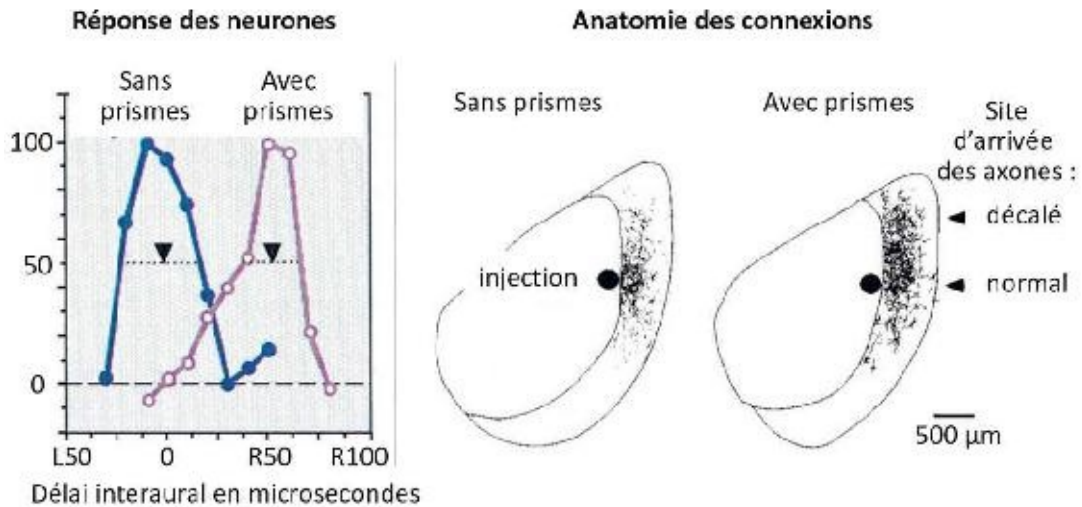
Le bon côté de la chose, c'est que, parce que nos circuits se figent, nous gardons toute notre vie une trace synaptique inconsciente de certains apprentissages précoces. Même si ceux-ci sont rendus obsolètes, par exemple parce que nous réapprenons quelque chose de nouveau, nos circuits cérébraux conservent une trace dormante de nos débuts. Un exemple remarquable est le cas des enfants adoptés, qui réapprennent une nouvelle langue maternelle. Confrontée à de nombreux enfants des rues, victimes de la guerre, de la séparation ou de la misère, la Corée est l'un des pays qui a eu massivement recours à l'adoption internationale. Sur quarante ans depuis 1958, ce sont près de 180 000 enfants qui ont été adoptés, dont 130 000 à l'étranger et plus de 10 000 en France. Christophe Pallier et moi avons scanné une vingtaine d'entre eux à l'âge adulte. Ils étaient arrivés en France entre 5 et 9 ans, et ils n'avaient pratiquement aucun souvenir conscient de leur pays natal, si ce n'est quelques odeurs. En première approximation, leur cerveau se comportait comme celui d'un enfant né en France<sup>131</sup> : leurs aires du langage, dans l'hémisphère gauche, répondaient fortement aux phrases françaises, mais plus du tout aux phrases coréennes (en tout cas, pas plus qu'à une autre langue inconnue, comme le japonais). Au niveau lexical et syntaxique, donc la nouvelle langue avait supplanté l'ancienne.

Et pourtant... les travaux d'un autre groupe de chercheurs ont montré que des Chinois adoptés au Canada conservaient, au plus profond de leur cortex, une trace dormante des tons de leur langue d'origine<sup>132</sup>. Alors qu'un Canadien de souche les traite comme une vague mélodie par le biais de l'hémisphère droit, eux les traitent comme des sons du langage dans une région phonologique de l'hémisphère gauche appelée planum temporale (figure 17). Ce circuit s'inscrit dès la première année de vie et, visiblement, ne se renverse plus par la suite.



Ce n'est pas le seul exemple. Je vous ai déjà expliqué que le strabisme de l'enfant, s'il n'est pas corrigé, modifie définitivement ses circuits visuels. L'éthologue et neurophysiologiste Eric Knudsen a étudié un modèle animal de cette période sensible. Il a élevé des chouettes en leur faisant porter des prismes qui décalent l'ensemble du champ visuel d'une vingtaine de degrés vers la droite. Avec ses chouettes équipées de lunettes, il a réalisé les plus fines études des mécanismes neuronaux de la période sensible<sup>133</sup>. Seules les chouettes qui portaient des prismes pendant l'enfance parvenaient à s'ajuster à cette entrée sensorielle inhabituelle : leurs réponses auditives se décalaient pour s'aligner sur la rétine de sorte qu'elles puissent chasser en s'appuyant simultanément sur l'ouïe et sur la vision nocturne. Les vieilles chouettes, même après avoir porté des prismes pendant des semaines, s'accommodaient moins bien, voire pas du tout. Mais, surtout, les animaux entraînés pendant l'enfance conservaient, pour le reste de leur vie, une trace neuronale permanente de leur expérience précoce. Après l'apprentissage, on observait un double circuit : certains axones des neurones auditifs, dans le colliculus inférieur, avaient conservé leur position normale, tandis que d'autres s'étaient réorientés afin de s'aligner sur la carte visuelle (figure 18). Lorsqu'on ôtait à l'animal ses prismes, il réapprenait à s'orienter correctement ; mais dès qu'on les lui remettait, il se décalait aussitôt de 20 degrés : comme un *full bilingual*, il parvenait à *switcher* d'un *language* à l'autre. Son cerveau avait enregistré, une fois pour toutes, deux jeux de paramètres qui lui permettaient de changer de configuration sans coup férir – exactement comme les Chinois adoptés au Canada conservent une trace cérébrale de la sonorité de leur langue d'origine.

Dans notre espèce aussi, les apprentissages précoces laissent une trace permanente, qu'il s'agisse de la pratique du piano, de la vision binoculaire, ou même du lexique mental. Adultes, nous reconnaissons toujours avec une plus grande vitesse les mots que nous avons entendus dès l'enfance, comme « biberon », « papa », « couche » – tous ceux que la plasticité synaptique précoce a définitivement gravés dans notre mémoire<sup>134</sup>. Le cortex juvénile enregistre presque tout sans effort, et sauvegarde ce savoir dans la mémoire permanente de la géométrie de ses axones et de ses dendrites.



**Figure 18.** Les expériences précoces peuvent s'imprimer profondément dans nos circuits neuro-naux. Une chouette parvient à s'ajuster au port de prismes qui décalent la vision – mais uniquement lorsque cette expérience a lieu dans l'enfance. Ses neurones auditifs, qui localisent les objets en s'appuyant sur le minuscule délai entre les sons qui parviennent aux oreilles droite et gauche, s'ajustent afin de s'aligner sur la vision. Cela se traduit par un déplacement des axones de l'ordre d'un demi-millimètre. Notez que le circuit conserve les deux options : décalée ou normale.

## *Un miracle à Bucarest*

L'intense plasticité cérébrale de nos premières années de vie donne raison à tous les progressistes : investir dans la petite enfance, et notamment dans l'éducation en primaire et en maternelle, est une priorité, car c'est dans ce moment de grâce que le cerveau de l'enfant se transforme le plus aisément et le plus massivement. Plus tard, avec la fermeture progressive des périodes sensibles, l'apprentissage deviendra plus malaisé – mais n'oublions pas que c'est aussi grâce à cette rigidification que notre cerveau conservera, tout au long de sa vie, la trace des modifications synaptiques qui s'y sont opérées dans l'enfance.

Fort heureusement, la fermeture des périodes sensibles ne tombe pas comme un couperet : la plasticité ne disparaît pas du jour au lendemain. De plus, elle varie selon les aires cérébrales. La plasticité se ferme en quelques années dans les aires sensorielles, mais elle s'étend au moins jusqu'à la puberté pour la phonologie et la syntaxe d'une seconde langue. Dans certains domaines, comme le vocabulaire ou les connaissances sémantiques, la faculté d'apprendre semble ne jamais s'affaiblir.

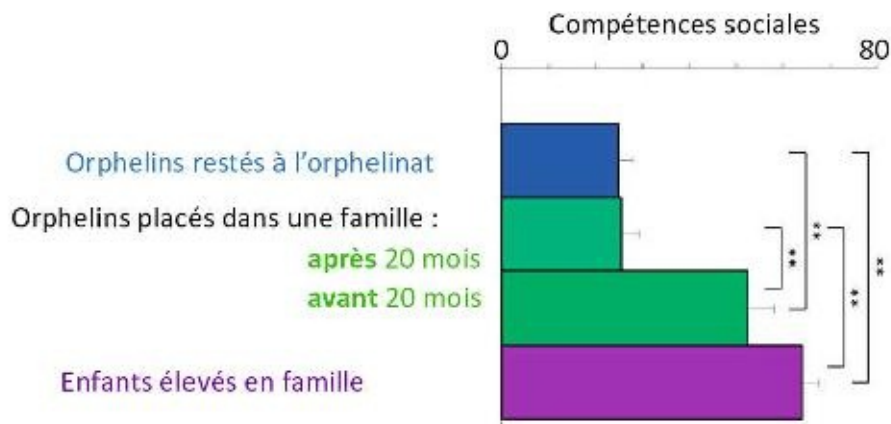
S'il faut se dépêcher d'apprendre, il serait donc profondément faux de conclure que tout se joue avant 3 ans. Le cerveau conserve sa résilience pendant de nombreuses années. Passé la période bénie de la petite enfance, sa plasticité diminue sans toutefois disparaître. Elle s'affaiblit doucement au fil du temps, en commençant par les aires sensorielles périphériques. Les aires corticales de haut niveau conservent toute leur vie leur potentiel d'adaptation. C'est pourquoi une intervention pédagogique fait parfois des miracles, surtout lorsqu'elle intervient le plus précocement possible. Elle ne restaurera peut-être pas toutes les subtilités du mouvement syntaxique ou de la perception des tons du chinois, mais elle parviendra à transformer l'enfant à risque en un jeune adulte épanoui et responsable.

Nous en avons un exemple émouvant, celui des orphelins de Bucarest. En décembre 1989, la Roumanie, alors sous le joug communiste, s'insurgeait soudain. En moins d'une semaine, le peuple révolté chassait le dictateur Nicolae Ceausescu et son épouse, fusillés le jour de Noël. Effarés, nous découvriions les épouvantables conditions de vie des habitants de ce petit morceau d'Europe. L'une des visions les plus insoutenables était celle de ces jeunes enfants aux

corps émaciés, aux yeux morts, abandonnés dans près de 600 orphelinats, véritables mouvoirs où près de 150 000 d'entre eux s'entassaient, presque entièrement livrés à eux-mêmes (figure 19). Convaincu que la force d'un pays réside dans sa jeunesse, le régime de Ceausescu avait promulgué une délirante politique nataliste : taxation massive des célibataires et des couples sans enfants, interdiction de la contraception et de l'avortement, peine de mort pour les avorteurs... tout était fait pour assurer des naissances par milliers. Les couples qui ne pouvaient pas assurer la subsistance de leurs enfants devaient les confier aux services de l'État – d'où ces centaines d'orphelinats qui ne parvenaient pas à assurer l'hygiène, l'alimentation, le chauffage, ni le minimum de chaleur humaine et de stimulation cognitive indispensables à l'épanouissement des petits. Résultat : de véritables enfants zombies qui présentaient des déficits cognitifs et émotionnels majeurs dans tous les domaines.

Après l'ouverture du pays, de nombreuses ONG se sont penchées sur cet immense gâchis. C'est dans ce contexte qu'est né un projet de recherche très particulier, le Bucharest Early Intervention Project<sup>135</sup>. Sous l'égide du secrétariat d'État roumain à la Protection de l'enfance, un chercheur de Harvard, Charles Nelson, a décidé d'étudier scientifiquement les conséquences de la vie en orphelinat, et la possibilité de sauver ces enfants en les plaçant dans des familles d'accueil. Comme il n'y avait aucun programme de placement digne de ce nom en Roumanie, il a créé son propre système de recrutement et est parvenu à enrôler 56 familles volontaires, prêtes à accueillir un ou deux orphelins au sein de leur fratrie. Mais 56 familles, c'était une goutte d'eau face aux sombres profondeurs des orphelinats roumains : seuls 68 enfants pouvaient être accueillis. L'article de *Science* décrit en détail ce moment, qu'on devine dramatique, où l'on réunit 136 enfants, qu'on numérotait de 1 à 136, et dont on tira les numéros d'un grand chapeau : 68 d'entre eux allaient rester en orphelinat, tandis que 68 autres allaient enfin connaître une famille. Cette façon de procéder peut sembler choquante, mais comment faire mieux ? Les ressources humaines étant limitées, le tirage au sort était sans doute la solution la plus équitable. L'équipe continua d'ailleurs de lever des fonds pour sortir toujours plus d'enfants de leur misère, et

un second article de *Science* jugea que les conditions éthiques de l'étude initiale étaient irréprochables<sup>136</sup>.



**Figure 19.** Les traumatismes de l'enfance laissent des traces dans le cerveau, mais une intervention précoce permet de les minimiser. Dans les orphelinats de la Roumanie du dictateur Ceausescu, les enfants étaient maltraités et privés d'interactions avec les adultes. À l'âge de 8 ans, ces orphelins présentaient d'énormes déficits des compétences sociales, même s'ils avaient été placés dans des familles après l'âge de 20 mois. Cependant, ceux qui avaient pu être placés en famille d'accueil avant 20 mois avaient considérablement rattrapé leur retard.

Le tirage au sort permettait de poser une question rigoureuse : toutes choses égales par ailleurs, le placement précoce dans une famille d'accueil permettait-il

aux enfants de remonter la pente ? La réponse fut positive, mais fortement dépendante de l'âge : ce n'est que lorsqu'ils étaient placés avant l'âge de 20 mois que les enfants allaient beaucoup mieux que s'ils étaient restés à l'orphelinat.

Des dizaines d'études ont documenté les effets dramatiques de l'isolement affectif et social sur le développement cérébral, et celle de Bucarest ne fait pas exception : en comparaison avec les enfants nés dans une famille moyenne, tous les orphelins présentaient de graves déficits des fonctions cognitives et même des aspects les plus fondamentaux du fonctionnement cérébral, comme le métabolisme du glucose ou le volume total de matière grise. Après le placement en famille, cependant, une partie de ces indices revenait progressivement à la normale. Six ans plus tard, vers l'âge de 8 ans, les enfants qui avaient été placés avant qu'ils n'atteignent 20 mois avaient nettement progressé par rapport au groupe contrôle et ne présentent plus aucune différence avec les enfants restés en famille. La force des ondes alpha dans leur électroencéphalogramme, qui est un marqueur de l'attention, avait retrouvé son niveau normal, tout comme leurs compétences sociales (figure 19 ). Leur vocabulaire avait également considérablement progressé.

Ces progrès spectaculaires ne doivent pas faire oublier qu'il subsistait un déficit de matière grise – et aussi que les enfants adoptés après 20 mois continuaient, eux, de présenter d'intenses retards dans tous ces domaines. Rien ne remplace vraiment vingt mois d'amour perdu, et ces enfants porteront toujours, dans leur cerveau, les stigmates des graves privations qu'ils ont subies. Mais l'étude des orphelins de Bucarest, comme celle des adoptés de Corée, souligne à quel point la plasticité du cerveau facilite la résilience et la réversibilité des traumatismes précoces : pourvu qu'ils soient traités à temps, de nombreux déficits sont loin d'être irréversibles.

# CHAPITRE 6

## Recyclez votre cerveau

---

Résumons-nous. Tous les bébés viennent au monde avec un riche savoir, un jeu d'hypothèses universelles. Leurs circuits cérébraux sont bien organisés, ce qui leur confère des intuitions fortes dans toutes sortes de domaines : objets, personnes, temps, espace, nombres... Leurs compétences statistiques sont remarquables et c'est en véritables scientifiques en herbe qu'ils sélectionnent les meilleurs modèles du monde.

Dès la naissance, tous les grands faisceaux de connexions sont en place, mais la plasticité cérébrale peut en réorganiser les terminaisons. Synapses et faisceaux de connexions se modifient en permanence chaque fois que nous apprenons des connaissances nouvelles. Envoyer un enfant à l'école, enrichir son environnement, c'est modifier profondément son cerveau et y graver des compétences qu'il gardera tout au long de sa vie. Cette plasticité est toutefois contrainte, à la fois dans l'espace (quelques millimètres) et dans le temps : beaucoup de circuits se ferment à l'apprentissage après quelques mois ou quelques années de vie.

Il nous reste à comprendre comment l'éducation joue avec la plasticité cérébrale précoce. Comment se fait-il que l'espèce humaine puisse étendre ses

capacités dans des directions jamais anticipées par l'évolution ? Comment ne pas être fasciné de voir le primate humain, seul en son genre, prendre une craie ou un clavier et se mettre à écrire ou à calculer ? Que le singe que nous sommes puisse apprendre à lire ne devrait jamais cesser de nous émerveiller. Comme le dit si bien Vladimir Nabokov, « nous sommes absurdemment accoutumés au miracle de quelques signes écrits capables de contenir une imagerie immortelle, des tours de pensée, des mondes nouveaux avec des personnes vivantes qui parlent, pleurent, rient. [...] Et si un jour nous allions nous réveiller, tous autant que nous sommes, et nous trouver dans l'impossibilité absolue de lire<sup>137</sup> ? ».

J'ai longuement étudié le cerveau des analphabètes, que ce soit au Portugal, au Brésil ou encore en Amazonie – des gens qui n'ont jamais eu la chance d'aller à l'école, parce que leur famille ne pouvait pas se le permettre, ou parce qu'il n'y avait aucune école à proximité. La recherche montre que leurs compétences sont, par certains côtés, profondément différentes<sup>138</sup> : non seulement les lettres leur échappent, mais ils éprouvent également des difficultés à reconnaître les formes, à distinguer deux images en miroir<sup>139</sup>, à faire attention à une partie d'un visage<sup>140</sup> ou à mémoriser et à distinguer les mots parlés<sup>141</sup>. Autant pour Platon, qui croyait naïvement que l'apprentissage de la lecture, en nous permettant de nous reposer sur la mémoire externe du livre, allait ruiner notre mémoire interne. Rien n'est plus faux. Le mythe du barde ou du griot qui, bien qu'illettré, posséderait sans effort une immense mémoire, a vécu. Tous, nous devons exercer notre mémoire – et le fait d'être allé à l'école et d'avoir appris l'alphabet aide énormément.

L'impact de l'éducation est encore plus frappant dans le domaine des mathématiques<sup>142</sup>. Un Indien d'Amazonie qui n'est jamais allé à l'école et n'a jamais appris à compter dispose certes d'un jeu universel d'intuitions arithmétiques et géométriques, le même que celui des très jeunes enfants : il distingue les formes géométriques de base, comprend l'organisation de l'espace, sait y naviguer en ligne droite, perçoit la différence entre des quantités comme 2 et 4 ou 30 et 50, et sait qu'on peut les ordonner de gauche à droite. Toutes ces compétences innées, nous les partageons d'ailleurs avec d'autres animaux aussi divers que le corbeau, le singe macaque ou le poussin fraîchement éclos – elles



se mettent en place sur une base génétique, et nous les héritons de notre évolution. Cependant, l'éducation décuple ces compétences initiales. C'est grâce à elle que nous comprenons, par exemple, qu'il y a le même écart entre deux nombres successifs. Sans éducation, nous en serions tous restés à l'idée naïve, enfantine, que 1 est bien plus différent de 2 que 8 ne l'est de 9. C'est la pratique du comptage et de l'arithmétique exacte qui, en nous montrant que chaque nombre  $n$  possède un successeur  $n+1$ , nous fait comprendre que tous ces nombres sont équidistants et forment une échelle linéaire, la ligne numérique – alors que les animaux, les très jeunes enfants et les illettrés considèrent que cette ligne se comprime, parce que les grands nombres sont plus proches les uns des autres<sup>143</sup>.

Si nous ne disposions que d'un sens approximatif du nombre, comme les autres animaux, nous serions incapables de distinguer 11 de 12. La précision raffinée de notre sens des nombres, nous la devons à l'éducation – et sur elle repose toute la construction des mathématiques. Sans éducation, il ne nous viendrait pas à l'idée de distinguer un nombre pair d'un nombre impair, pour ne pas parler des concepts de nombre premier ou d'individus aussi saugrenus que  $\sqrt{2}$ ,  $\pi$ ,  $e$ ,  $+\infty\dots$

## *L'hypothèse du recyclage neuronal*

Comment l'éducation parvient-elle ainsi à révolutionner nos représentations mentales pour devenir des primates lecteurs de Proust, Perec, Grothendieck ou Einstein ? Comme nous l'avons vu, tout ce que nous apprenons doit passer par la modification de circuits cérébraux précontraints, déjà largement organisés à la naissance, mais capables de se modifier à l'échelle de quelques millimètres. Toute la diversité des cultures humaines s'inscrit au sein des contraintes qu'impose notre nature d'homme neuronal.

C'est pour résoudre ce paradoxe que j'ai proposé l'hypothèse du recyclage neuronal<sup>144</sup>. L'idée est simple : même si la plasticité synaptique est importante,

surtout dans l'espèce humaine où l'enfance se prolonge pendant près de quinze ans, le cerveau humain reste soumis à des contraintes anatomiques fortes, héritées de son évolution. Dès lors, chaque nouvel objet culturel que nous inventons, tel que l'alphabet ou les chiffres arabes, doit trouver sa niche neuronale dans le cerveau : un circuit dont la fonction initiale est suffisamment similaire pour s'adapter à la nouvelle invention, et dont la flexibilité est suffisante pour être reconverti à ce nouvel usage. Toute acquisition culturelle nouvelle n'est possible que dans la mesure où elle s'appuie sur une architecture neuronale préexistante, qu'elle recycle. L'éducation doit composer avec nos limites neuronales, en s'aidant de la diversité de nos circuits innés et des longues années d'immaturation cérébrale caractéristiques de notre espèce.

Selon cette hypothèse, s'éduquer, c'est recycler son cerveau. Au fil des millénaires, nous avons appris à faire du neuf avec du vieux. Chaque apprentissage scolaire réoriente un circuit neuronal préexistant dans une direction nouvelle. Pour lire ou pour calculer, les enfants s'appuient sur des circuits antérieurs qui ont évolué pour un autre usage, mais qui, grâce leur marge de plasticité, parviennent à se réorganiser et à se spécialiser pour cette nouvelle fonction culturelle.

Pourquoi ce terme étrange de recyclage neuronal ? Parce que ce mot unit deux idées qui évoquent ce qui se produit dans notre cerveau : la reconversion vers un nouveau métier, et la réutilisation d'un matériau doté de qualités propres.

- Se recycler, pour une personne, c'est recevoir une formation complémentaire afin de s'adapter à l'évolution de sa profession, ou de se reconvertir dans une nouvelle activité. C'est exactement ce qui se produit lorsque nous apprenons à lire ou à calculer. Tout apprentissage culturel confère à notre cortex des compétences nouvelles, qui débordent de la sphère normale de compétences du cerveau des primates. L'éducation recycle le cortex, exactement comme un employé se recycle en vue d'un nouvel emploi.
- Être recyclé, pour un matériau, c'est être réintroduit dans un nouveau cycle de production, afin d'en faire un objet neuf. Mais cette réutilisation

est contrainte : pas question de construire une voiture en papier recyclé ! Chaque matériau possède des qualités intrinsèques qui le rendent plus ou moins propre à d'autres usages. De même, chaque région du cortex, par ses propriétés moléculaires, ses circuits locaux et ses connexions à longue distance, est dotée dès la naissance de caractères qui lui sont propres. L'apprentissage doit composer avec ces contraintes matérielles.

En introduisant l'expression « recyclage neuronal », j'ai voulu distinguer l'acquisition d'une nouvelle fonction culturelle de toutes les autres situations où la biologie fait du neuf avec du vieux. En effet l'évolution, selon le biologiste français François Jacob, ne cesse de bricoler : faisant feu de tout bois, elle retape d'antiques organes pour en faire d'élégants engins d'avant-garde. Les plumes des oiseaux ? D'anciens régulateurs thermiques reconvertis en volets aérodynamiques. Les pattes des reptiles et des mammifères ? D'antédiluviennes nageoires. Dans l'atelier de sculpteur de l'évolution, les poumons deviennent un organe de flottaison, un vieux bout de mâchoire reptilienne se transforme en oreille interne, et même la contraction des babines carnassières se convertit en délicieux sourire de la Joconde.

Le cerveau ne fait pas exception : les circuits du langage sont probablement apparus par duplication et réutilisation de cartes corticales archaïques au cours de l'évolution du cortex<sup>145</sup>. Mais ne confondons pas tout : de telles modifications lentes, d'origine génétique, n'entrent pas dans ma définition du recyclage neuronal. Le terme approprié est « exaptation », un néologisme forgé par l'évolutionniste de Harvard Stephen Jay Gould, sur le modèle du mot « adaptation ». Exapter, c'est, au fil de l'évolution, donner à un ancien mécanisme une nouvelle utilité (rendre un ex apte à d'autres tâches !). L'exaptation court sur des dizaines de milliers d'années. Le recyclage neuronal, lui, survient en un temps bien plus court, à l'échelle de quelques jours ou, au plus, de quelques années. Recycler un circuit cérébral, c'est, en peu de temps, par simple apprentissage, sans modification génétique, réorienter en partie sa fonction.

J'ai formulé l'hypothèse du recyclage neuronal dans le but d'expliquer le talent singulier de notre espèce à sortir de sa niche écologique. Si nous

parvenons à acquérir des compétences singulières qu'aucun autre primate ne possède, telles que lire, écrire, compter, calculer, chanter, se vêtir, monter à cheval ou conduire une voiture, c'est que notre plasticité cérébrale s'étend sur de nombreuses années. Cela nous confère une remarquable faculté d'adaptation, que nous amplifions encore en soumettant nos enfants, jour après jour, au régime draconien de l'école. Cependant, le recyclage neuronal, à une plus petite échelle, existe également chez d'autres animaux. Depuis une dizaine d'années, l'idée a été mise à l'épreuve chez le singe macaque, où elle a pu être testée à une échelle inégalée.

En effet, les technologies d'aujourd'hui permettent d'enregistrer l'activité d'une centaine de neurones pendant plusieurs semaines, alors que le singe acquiert une compétence nouvelle. Ces expériences ont donc permis de répondre à une question simple mais profonde : avec l'apprentissage, le code neural est-il radicalement bouleversé, ou bien est-il simplement recyclé ?

Dans une expérience très récente, grâce à une interface cerveau-machine, les chercheurs ont demandé à un singe d'apprendre à contrôler son propre cerveau : si tu veux que le curseur aille à droite, active ces dix neurones-là, si tu veux qu'il se déplace vers le haut, allume ces dix autres cellules, et ainsi de suite<sup>146</sup>. Et cela fonctionne : en quelques semaines, l'animal apprend à infléchir l'activité d'une dizaine de neurones choisis arbitrairement afin de faire bouger un curseur selon ses désirs. Cependant – et c'est là qu'est la clé –, il n'y parvient que si les décharges neuronales qu'on lui demande de produire ne s'écartent pas trop de celles que son cortex produit déjà spontanément. Autrement dit, ce qu'il apprend doit déjà appartenir au répertoire du cortex qu'on lui demande de recycler.

Pour comprendre ce qu'ont montré les chercheurs, il faut prendre conscience que la dynamique du cortex est contrainte. Le cerveau n'explore pas toutes les configurations d'activité auxquelles il pourrait avoir accès. En théorie, dans un ensemble de 100 neurones, l'activité pourrait se promener dans un espace à 100 dimensions, ce qui fait un nombre gigantesque d'états (au minimum  $2^{100}$ , si l'on considère que chaque neurone peut être silencieux ou actif). Mais, dans la réalité, ce n'est pas le cas : les vrais neurones ne parcourent qu'un sous-espace restreint qui ne comprend qu'une dizaine de dimensions. La contrainte sur

l'apprentissage est très simple à comprendre : soit ce que l'on demande au cortex « tient » dans cet espace préexistant, et le singe apprend sans difficulté ; soit on exige de lui une configuration d'activité qui dépasse ces bornes, et il ne parvient pas à l'apprendre. Le comportement que l'animal apprend peut être radicalement nouveau dans l'évolution (un primate contrôle un écran d'ordinateur !), mais l'état neuronal qui l'implémente doit rester confiné dans d'étroites limites, celles des dimensions préexistantes de l'activité corticale. Ce résultat valide directement une prédiction de l'hypothèse du recyclage neuronal.

Il est de plus en plus clair que chaque région du cerveau impose ses propres contraintes. Dans une région du cortex pariétal, par exemple, certains neurones se cantonnent à une seule dimension, une ligne droite<sup>147</sup>. Ils codent toutes les données sur un axe allant de peu à beaucoup, et sont donc parfaits pour signaler s'il y a plus ou moins d'une certaine quantité. Cette dynamique neuronale peut sembler extraordinairement limitée – mais ce qui semble un handicap se révèle un avantage lorsqu'il s'agit de coder des quantités comme la taille, le nombre, la surface ou tout autre paramètre ordonné du plus petit au plus grand. D'une certaine manière, ce cortex est précâblé pour coder les quantités – et de fait, l'expérience montre qu'il intervient systématiquement dès que nous apprenons à manipuler des quantités sur un axe linéaire, depuis les nombres jusqu'au statut social (qui est au-dessus de qui)<sup>148</sup>.

Autre exemple : une région du cortex temporal, le cortex entorhinal, contient les fameuses cellules de grille qui pavent l'espace, et dont je vous ai déjà parlé. Dans cette région, le code neural est bidimensionnel : même s'ils sont des millions, les neurones ne peuvent pas s'empêcher de rester confinés à un plan<sup>149</sup>. C'est évidemment parfait pour former une carte de l'environnement, un plan vu du dessus – et de fait, c'est cette région qui représente la carte mentale avec laquelle un rat se repère dans l'espace. Mais des travaux récents ont montré que cette même région, avec quelques autres, s'allume dès que nous devons apprendre à représenter des données quelconques sur un plan<sup>150</sup>. Ce peut être, par exemple, des oiseaux dont la longueur du cou constitue la première dimension, et celle des pattes la seconde dimension : après apprentissage, les

personnes qui parviennent à naviguer dans cet espace conceptuel le codent dans leur cortex entorhinal.

La liste pourrait continuer : le cortex visuel représente des assemblages de traits et de formes, la région de Broca code pour des arbres syntaxiques<sup>151</sup> ... Chaque région possède sa dynamique propre et qui ne change guère. Chacune d'elles projette sur le monde son espace d'hypothèses : l'une cherche à faire rentrer les données sur une ligne, l'autre à les afficher sur un plan, la troisième sur un arbre... Chacun de ces espaces précède l'apprentissage et, d'une certaine manière, le rend possible. Nous pouvons bien sûr apprendre des faits nouveaux, mais encore faut-il qu'ils trouvent leur niche neuronale, un espace de représentation adaptée à leur organisation naturelle.

Voyons à présent comment cette idée se décline dans le domaine des apprentissages scolaires les plus fondamentaux : l'arithmétique et la lecture.

## *Les mathématiques recyclent les circuits du nombre*

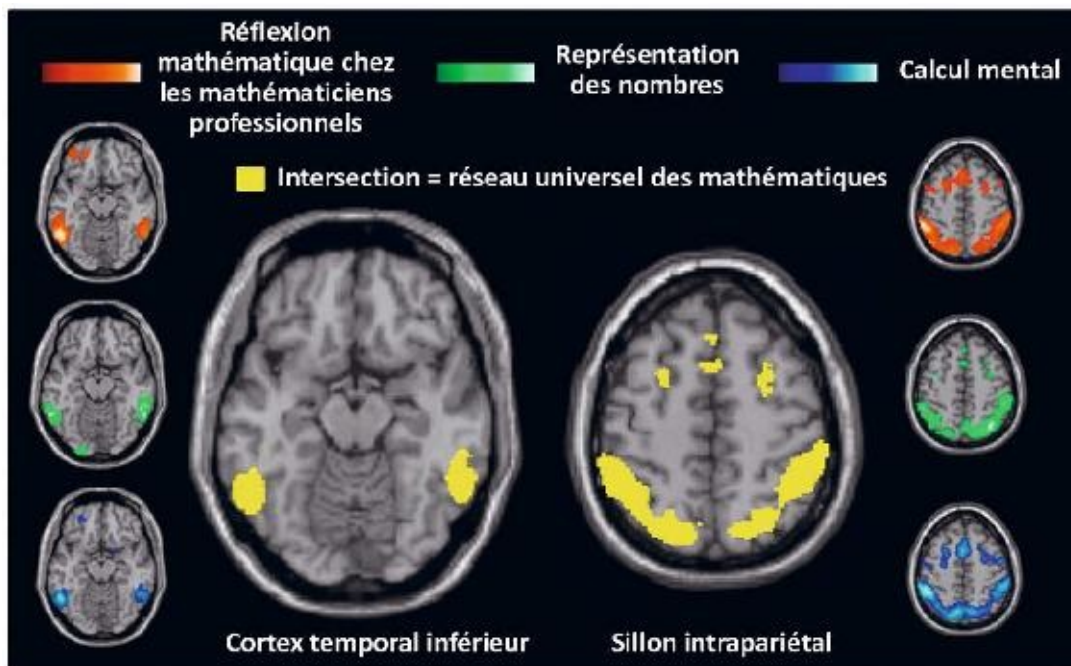
Prenons d'abord l'exemple des mathématiques. Comme je l'ai expliqué dans mon livre *La Bosse des maths*<sup>152</sup>, dans le cas de l'arithmétique, nous avons de nombreuses preuves que l'éducation au calcul mental ne s'imprime pas dans le cerveau comme dans un bloc de cire vierge : elle s'appuie au contraire sur une représentation innée des quantités numériques, qu'elle étend et raffine.

Chez l'homme comme chez le singe macaque, les lobes pariétaux et préfrontaux abritent un circuit neuronal qui représente les nombres sous une forme approximative. Avant tout apprentissage, ce circuit comprend déjà des neurones sensibles au nombre d'objets<sup>153</sup>, une ligne numérique mentale spontanée. Que fait l'apprentissage ? Chez l'animal entraîné à comparer des quantités, les neurones détecteurs de nombre deviennent plus nombreux dans le lobe frontal<sup>154</sup>. Surtout, avec l'éducation aux symboles des nombres, une

fraction de ces neurones devient sélective aux chiffres arabes<sup>155</sup>. C'est un bon exemple de recyclage neuronal : la reconversion (partielle) d'un circuit qui absorbe une invention culturelle nouvelle, les symboles des nombres.

Nous réutilisons également, pour calculer, les régions du lobe pariétal qui nous servent à déplacer l'attention : additionner, c'est déplacer son attention vers la droite, en direction des grands nombres, tandis que soustraire fait appel aux circuits qui déplacent l'attention vers la gauche<sup>156</sup>. Nous avons tous, dans la tête, une sorte de ligne numérique, une carte mentale des nombres sur laquelle nous avons appris à nous déplacer avec précision pour faire des calculs.

Récemment, mon équipe de recherche est allée beaucoup plus loin : avec une jeune chercheuse mathématicienne, Marie Amalric, nous avons examiné si ces mêmes circuits continuent d'être utilisés pour réfléchir aux concepts les plus élevés des mathématiques<sup>157</sup>. Nous avons donc recruté quinze mathématiciens de haut niveau, et nous les avons scannés en IRM alors qu'on leur présentait des expressions mathématiques abstruses qu'eux seuls pouvaient comprendre, telles que  $\int_{\mathbb{R}^n \times \mathbb{R}^n}$ , ou encore des propositions du genre « toute matrice carrée à coefficients dans un corps est combinaison linéaire de trois projecteurs ». Comme nous le prédisions, ces objets mathématiques de haut niveau continuaient d'activer le même réseau cérébral – celui-là même qui, dès la plus tendre enfance, s'active lorsqu'un bébé voit trois objets<sup>158</sup>, ou qu'un enfant apprend à compter (figure 20)<sup>159</sup>. Intégrales doubles, martingales, nombres complexes, espaces de fonctions et variétés en dimension infinie sont des constructions conceptuelles dont les racines puisent, en dernière analyse, dans la recombinaison de circuits neuronaux élémentaires présents dans l'enfance. Tout au long de la construction culturelle des mathématiques, de l'écolier jusqu'au titulaire de la médaille Fields, nous ne cessons de raffiner le code neural du même circuit cérébral.



**Figure 20.** L'éducation consiste à recycler des circuits cérébraux anciens pour les réorienter vers des activités nouvelles. Nous possédons tous, dès la petite enfance, des circuits de représentation des nombres (en vert) et de calcul mental (en bleu). Ce sont ces mêmes circuits que les mathématiciens professionnels continuent d'utiliser lorsqu'ils réfléchissent aux concepts les plus élevés des mathématiques (en rouge). Ces réseaux neuronaux répondent initialement aux ensembles concrets d'objets, mais parviennent à se recycler pour des concepts plus abstraits.

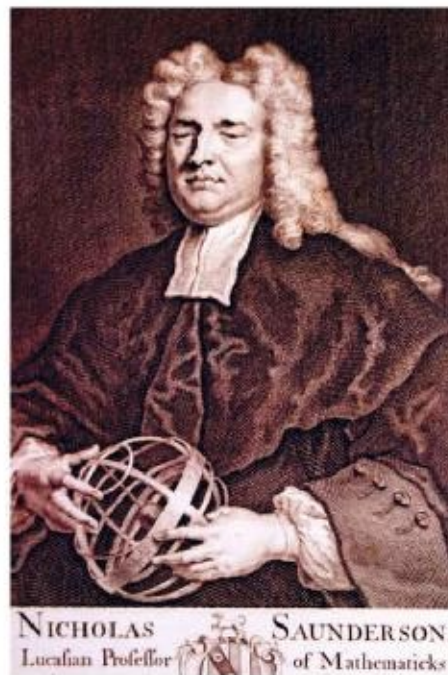
Et ce circuit est fortement contraint par les gènes : l'apprentissage lui permet certes d'accommoder des concepts nouveaux, mais son architecture reste globalement la même. Nous en avons eu la preuve en étudiant des matheux dont l'expérience sensorielle, depuis l'enfance, était radicalement différente : des mathématiciens professionnels aveugles<sup>160</sup>. Car, c'est un fait méconnu, il n'est pas rare qu'un aveugle devienne un excellent mathématicien. Le plus connu de ces mathématiciens non voyants est sans doute Nicholas Saunderson (1682-1739), qui occupa la chaire d'Isaac Newton à l'Université de Cambridge, et était devenu aveugle à l'âge de 8 ans.

Marie Amalric et moi avons repéré trois mathématiciens contemporains aveugles, tous professeurs d'université en France. L'un d'eux, Emmanuel Giroux, est un très grand mathématicien, directeur d'un laboratoire de soixante



personnes à l'École normale supérieure de Lyon. Aveugle depuis l'âge de 11 ans, il est réputé notamment pour sa splendide démonstration d'un grand théorème de géométrie de contact. L'existence même de tels mathématiciens aveugles réfute la vision empiriste du cerveau comme une page blanche qui se remplit sous l'influence de l'expérience sensorielle. En effet, comment un aveugle pourrait-il, d'une expérience aussi restreinte, tirer les mêmes concepts abstraits s'il ne possédait pas déjà, dans son cerveau, les circuits capables de les engendrer, au moins en hypothèse ? Comme le dit Emmanuel Giroux, paraphrasant *Le Petit Prince*, « en géométrie, l'essentiel est invisible avec les yeux, on ne voit bien qu'avec l'esprit » !

Si l'expérience déterminait l'organisation du cortex, un mathématicien aveugle, qui a tout appris par le sens du toucher, devrait activer des aires cérébrales très différentes de celles d'un mathématicien voyant lorsqu'il fait des mathématiques. La théorie du recyclage neuronal, au contraire, prédit que les circuits neuronaux des mathématiques sont fixes. Et c'est ce dernier résultat que nous avons observé lorsque nous avons scanné nos trois mathématiciens aveugles. Comme nous nous y attendions, lorsqu'ils « voient » ce que signifie un théorème de maths, ils le font avec les mêmes circuits des lobes pariétaux et frontaux que ceux d'un mathématicien voyant ([figure 21](#)).



15 mathématiciens voyants :



3 mathématiciens aveugles :



Activation supplémentaire du cortex visuel chez les aveugles :



**Figure 21.** L'apprentissage des mathématiques s'appuie sur des connaissances universelles, largement indépendantes de l'expérience sensorielle. Chez les mathématiciens aveugles, ce sont les mêmes régions du cortex pariétal, temporal et frontal qui s'activent pendant la réflexion mathématique que chez les mathématiciens voyants. Seule différence : les aveugles parviennent également à recycler leur cortex visuel pour faire des maths.

L'expérience sensorielle ne fait rien à l'affaire : seul ce circuit est capable de se recycler pour faire des mathématiques.

La seule différence est que nos trois aveugles, lorsqu'ils cogitent à leur domaine de prédilection, recrutent également les aires visuelles du cortex. Cédric Villani (un autre mathématicien génial, médaille Fields) en avait eu l'intuition. Lorsque nous avons discuté de cette expérience, il m'avait lancé cette boutade : « Tu sais, Emmanuel Giroux, c'est un très grand mathématicien, mais il a de la chance : comme il est aveugle, il peut consacrer encore plus de cortex aux maths ! » Il avait raison, et c'est encore un magnifique exemple de recyclage neuronal. Chez les aveugles, le cortex occipital, normalement consacré à la vision, ne reste pas inactif : il se consacre à de nouvelles fonctions, et notamment au calcul mental et aux mathématiques<sup>161</sup>. Chez les aveugles de naissance, la réorganisation est encore plus extrême, puisqu'on observe dans leur cortex visuel des réponses vraiment inattendues à la grammaire du langage parlé, similaires à celles de l'aire de Broca<sup>162</sup>.

La présence de réponses abstraites dans le cortex visuel des aveugles fait encore l'objet d'un débat théorique<sup>163</sup> : s'agit-il vraiment de recyclage, ou bien d'une preuve extrême de plasticité, qui conduit à la réorganisation totale du cortex ? Il me semble que la balance penche en faveur de l'hypothèse du recyclage neuronal, notamment parce que l'organisation préexistante de cette région n'est pas complètement effacée, comme le serait une ardoise sous les coups d'éponge d'un environnement radicalement différent. En effet, le cortex visuel des aveugles conserve l'organisation de ses connexions et de ses cartes neuronales<sup>164</sup>, tout en les réorientant vers d'autres domaines. En fait, comme ce cortex est très vaste, on y trouve, chez l'aveugle, de nombreuses régions qui répondent non seulement aux maths et au langage, mais aussi aux lettres et aux nombres (présentés en braille), aux objets, aux lieux, aux animaux<sup>165</sup> ... et ces catégories se retrouvent, la plupart du temps, à des endroits similaires du cortex, que l'on soit voyant ou non-voyant. La « boîte aux lettres » du cerveau, par exemple, est située pratiquement au même endroit chez un lecteur voyant et chez un aveugle qui a appris le braille : la fonction de cette région semble largement

déterminée par ses connexions avec les aires du langage, ainsi sans doute que quelques autres propriétés innées<sup>166</sup>.

Pour revenir aux mathématiques, l'hypothèse du recyclage neuronal n'est pas seulement soutenue par la localisation cérébrale identique des concepts les plus élémentaires ( $1 + 1 = 2$ ) et des idées mathématiques les plus avancées ( $e^{-it} + 1 = 0$ ), chez les voyants comme chez les aveugles. D'autres découvertes, purement psychologiques, indiquent que les mathématiques que l'on apprend à l'école reposent sur le recyclage de circuits anciens et consacrés aux quantités approximatives.

Pensez au chiffre 5. En ce moment même, votre cerveau réactive une représentation des quantités, la même que celle que nous partageons avec les autres primates. Maintenant, essayez de décider si 5 est plus grand ou plus petit que 6. L'expérience montre que vous êtes bien plus lent lorsque les nombres sont proches, comme 5 et 6, que lorsqu'ils sont distants, comme 5 et 9. Cet effet de distance<sup>167</sup> est l'une des signatures de la représentation ancienne des nombres que vous avez recyclée lorsque vous avez appris à compter et à calculer. Vous essayez de vous concentrer sur les symboles, mais vous ne pouvez pas vous empêcher d'activer les quantités correspondantes, et celles-ci se recouvrent d'autant plus que les nombres sont proches. Même lorsque vous devez décider si deux chiffres comme 8 et 9 sont différents l'un de l'autre, ce qui devrait être immédiat, vous continuez d'être influencé par la distance qui les sépare – et il en va exactement de même pour les singes qui ont appris à reconnaître les symboles des chiffres arabes<sup>168</sup>.

Je pourrais multiplier les exemples. Quand nous soustrayons deux nombres, le temps que nous mettons varie en proportion directe de la grandeur du nombre soustrait<sup>169</sup>. Tout se passe comme si, dans notre cerveau, nous nous déplaçons mentalement d'un nombre à l'autre : plus nous allons loin, plus nous mettons de temps. De même, quand nous pensons à un prix, nous ne pouvons pas nous empêcher de lui attribuer une imprécision d'autant plus élevée que le nombre est grand : les nombres les plus élevés sont aussi les plus flous<sup>170</sup>. C'est pour cela que, quand nous négocions, contre toute rationalité, nous sommes prêts à lâcher quelques milliers d'euros sur le prix d'un appartement et, le jour même,

barguigner sur le prix trop élevé d'une baguette de pain : l'imprécision qui nous paraît tolérable est proportionnelle au nombre représenté, chez nous comme chez le singe macaque.

Et la liste pourrait continuer : parité, nombres négatifs, fractions... tous ces concepts s'appuient sur la représentation initiale des quantités<sup>171</sup>. Contrairement à un ordinateur numérique, nous sommes incapables de manipuler des symboles dans l'abstrait, nous les ancrons toujours dans le concret des quantités. La persistance de tels effets analogiques dans un cerveau éduqué trahit les attaches anciennes de notre concept de nombre.

Le nombre approximatif est l'un des piliers sur lesquels s'appuie la construction des mathématiques. Cependant, l'éducation l'enrichit. Lorsque nous apprenons à compter et à calculer, cette acculturation mathématique introduit des symboles précis là où l'évolution se contentait, depuis des millions d'années, de quantités floues. C'est un puissant facteur de changement : tous nos circuits arithmétiques se modifient minimalement afin de permettre la manipulation des symboles. De plus, ce sens inné des quantités, même si je l'ai appelé facétieusement la « bosse des maths », n'est certainement pas le seul fondement des mathématiques. Comme nous l'avons vu plus haut, nous héritons également de notre évolution un sens de l'espace, avec ses propres circuits neuronaux et leurs cellules de lieu, de grille et de direction. Nous possédons également un sens des formes, qui permet à un petit enfant de reconnaître un rectangle ou un triangle. D'une manière qui n'est pas encore parfaitement comprise, sous l'influence des symboles que sont les mots et les chiffres, tous ces concepts se recyclent lorsque nous apprenons à faire des mathématiques : nous les recombinaisons sans cesse, dans un langage de la pensée, pour forger des concepts nouveaux<sup>172</sup>. Les primitives de base qui nous sont données par l'évolution sont comme les mots d'un langage nouveau, productif, dont les mathématiciens écrivent chaque jour de nouvelles pages.

## *La lecture recycle les circuits de la vision et du langage parlé*

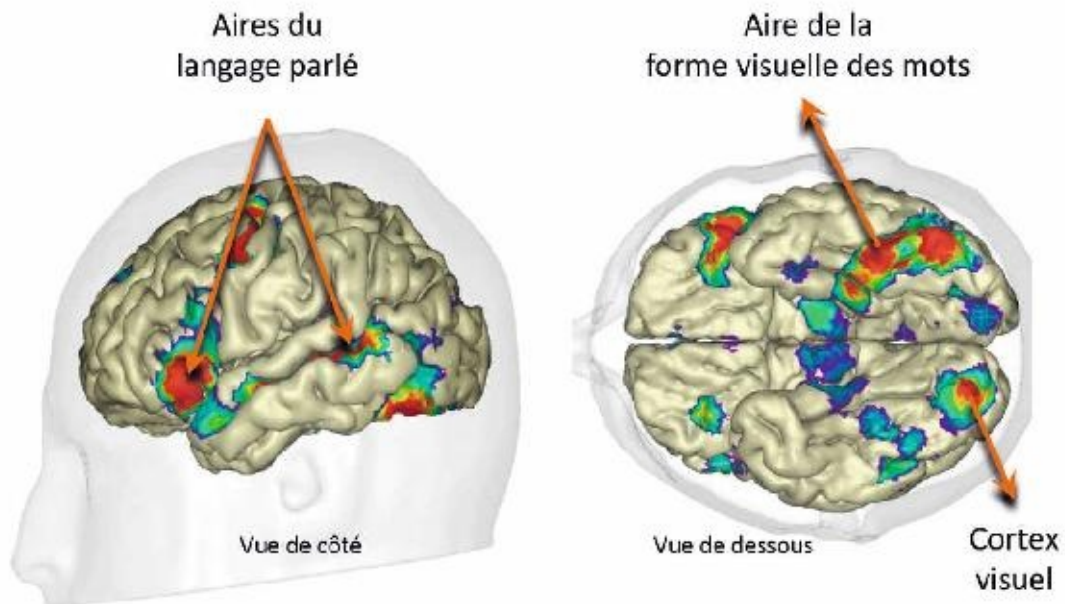
Qu'en est-il de l'apprentissage de la lecture ? Dans mon deuxième livre, *Les Neurones de la lecture*<sup>173</sup>, j'ai décrit en détail les circuits qui nous permettent d'apprendre à lire. C'est une autre preuve du recyclage neuronal : pour lire, nous réutilisons un vaste ensemble d'aires cérébrales qui sont initialement dédiées à la vision et au traitement du langage parlé. Lorsque nous apprenons à lire, nous réorientons ces circuits afin que les régions visuelles reconnaissent les chaînes de lettres et qu'elles les envoient vers les aires du langage. Le résultat, c'est que chez un bon lecteur, les mots lus sont traités exactement comme des mots entendus : l'alphabétisation crée une nouvelle porte d'entrée visuelle vers les circuits du langage.

Bien avant d'apprendre à lire, tout enfant dispose d'un système visuel sophistiqué qui lui permet de reconnaître et de nommer les objets, les animaux et les personnes. Il reconnaît une image quelle que soit sa taille, sa position ou son orientation en trois dimensions, et il sait lui associer un nom. La lecture recycle une partie de ce circuit : l'éducation à la lecture développe une région très spécialisée du cortex visuel, que mon collègue Laurent Cohen et moi-même avons nommée l'aire de la forme visuelle des mots et qui devient la véritable « boîte aux lettres » de notre cerveau. C'est elle qui nous permet de reconnaître les mots, quelle que soit leur taille, leur <sup>position</sup>, leur CaSsE (MAJUSCULE ou minuscule) ou leur **police**<sup>174</sup>. Grâce à ses connexions très directes vers les aires du langage<sup>175</sup>, elle apprend à traduire rapidement ces informations visuelles en sons et en sens.

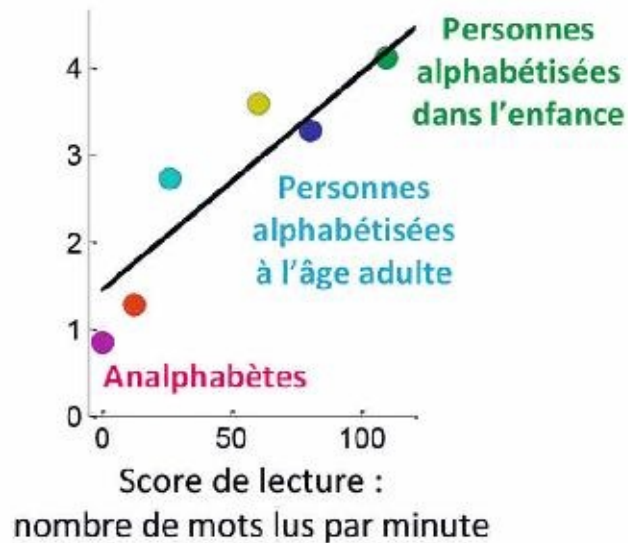
Que se passerait-il si l'on scannait un cerveau d'analphabète à mesure qu'il devient lettré ? Si la théorie est correcte, alors nous devrions littéralement voir le cortex visuel se réorganiser. La théorie du recyclage neuronal prédit que la lecture envahit une région du cortex normalement consacrée à des activités proches, et qu'elle les réoriente vers cette nouvelle tâche. La prédiction est immédiate : cet apprentissage devrait entrer en compétition avec les autres

fonctions préexistantes du cortex. Se pourrait-il que nous perdions certaines fonctions héritées de notre évolution à mesure que nous apprenons à lire ? Ou, au minimum, qu'elles se réorganisent sous les coups de boutoir de la lecture ?

C'est cette prédiction contre-intuitive que mes collègues et moi avons testée dans une série d'expériences. Nous avons commencé par étudier le cerveau d'adultes analphabètes, au Portugal et au Brésil, en le comparant à celui de personnes des mêmes villages mais qui avaient eu la chance d'apprendre à lire, soit à l'école dans l'enfance, soit à l'âge adulte<sup>176</sup>. Les résultats nous ont permis de dresser une carte complète de toutes les régions du cerveau dont la réponse augmente avec l'apprentissage (figure 22). C'est un résultat facile à comprendre. Présentez une série de mots à un illettré. Vous verrez que son cerveau ne répond guère. C'est évident : il ne sait pas les lire, donc l'activité s'arrête rapidement, au niveau des aires visuelles. Maintenant, présentez la même phrase écrite à une personne qui a appris à lire : un très vaste circuit cortical s'allume, en proportion directe de son score de lecture. Les changements s'étendent depuis les aires visuelles primaires, spécialisées dans la reconnaissance des petits caractères<sup>177</sup>, jusqu'aux régions du langage associées à la compréhension des phrases, en passant par la fameuse « boîte aux lettres » du cerveau dans le cortex occipito-temporal de l'hémisphère gauche. Mieux on sait lire, plus ces régions s'activent, et leurs connexions s'améliorent également<sup>178</sup> : automatiser la lecture, c'est fluidifier la relation directe entre lettres et sons du langage.



### Activité cérébrale en réponse aux phrases écrites

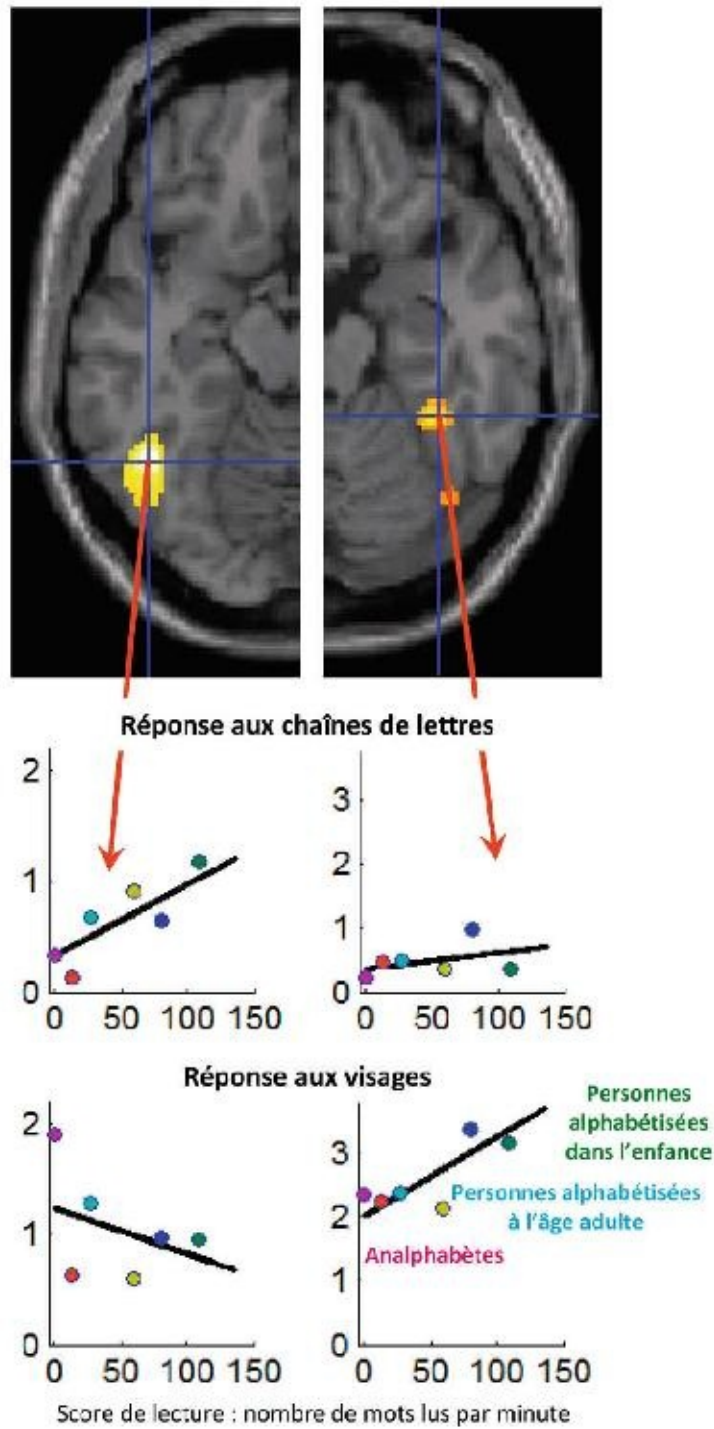


**Figure 22.** L'apprentissage de la lecture recycle un réseau d'aires cérébrales impliquées dans la vision et dans le traitement du langage parlé. Dans toutes les régions en couleurs, l'activité cérébrale évoquée par des phrases écrites augmente avec le score de lecture, depuis les analphabètes complets jusqu'aux lecteurs experts. On voit que la lecture a un double effet : elle spécialise les aires visuelles – particulièrement une région de l'hémisphère gauche appelée « aire de la forme visuelle des mots » – et elle active les aires du langage parlé par le biais de la vision.



Mais on peut également poser la question en sens inverse : existe-t-il des régions qui s'activent plus chez les mauvais lecteurs, et dont l'activité diminue à mesure qu'on apprend à lire ? Nous avons découvert que oui : les analphabètes présentent des activations supérieures pour les visages. Mieux on sait lire, plus cette activité diminue dans l'hémisphère gauche, à l'endroit exact où s'installent les mots écrits. Tout se passe comme s'il fallait faire de la place pour les lettres dans le cortex : la lecture interfère avec une fonction beaucoup plus ancienne, celle de la reconnaissance des visages. Et comme il faut bien apprendre à les reconnaître, nous avons également observé qu'avec l'apprentissage de la lecture, la réponse aux visages augmente dans l'hémisphère droit. Chassés de l'hémisphère gauche, qui est le siège du langage chez la majorité d'entre nous, les visages se réfugient de l'autre côté (figure 23)<sup>179</sup>.

Nous avons fait cette observation chez des adultes lettrés ou illettrés, puis ces résultats ont été répliqués chez l'enfant en train d'apprendre<sup>180</sup>. Dès qu'un enfant commence à lire, l'aire de la forme visuelle des mots se développe dans l'hémisphère gauche – et comme en miroir, la représentation des visages s'accroît dans l'hémisphère droit (figure 24).

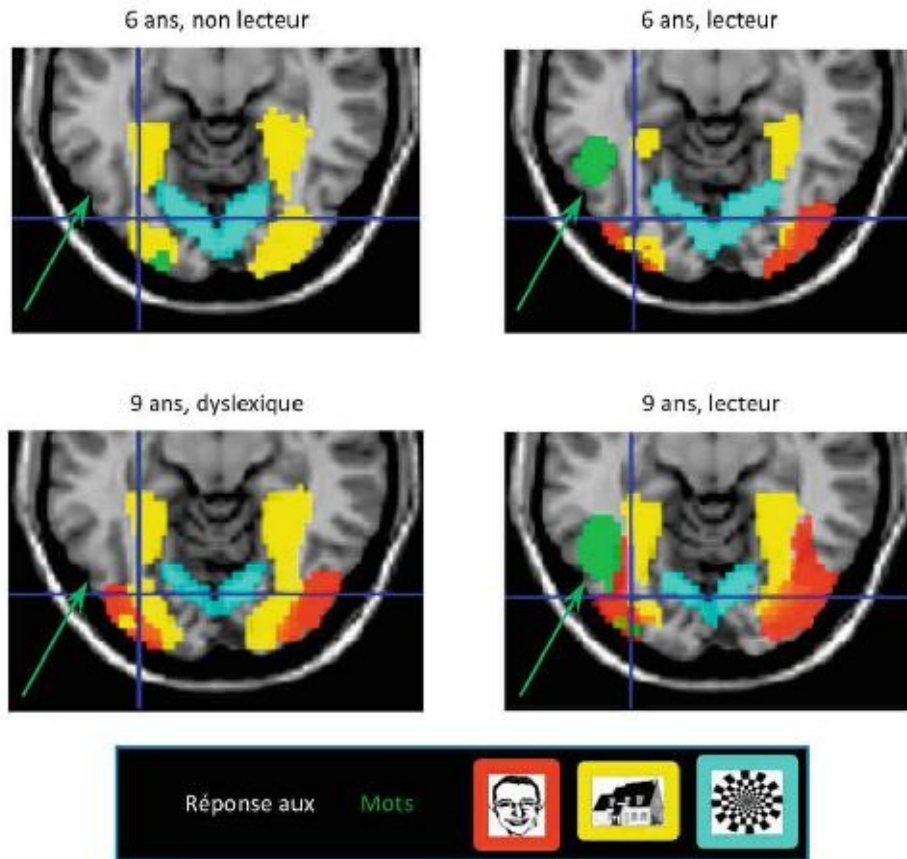
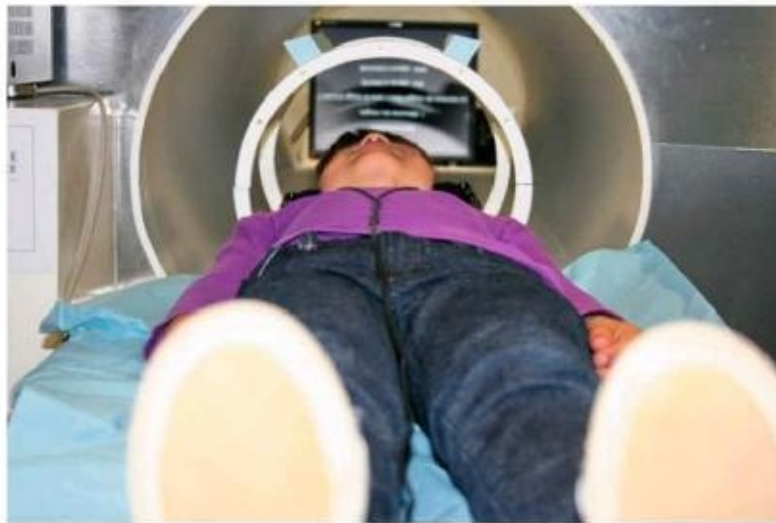


**Figure 23.** Conformément à l'hypothèse du recyclage neuronal, l'apprentissage de la lecture entre en compétition avec les fonctions antérieures du cortex visuel – en l'occurrence la reconnaissance des visages. Plus le score de lecture augmente, plus l'activation évoquée par les mots écrits augmente dans l'hémisphère gauche – et plus l'activation évoquée par les visages se déplace de l'hémisphère gauche vers le droit.

L'effet est tel qu'à âge égal, rien qu'en examinant l'activité évoquée par les visages, un algorithme parvient à décider si un enfant a ou n'a pas encore appris à lire. Et si l'enfant souffre de dyslexie, ces régions ne se développent pas normalement – ni à gauche, dans l'aire de la forme visuelle des mots, ni à droite, dans la région du cortex fusiforme spécialisée pour la reconnaissance des visages<sup>181</sup>. L'activité anormale du cortex occipito-temporal gauche est un marqueur universel des difficultés de lecture, dans tous les pays où cela a été testé<sup>182</sup>.

Récemment, nous avons obtenu l'autorisation de réaliser une expérience audacieuse. Nous souhaitons faire venir les mêmes enfants tous les deux mois dans notre centre d'imagerie cérébrale, afin de les scanner en fin de maternelle et tout au long de la première année d'école, et de voir émerger les circuits de la lecture. Les résultats ont été à la hauteur de nos espérances. Au début, rien à signaler : tant que l'enfant, encore en maternelle, n'a pas appris à lire, son cortex répond aux objets, aux visages, aux maisons, mais pas aux lettres. Après deux mois d'école, par contre, on voit apparaître une réponse spécifique aux mots écrits, à la même position que chez l'adulte dans le cortex occipito-temporal gauche. Et très lentement, la représentation des visages se modifie : au fur et à mesure que l'enfant s'alphabétise, la représentation des visages augmente dans l'hémisphère droit, en proportion directe du score de lecture. Lire entre en compétition avec une autre fonction du cortex, la reconnaissance visuelle des visages, comme le prédit l'hypothèse du recyclage neuronal.

Nous nous sommes rendu compte en réalisant ce travail que cette compétition pouvait s'expliquer de deux manières différentes. Première possibilité, le knock-out : dès la naissance, les visages commencent à s'installer dans le cortex visuel de l'hémisphère gauche, mais la lecture les en déloge, telle une opération coup de poing. Seconde possibilité, le blocus : le cortex se développe lentement, en se spécialisant progressivement pour les visages, les objets ou les maisons. Lorsque les lettres pénètrent dans ce territoire en développement, elles occupent une partie du terrain et bloquent l'expansion des autres catégories visuelles.

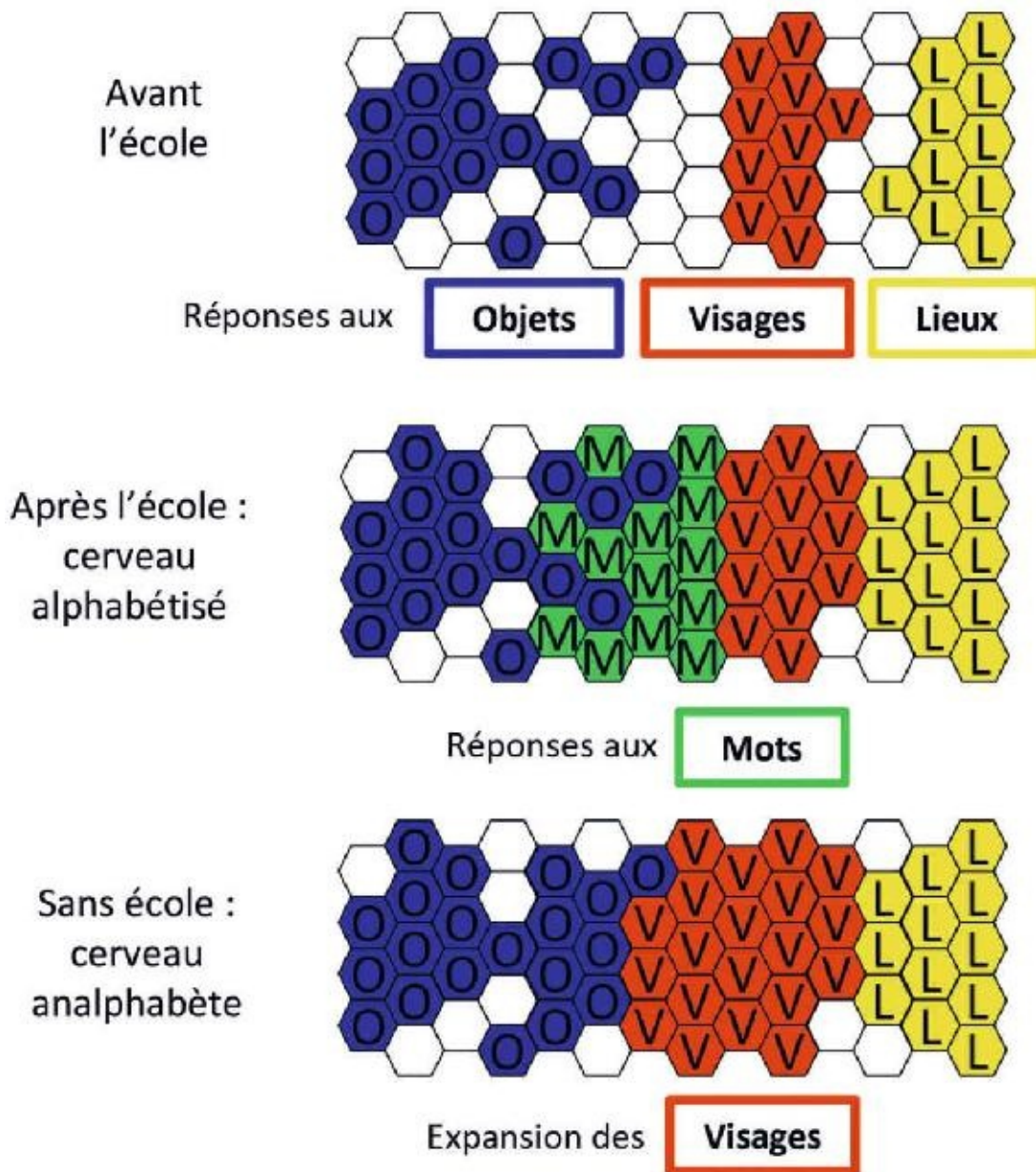


**Figure 24.** L'IRM permet de suivre le développement de la lecture chez l'enfant. Dès qu'un enfant apprend à lire, on voit apparaître une région qui répond aux lettres et aux chaînes de lettres. La lecture recycle une partie de la mosaïque de régions qui, chez tous les primates, servent à reconnaître les visages, les objets et les lieux.

L'alphabétisation conduit-elle à un knock-out ou à un blocus du cortex ? L'expérience montre que c'est la seconde forme de compétition, le blocus, qui semble se produire (figure 25). Grâce aux IRM que nous avons acquises tous les deux mois, nous avons vu la spécialisation du cortex visuel s'accroître très lentement. À l'âge où l'on apprend à lire, vers 6 ou 7 ans, elle est encore loin d'être achevée. Il existe certes des territoires spécialisés pour les visages, les objets ou les lieux, mais aussi beaucoup de régions corticales qui ne se sont pas encore spécialisées pour l'une ou l'autre de ces catégories. Et les lettres, précisément, envahissent une de ces régions demeurées flexibles. Elles n'écrasent donc pas l'aire des visages, mais s'installent juste à côté, dans une région voisine, un peu comme un supermarché agressif s'installe juste à côté d'un petit commerce. L'expansion de l'un bloque celle de l'autre<sup>183</sup>.

En bref, ce n'est sûrement pas un hasard si nous apprenons à lire dans la petite enfance. L'apprentissage de la lecture tire parti de cette période bénie où le cortex visuel est encore labile. Son architecture est loin d'être dépourvue de contraintes, mais elle a évolué pour apprendre à reconnaître toutes sortes d'images et s'adapter à leur forme. Lorsqu'on l'expose à des milliers de mots écrits, elle s'adapte et se recycle pour cette nouvelle activité.

Avec l'âge, très progressivement, un nombre croissant de colonnes corticales se figent et cessent de se modifier. Il devient alors de plus en plus difficile d'automatiser la reconnaissance des lettres et de leurs combinaisons. Mes collègues et moi avons étudié plusieurs personnes qui ont tenté d'apprendre à lire à l'âge adulte : l'une d'entre elles n'avait jamais eu la chance d'aller à l'école, et l'autre avait subi un petit accident vasculaire juste dans la région de la forme visuelle des mots, qui l'avait rendue alexique, incapable de lire un mot. Nous les avons suivies et scannées régulièrement pendant deux ans<sup>184</sup>. Leurs progrès étaient incroyablement lents. La première finit par développer une région spécialisée pour les lettres, mais sans pour autant que les visages ne soient affectés – ce qui était imprimé ne bougeait plus. La seconde, elle, ne parvint jamais à faire émerger une nouvelle « boîte aux lettres » dans son cortex visuel. Sa lecture, tout en s'améliorant, restait lente et semblable au déchiffrement laborieux d'un débutant.



**Figure 25.** L'apprentissage est plus facile dans l'enfance, pendant que le cortex est encore malléable. Avant qu'un jeune enfant n'aille à l'école, certaines régions visuelles de son cerveau se sont déjà spécialisées pour la reconnaissance des objets, des visages ou des lieux – mais il reste de grandes plages de régions demeurées flexibles (symbolisées par des hexagones vides). L'apprentissage de la lecture (en vert) envahit ces régions labiles et bloque la croissance des autres catégories d'objets. Si l'enfant n'apprend pas à lire, ces secteurs s'investissent dans la reconnaissance des visages et des objets, et perdent progressivement leur capacité d'apprendre les lettres.

## *Partitions, équations et visages*

La conclusion est simple : pour recycler profondément notre cortex visuel, et devenir un excellent lecteur, il faut s'appuyer sur la période de plasticité maximale qu'offre la petite enfance. Nos recherches en offrent bien d'autres exemples. Prenons l'apprentissage de la lecture musicale : un musicien qui a appris à déchiffrer des partitions double pratiquement la surface de son cortex visuel dédiée aux portées, par rapport à quelqu'un qui n'a jamais appris la musique. Cette croissance massive et précoce occupe de la place à la surface du cortex, et elle semble y déloger l'aire de la forme visuelle des mots : chez les musiciens, la région corticale qui répond aux lettres, boîte aux lettres du cerveau, se déplace de près d'un centimètre par rapport à sa position normale chez les non-musiciens<sup>185</sup>.

Autre exemple : la reconnaissance des équations mathématiques. Un mathématicien accompli doit savoir analyser, au premier coup d'œil, des expressions aussi absconses que

$$\pi = 3,141592\dots, \phi = 1,61803394\dots, f(x) = \sum_{n=1}^{\infty} \left( a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right)$$

ou  $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$  – exactement comme nous lisons un roman. C'est une expertise visuelle de haut vol. Pour la petite histoire, j'ai assisté un jour à une conférence où le mathématicien Alain Connes (autre médaille Fields) exhiba une équation-univers qui faisait 25 lignes d'une extraordinaire densité. Y figurait, expliqua-t-il, l'expression mathématique de tous les effets physiques de toutes les particules élémentaires connues. Un second mathématicien leva le doigt : il y a une erreur à la ligne 13... Non, répond immédiatement Alain Connes, sans se laisser démonter : elle est compensée par un autre terme à la ligne 14 !

Comment une telle aisance pour les formules se reflète-t-elle dans le cerveau des mathématiciens ? L'imagerie cérébrale montre que ces objets mathématiques envahissent leur cortex jusqu'à occuper une bonne partie des régions occipitales latérales des deux hémisphères – bien plus que chez un non-mathématicien. Et

l'on voit, en parallèle, la surface de cortex disponible pour les visages se réduire comme peau de chagrin – toujours dans les deux hémisphères<sup>186</sup>. Autrement dit, alors que l'alphabétisation chasse les visages de l'hémisphère gauche et les déplace vers le droit, la pratique des chiffres et des équations semble interférer avec la représentation des visages aussi bien à gauche qu'à droite.

Il est tentant de faire le lien avec le syndrome du savant Cosinus, incapable de s'intéresser à autre chose qu'à ses équations, et qui ne reconnaît même plus sa voisine, son chien ou son reflet dans le miroir. Les anecdotes sur les mathématiciens abondent, et vous connaissez peut-être la plaisanterie : « Savez-vous à quoi on reconnaît un mathématicien extraverti ? C'est celui qui regarde *vos* chaussures... »

En réalité, nous ne savons pas encore si cette réduction de la réponse du cortex aux visages, chez les matheux, joue un rôle direct dans l'impression fréquente que la reconnaissance des autres n'est pas leur point fort. La causalité reste à déterminer : passer sa vie dans les formules mathématiques réduit-il la réponse aux visages ? Ou bien, au contraire, accepte-t-on de se plonger dans les calculs parce qu'on les trouve, d'emblée, plus attrayants que les interactions sociales ? Quoi qu'il en soit, la compétition est avérée, et la représentation corticale des visages s'avère remarquablement sensible aux apprentissages scolaires, au point de fournir un reflet de l'éducation qu'un enfant a reçue, en maths, en musique ou en lecture. C'est une autre preuve du recyclage neuronal.

## *Les bénéfices d'un environnement enrichi*

La principale conclusion qu'il faut retenir, c'est que, aux antipodes de la caricature du tout inné ou du tout acquis, le cerveau de l'enfant est à la fois structuré et plastique. Dès la naissance, il est doté d'une panoplie de circuits spécialisés, façonnés par les gènes, et eux-mêmes sélectionnés par des dizaines



de millions d'années d'évolution. Cette auto-organisation lui confère des intuitions profondes de tous les grands domaines du savoir : une physique naïve des objets et de leur comportement, un sens des cartes et de la navigation dans l'espace, des intuitions du nombre, de la probabilité et des mathématiques, un penchant pour les autres êtres humains et même, dans notre espèce, un véritable génie pour les langues. Rien à voir avec une ardoise vierge ! Mais l'évolution a également laissé la porte ouverte à de nombreux apprentissages. Tout n'est pas prédéterminé dans le cerveau de l'enfant, bien au contraire : le détail des circuits neuronaux, à l'échelle de quelques millimètres, est largement ouvert aux interactions avec le monde extérieur.

Les gènes produisent, pendant les premières années de la vie, une exubérante surproduction de circuits neuronaux : deux fois plus de synapses que nécessaire. D'une manière que nous ne comprenons pas encore bien, cette combinatoire foisonnante ouvre un immense espace de modèles mentaux du monde – un véritable langage intérieur. Grâce à elle, le cerveau de l'enfant bouillonne de possibilités et explore un ensemble d'hypothèses bien plus vaste que celui de l'adulte. Chaque bébé est ouvert à toutes les langues, toutes les écritures, toutes les mathématiques possibles – à l'intérieur, bien sûr, de l'enveloppe génétique propre à notre espèce.

De puissants algorithmes d'apprentissage interviennent ensuite pour sélectionner les synapses et les circuits les plus utiles. En quelques années, le cerveau se spécialise et se fige. Les premières régions à se fixer sont les aires sensorielles : dès la première année de vie, les aires auditives commencent à converger vers les voyelles et les consonnes de notre langue maternelle. En quelques années, à mesure que se ferment successivement les différentes périodes sensibles de la plasticité cérébrale, nous devenons, à tout jamais, des natifs d'une langue, d'une écriture, d'une culture particulières. Et si nous sommes privés de stimulation dans un certain domaine, que nous soyons orphelins à Bucarest ou analphabètes dans la banlieue de Brasilia, nous risquons de perdre à jamais notre fluidité mentale dans ce champ du savoir. Intervenir reste toujours souhaitable, car le cerveau conserve toute sa vie une partie de sa plasticité, notamment dans les régions cérébrales du plus haut niveau telles que

le cortex préfrontal. Cependant, les interventions les plus précoces sont aussi les plus efficaces : qu'il s'agisse de faire porter des lunettes à une chouette, d'enseigner une seconde langue à un enfant adopté, de s'ajuster à la surdité, à la cécité ou à la perte de tout un hémisphère cérébral, le plus tôt est toujours le mieux.

L'école est une institution conçue pour tirer le meilleur parti de l'extraordinaire plasticité initiale de notre million de milliards de synapses ( $10^{15}$ ). Elle s'appuie sur la spectaculaire flexibilité du cerveau de l'enfant pour le réorienter, le plus tôt possible, vers des activités nouvelles telles que la lecture ou les mathématiques. De nombreuses expérimentations montrent que, lorsqu'elle intervient précocement, la scolarisation transforme la vie<sup>187</sup> : suivis pendant des décennies, les enfants de milieu défavorisé qui bénéficient d'interventions précoces voient leur vie bouleversée dans tous les domaines, qu'il s'agisse du QI, du revenu moyen, de la santé ou même de la criminalité.

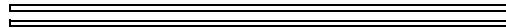
Mais l'école ne peut pas tout. Parents et familles, nous avons tous un devoir de stimulation du cerveau de l'enfant afin d'enrichir au maximum son environnement. Chaque bébé est un physicien en herbe qui adore expérimenter avec la gravité et la chute des corps ; encore faut-il qu'on le laisse bricoler, construire, échouer et recommencer, plutôt que de le sangler dans un siège pendant des heures. Chaque enfant est un mathématicien qui adore compter, mesurer, tracer des droites et des cercles, assembler des formes – pourvu qu'on lui fournisse règles, compas, papier, pliages... Chaque nourrisson est un linguiste : à partir de 18 mois, il apprend un minimum de 10 à 20 mots par jour – encore faut-il qu'on lui parle, que son entourage respecte son appétit de savoir et n'hésite pas à le nourrir d'un vocabulaire enrichi. Le vocabulaire qu'il maîtrisera à 3-4 ans dépend directement des paroles qu'on lui aura adressées, et pas simplement de l'exposition aux paroles ambiantes<sup>188</sup>.

Toutes les recherches le montrent : enrichir l'environnement des très jeunes enfants, c'est construire l'avenir de leur cerveau. L'enfant de 4 ans à qui ses parents lisent tous les jours des histoires montre, en imagerie cérébrale, des activations supérieures à celles des autres bambins dans les réseaux cérébraux du langage parlé, les régions mêmes qui lui permettront plus tard de comprendre

des textes ou de formuler des pensées complexes<sup>189</sup> . Celui qui a la chance de naître dans une famille bilingue, et à qui ses parents font le don merveilleux de lui parler chacun dans sa langue maternelle, acquiert sans effort deux lexiques, deux grammaires, deux cultures<sup>190</sup> . Il gardera, toute sa vie, de meilleures capacités d'analyse du langage et d'apprentissage d'une troisième ou d'une quatrième langue. Et lorsque viendra la vieillesse, il semble que son cerveau résistera plus longtemps aux ravages de la maladie d'Alzheimer. Le confronter, dès le plus jeune âge, à des environnements stimulants, c'est lui permettre de conserver des synapses plus nombreuses, des dendrites plus vastes, des circuits plus flexibles et plus redondants<sup>191</sup> – comme la chouette qui, ayant très tôt appris à porter des prismes, conserve toute sa vie des axones plus diversifiés et capables de passer d'un comportement à l'autre. L'épanouissement du cerveau de l'enfant passe par l'enrichissement de son environnement.

TROISIÈME PARTIE

# LES QUATRE PILIERS DE L'APPRENTISSAGE



Nous venons de voir que tous nos apprentissages reposent sur la modification des circuits de notre cerveau pendant une période sensible où ceux-ci sont encore flexibles et possèdent une marge de plasticité. Cependant, l'existence de la plasticité synaptique ne suffit pas à expliquer l'extraordinaire succès de notre espèce. En effet, cette plasticité est omniprésente dans le monde animal : même la mouche, le ver nématode ou la limace de mer possèdent des synapses modifiables.

Si nous sommes devenus *Homo docens* , si l'apprentissage est devenu notre niche écologique et la principale raison de notre succès planétaire, c'est que notre cerveau contient toute une panoplie d'astuces supplémentaires. Au fil de son évolution, notre cerveau s'est doté de quatre fonctions majeures qui maximisent la vitesse avec laquelle nous parvenons à extraire des informations de notre environnement. Je les appelle les quatre piliers de l'apprentissage, parce que chacun d'eux est indispensable à nos constructions mentales : qu'un seul de ces piliers chancelle et c'est l'ensemble de l'édifice qui vacille. Inversement, chaque fois que nous avons besoin d'apprendre, et d'apprendre vite, nous pouvons nous appuyer sur eux pour optimiser nos efforts. Ces piliers sont : – l'attention, qui amplifie l'information sur laquelle nous nous concentrons ; – l'engagement actif, un algorithme qu'on appelle également « curiosité », et qui incite notre cerveau à évaluer sans relâche de nouvelles hypothèses ; – le retour sur erreur, qui compare nos prédictions avec la réalité et corrige nos modèles du

monde ; – la consolidation, qui automatise et fluidifie ce que nous avons appris, notamment pendant le sommeil.

Au cours de l'évolution, ces fonctions ne sont pas apparues en un jour, et elles ne sont donc pas uniques à notre espèce. Cependant, grâce à notre cerveau social et à notre faculté de langage, nous les exploitons de façon plus efficace que tous les autres animaux – notamment dans nos familles, nos écoles et nos universités.

Faire attention, s'engager, se mettre à l'épreuve et savoir consolider ses acquis sont les secrets d'un apprentissage réussi. L'enseignant qui parvient à mobiliser ces quatre fonctions chez chacun de ses élèves est certain de maximiser la vitesse et l'efficacité avec lesquelles sa classe apprend. Chacun de nous devrait donc apprendre à les apprivoiser – ce qui nécessite de bien comprendre comment chacun de ces systèmes fonctionne et à quoi il sert.

# CHAPITRE 7

## L'attention

---

Imaginez que vous arriviez à l'aéroport juste à temps pour prendre un avion. Tout, dans votre comportement, met en évidence la concentration de votre attention. L'esprit en alerte, vous recherchez le panneau des départs, sans vous laisser distraire par le flot de passagers, puis vous identifiez la ligne qui indique votre vol. Des publicités criardes vous interpellent, mais vous ne les voyez même pas : vous vous dirigez en droite ligne vers le guichet d'enregistrement. Soudain, vous vous retournez, car un ami vient de prononcer votre prénom : ce message, jugé prioritaire par votre cerveau, s'empare de votre attention et envahit votre conscience... vous faisant oublier le numéro du guichet.

Telles sont quelques-unes des fonctions clés de l'attention : éveil et alerte, sélection et distraction, orientation et filtrage. En sciences cognitives, on appelle « attention » l'ensemble des mécanismes par lesquels notre cerveau sélectionne une information, l'amplifie, la canalise et l'approfondit. Ce sont des mécanismes anciens dans l'évolution : le chien qui oriente ses oreilles, la souris qui se fige à l'écoute d'un craquement déploient des circuits attentionnels très proches des nôtres <sup>192</sup> .

Pourquoi l'évolution a-t-elle intégré, chez tous les animaux, des mécanismes attentionnels ? Parce que ces circuits résolvent un problème bien précis : la saturation en informations. Le cerveau est constamment bombardé de stimulations : les sens de la vue, de l'audition, de l'odorat, du toucher lui transmettent des millions de bits d'information par seconde. Tous ces messages sont initialement traités en parallèle par des neurones distincts – mais très vite, il devient impossible de les digérer tous : les ressources énergétiques du cerveau n'y suffiraient pas. C'est pourquoi une pyramide de mécanismes attentionnels, organisés comme un gigantesque filtre, opère un tri sélectif : à chaque étape, notre cerveau décide de l'importance qu'il convient d'accorder à telle ou telle entrée, et alloue des ressources aux informations qu'il juge les plus essentielles.

Sélectionner l'information pertinente est fondamental pour l'apprentissage. En l'absence d'attention, la découverte de régularités dans un monceau de données s'apparente à la recherche d'une aiguille dans une botte de foin. C'est l'une des raisons de la lenteur des réseaux de neurones artificiels conventionnels : ils perdent un temps considérable à analyser toutes les combinaisons possibles des données qu'on leur fournit. Leur lenteur vient du fait qu'ils ne trient pas les informations en fonction de leur pertinence. Ce n'est qu'en 2014 que deux chercheurs, le Canadien Joshua Bengio et le Coréen Kyunghyun Cho, ont montré comment intégrer l'attention à certains réseaux de neurones<sup>193</sup>. Leur premier modèle apprenait à traduire des phrases d'une langue à l'autre. Ils ont montré que l'attention apportait un immense bénéfice : un système doté de cette fonction apprenait mieux et plus vite, parce qu'il se concentrait, à chaque étape de la traduction, uniquement sur les mots pertinents de la phrase de départ.

Très vite, l'idée d'apprendre à faire attention s'est répandue comme un feu de broussaille en intelligence artificielle. Si les systèmes artificiels parviennent aujourd'hui à décrire une image (« une femme qui jette un frisbee dans un parc »), c'est parce qu'ils canalisent l'information en donnant une série de coups de projecteur. Au moment de décrire le frisbee, le réseau se concentre sur cet endroit de l'image et écarte temporairement tous les autres pixels, quitte à y revenir plus tard (figure 26)<sup>194</sup>. L'ingénieur en intelligence artificielle avisé ne connecte plus toutes les entrées avec toutes les sorties. Il remplace ce réseau



isotrope, où tout est connecté avec tout, par une architecture organisée, où l'apprentissage est décomposé en deux modules : le premier apprend à faire attention, le second à traiter les données filtrées par le premier.



Une femme qui jette un frisbee dans un parc.



Une petite fille assise dans un fauteuil avec un ours en peluche.



Une girafe debout dans une forêt, avec des arbres en arrière-plan.

**Figure 26.** L'attention est le premier pilier de l'apprentissage, un mécanisme tellement fondamental qu'il est aujourd'hui intégré aux réseaux de neurones artificiels les plus avancés. Ici, un réseau apprend à trouver les mots pour décrire une image. L'apprentissage repose sur un mécanisme d'attention sélective qui donne un coup de projecteur à certaines zones de l'image (en blanc dans

les images de droite). L'attention sélectionne une zone et y concentre, à un instant donné, toutes les ressources.

L'attention est indispensable, mais elle cause également des problèmes : lorsqu'elle est mal orientée, l'apprentissage est bloqué<sup>195</sup>. Tant que je ne prête pas attention à la femme au chapeau rouge, tout se passe comme si elle n'existait pas. Ses données sont écartées dès les premières étapes de traitement : elles restent confinées aux aires sensorielles précoces. Les images auxquelles nous ne prêtons pas attention ne conduisent qu'à des apprentissages lents et modestes<sup>196</sup>. Rien à voir avec l'extraordinairement amplification qui se produit dans notre cerveau lorsque nous prêtons attention à un objet et que nous en prenons conscience : les neurones sensoriels qui codent pour cet objet s'amplifient, et leurs décharges se propagent jusqu'au cortex préfrontal où se produit un véritable embrasement d'activité qui dure bien au-delà de l'événement initial<sup>197</sup>. Cette forte bouffée de décharges neuronales est exactement ce dont les synapses ont besoin pour changer leur valeur (ce que les neuroscientifiques appellent la « potentialisation à long terme ») et retenir ainsi l'épisode en question.

C'est pourquoi il est essentiel que chacun apprenne à faire attention – et que les enseignants fassent plus attention à l'attention ! Car si les élèves ne prêtent pas attention, ils n'ont aucune chance d'apprendre. Je dis souvent que le plus grand talent de l'enseignant consiste à canaliser et à captiver, à chaque instant, l'attention de l'enfant afin de l'orienter vers le niveau approprié.

L'attention joue un rôle tellement fondamental dans la sélection des informations pertinentes qu'elle est répliquée en de nombreux points du cerveau. Le psychologue américain Michael Posner distingue au moins trois grands systèmes attentionnels :

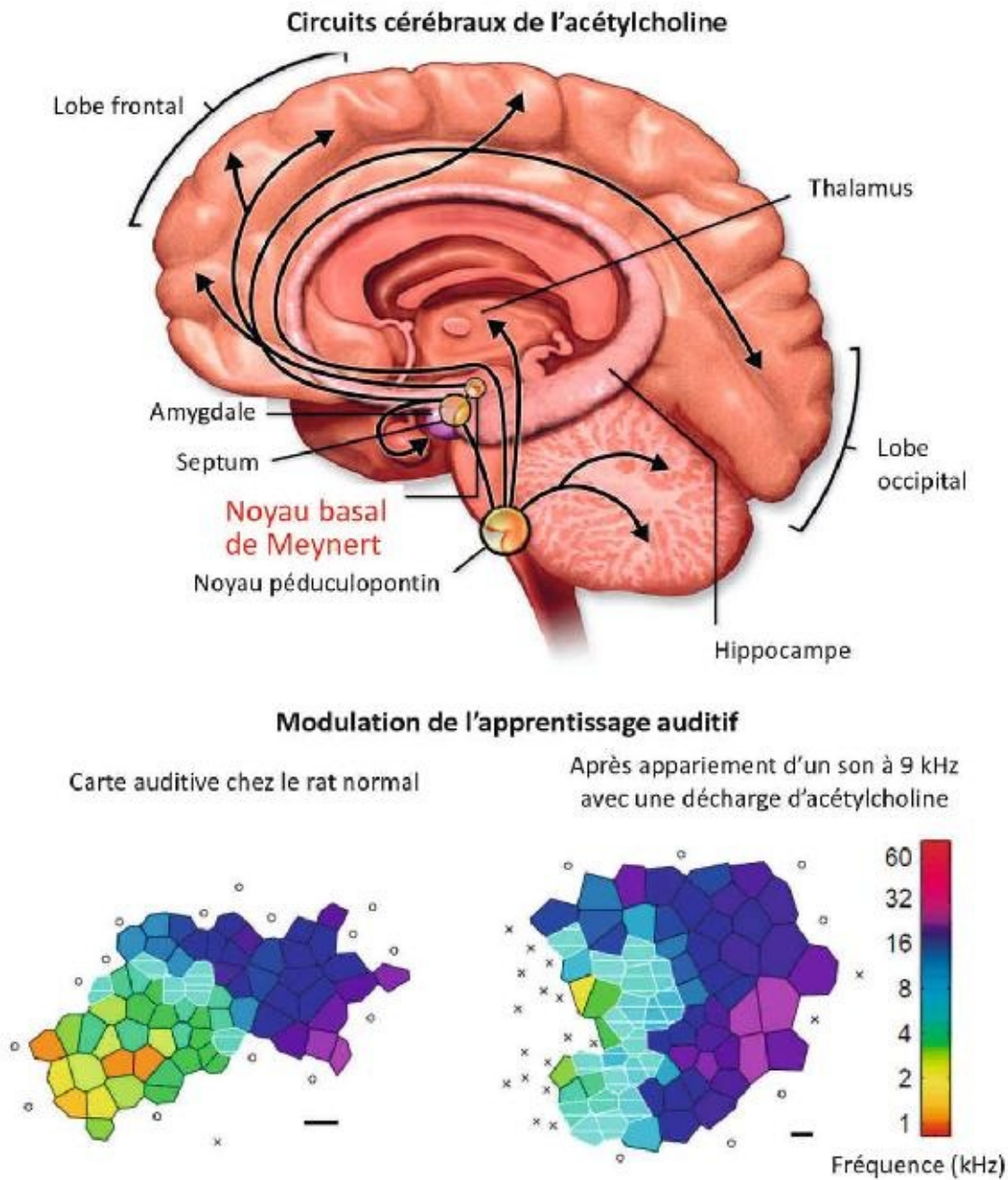
- 1 . L'alerte, qui indique *quand* faire attention et adapte notre niveau de vigilance.
- 2 . L'orientation de l'attention, qui signale à *quoi* faire attention et amplifie tout objet d'intérêt.

3 . Le contrôle exécutif, qui décide *comment* traiter l'information : il choisit la chaîne de traitements appropriée à une tâche donnée et en contrôle l'exécution.

Chacun de ces systèmes module massivement l'activité cérébrale et peut donc faciliter l'apprentissage, mais aussi l'orienter dans la mauvaise direction. Examinons-les un par un.

## *Alerte : l'éveil du cerveau*

Le premier système attentionnel, peut-être le plus ancien dans l'évolution, nous dit *quand* faire attention. C'est un signal d'alerte qui mobilise tout l'organisme lorsque les circonstances l'exigent. Qu'un prédateur s'approche, qu'une émotion forte nous submerge, et toute une série de noyaux sous-corticaux décident immédiatement d'augmenter l'éveil et la vigilance du cortex. Ils ordonnent la libération massive et diffuse de neuromodulateurs tels que la sérotonine, l'acétylcholine et la dopamine (figure 27 ). Par le truchement d'axones dotés de très nombreuses branches diffuses, ces messages atteignent la quasi-totalité du cortex. Ils vont y moduler, massivement, l'activité corticale et l'apprentissage.



**Figure 27.** Les signaux d'alerte déclenchent l'apprentissage. Certains neuromodulateurs, comme la sérotonine, l'acétylcholine et la dopamine, nous disent quand faire attention et semblent forcer le cerveau à apprendre. Dans cette expérience, un son banal, d'une fréquence de 9 kHz, est associé à la stimulation du noyau basal de Meynert (en haut), qui déclenche une décharge d'acétylcholine dans le cortex. Au bout de quelques jours de ce régime, tout le cortex auditif est envahi par cette fréquence et ses voisines (en bas, régions en bleu).

Les chercheurs parlent d'un signal *now print* , « imprimez maintenant », comme si ces systèmes commandaient directement au cortex d'imprimer en mémoire le contenu actuel de l'activité neuronale.

Des expériences chez l'animal montrent qu'effectivement, la décharge des systèmes d'alerte peut radicalement altérer la plasticité des cartes corticales (figure 27 ). Le neurophysiologiste américain Michael Merzenich a réalisé plusieurs expériences où l'on trompe l'animal en stimulant artificiellement ses circuits sous-corticaux de la dopamine ou de l'acétylcholine. Résultat : tous les neurones qui se trouvent activés à cet instant, même s'ils n'ont aucune importance objective, font l'objet d'une amplification massive et durable. Ainsi, lorsqu'un son, disons un *la* dièse, est systématiquement associé à un flash de dopamine ou d'acétylcholine, la plasticité cérébrale est tellement biaisée que toute la carte auditive se modifie et finit par être envahie par cette note arbitraire. L'animal discrimine de mieux en mieux les sons proches de cette note sensible, mais il perd la capacité de se représenter d'autres fréquences<sup>198</sup> .

Il est remarquable que cette plasticité corticale induite par le système d'alerte survienne même chez l'animal adulte. L'analyse des circuits impliqués montre que les neuromodulateurs comme la sérotonine et l'acétylcholine, par le biais notamment du récepteur nicotinique (sensible à la nicotine, cet autre acteur de l'éveil et de la vigilance), modulent l'activité des neurones inhibiteurs du cortex, et affectent ainsi l'équilibre entre l'excitation et l'inhibition dans le cortex<sup>199</sup> . Souvenez-vous que l'inhibition intervient dans la fermeture des périodes sensibles pour la plasticité synaptique. Désinhibés par les signaux d'alerte, les circuits corticaux semblent retrouver leur plasticité : la période sensible peut se rouvrir pour des signaux étiquetés comme cruciaux.

Et chez *Homo sapiens* ? Il est tentant de penser qu'une réorganisation similaire des cartes corticales se produit chaque fois qu'un compositeur ou un mathématicien se passionne pour son domaine de prédilection dès le plus jeune âge. Un Mozart, un Ramanujan sont peut-être tellement électrisés par la passion que leur cerveau se voit envahi de modèles mentaux de la musique ou des maths.

Même si tout le monde n'est pas Mozart, les mêmes circuits cérébraux de l'alerte et de la motivation sont présents dans tous les cerveaux. Quelles

circonstances de la vie quotidienne parviendraient à les mobiliser ? Faut-il attendre un traumatisme ou une émotion forte ? Peut-être pas. Certains travaux suggèrent que les jeux vidéo, notamment ceux d'action où l'on joue à la vie et à la mort, fournissent un moyen particulièrement efficace d'engager ces mécanismes attentionnels. En mobilisant les systèmes d'alerte et de récompense, ils modulent massivement les apprentissages. Le circuit de la dopamine, par exemple, s'embrase lorsque nous jouons à un jeu d'action<sup>200</sup>. La psychologue Daphné Bavelier a montré que cela se traduit par un apprentissage rapide<sup>201</sup>. Les jeux d'action les plus violents, peut-être parce qu'ils mobilisent le plus fortement les circuits cérébraux de l'alerte, semblent avoir les effets les plus massifs : même une dizaine d'heures de jeu améliore les capacités de détection visuelle, d'estimation rapide du nombre d'objets sur l'écran et de focalisation sur une cible sans se laisser distraire par d'autres. Un joueur de jeux vidéo prend des décisions ultrarapides, sans pour autant faire plus d'erreurs.

On ne dira donc plus que les enfants d'aujourd'hui, nourris aux ordinateurs, tablettes, consoles et autres Gameboy, ne savent que zapper avec distraction d'une activité à l'autre. Loin de toujours diminuer la capacité de concentration, les jeux vidéo peuvent aussi l'augmenter. Saurons-nous, à l'avenir, remobiliser la plasticité synaptique chez tous les enfants ? Les jeux vidéo sont un puissant stimulant, et c'est pourquoi mon laboratoire a développé, au fil des années, toute une panoplie de jeux sur tablette fondés sur les principes des sciences cognitives<sup>202</sup>. Mais ce n'est pas la seule voie à explorer : un professeur qui passionne ses élèves, un livre qui absorbe son lecteur, un film ou une pièce de théâtre qui transportent le spectateur peuvent être des signaux tout aussi puissants d'alerte et d'éveil de la plasticité cérébrale.

## *Orientation : le filtre du cerveau*

Le deuxième système attentionnel du cerveau détermine à *quoi* nous prêtons attention. Véritable coup de projecteur sur le monde extérieur, il choisit, parmi

les millions de stimulations qui nous bombardent, celle à laquelle nous choisissons d'attribuer nos ressources mentales, parce que nous la jugeons plus pertinente ou plus digne d'intérêt.

C'est le géant de la psychologie américaine William James qui, dans ses *Principes de psychologie* (1890), a le mieux défini cette fonction de l'attention : « Des millions d'éléments du monde extérieur se présentent à mes sens mais ne pénètrent jamais dans mon expérience consciente. Pourquoi ? Parce qu'ils n'ont aucun intérêt pour moi. Ce que je perçois est ce à quoi j'accepte de prêter attention. Seuls les éléments que je détecte affectent ma pensée – en l'absence de sélection, la perception est un chaos total. L'intérêt seul donne de l'accent et de l'emphase, de la lumière et de l'ombre, un arrière-plan, un premier plan, en un mot une perspective intelligible. L'attention varie d'une créature à l'autre, mais sans elle la conscience de chaque créature serait un fatras brumeux et sans discernement, qu'il est pratiquement impossible de concevoir. »

L'orientation de l'attention dans l'espace se traduit souvent par un mouvement de la tête et du regard : nous nous tournons vers ce qui attire notre attention. Ce faisant, nous amenons l'objet qui nous intéresse au cœur de la fovéa, c'est-à-dire la zone de très haute sensibilité qui se trouve au centre de notre rétine. Même sans bouger les yeux, nous parvenons également à faire attention à un objet ou à un point particulier de l'espace, ce qui amplifie tout ce qui s'y trouve<sup>203</sup>. Nous pouvons également appliquer le filtre de l'attention à des dimensions autres que l'espace visuel. La sélection attentionnelle s'applique à toutes les modalités sensorielles, même les plus abstraites. Par exemple, nous parvenons à prêter attention à l'espace sonore : les chiens bougent leurs oreilles, mais chez nous, seul s'oriente un pointeur interne à notre cerveau. Dans un cocktail animé et bruyant, nous parvenons à sélectionner une conversation parmi dix en nous appuyant sur la voix et le sens. Et rien n'interdit de faire attention à la couleur d'un tableau, à la forme d'une courbe, à la vitesse d'un coureur, ou encore au style d'un écrivain ou à la touche d'un peintre. Toutes les représentations de notre cerveau peuvent recevoir un coup de projecteur attentionnel.



Dans tous les cas, l'effet est le même : les neurones qui reçoivent ce coup de projecteur sont amplifiés, et leurs décharges augmentent. L'impact est double : l'attention rend ces neurones mieux à même de répondre aux informations que nous jugeons pertinentes, mais, surtout, elle augmente leur influence sur le reste du cerveau. Les circuits en aval se font l'écho du stimulus auquel nous prêtons l'œil, l'oreille ou l'esprit. *In fine*, c'est pratiquement l'ensemble du cortex qui s'oriente et amplifie l'objet de toutes nos attentions<sup>204</sup>. L'attention agit comme un amplificateur, un filtre sélectif.

« L'art de faire attention, qui est le grand art, dit le philosophe Alain, suppose l'art de ne pas faire attention, qui est l'art royal. » Faire attention, c'est effectivement choisir ce que l'on décide de négliger. Pour qu'un objet parvienne en pleine lumière, des milliers d'autres doivent rester dans l'ombre. Orienter l'attention, c'est choisir, filtrer, sélectionner : on parle ici d'attention sélective. L'attention accroît les signaux sélectionnés, mais elle réduit également dramatiquement tous ceux qui sont jugés non pertinents. Les scientifiques disent qu'elle « biaise la compétition » : parmi de nombreuses entrées sensorielles en compétition, elle en retient une et élimine les autres. C'est ici que la métaphore d'un projecteur qui éclabousse un objet de lumière trouve ses limites, car pour mieux allumer une région du cortex, le projecteur attentionnel de notre cerveau en éteint également d'autres. Les aires qui sont jugées dépourvues de pertinence sont envahies par des ondes assez lentes, dans les fréquences alpha (de l'ordre de 10 hertz), ce qui les empêche de développer une activité neuronale cohérente.

Faire attention, c'est donc sélectionner – et, en conséquence, prendre le risque d'être aveugle à ce que nous choisissons de ne pas voir. Aveugles, vraiment ? Le terme n'est pas trop fort : une expérience célèbre, celle du gorille invisible, illustre à merveille la cécité totale que cause l'inattention<sup>205</sup>. Dans cette expérience, on vous demande de regarder un petit film où des joueurs de basket, en blanc et en noir, se font des passes. Vous devez compter le nombre de passes de l'équipe blanche. Rien de plus facile, pensez-vous – et de fait, trente secondes plus tard, vous donnez triomphalement le bon compte. « Oui, mais... et le gorille ? » vous demande l'expérimentateur. « Le gorille ? Quel gorille ? » On rembobine le film et, à votre stupéfaction, vous découvrez qu'un acteur, déguisé

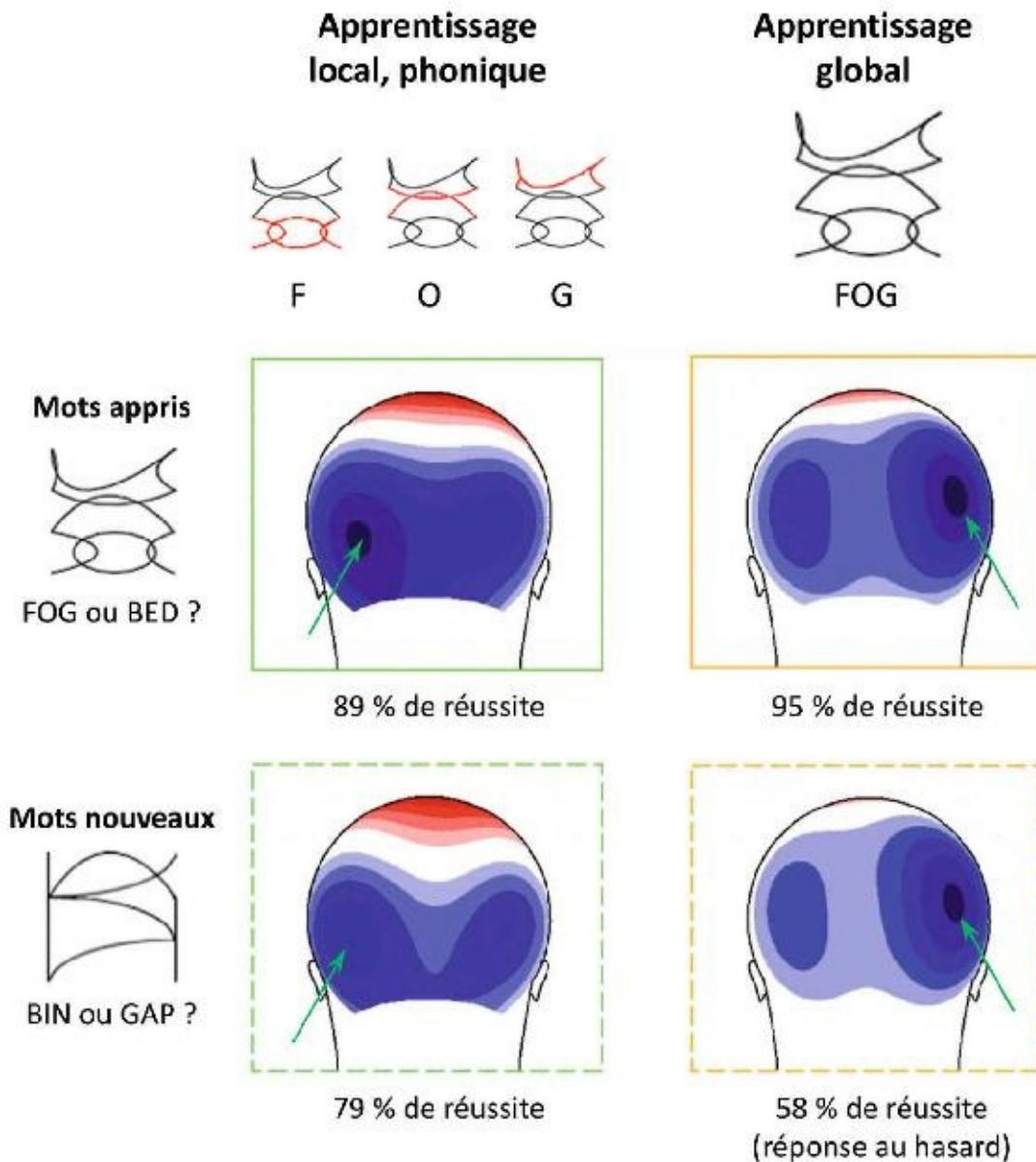
en gorille, vient de traverser toute la scène en se frappant la poitrine. Impossible de le manquer, et d'ailleurs on peut prouver que vos yeux se sont bien posés sur lui. Si vous ne l'avez pas vu, c'est que, concentré sur les joueurs de l'équipe blanche, vous étiez en train d'inhiber les personnages en noir... gorille compris ! Obsédé par la tâche de comptage, votre espace de travail mental était incapable de prendre conscience de cet incongru quadrumane.

L'expérience du gorille est une découverte fondamentale des sciences cognitives, maintes fois répliquée : le simple fait de focaliser son attention sur un objet de pensée rend aveugle à d'autres stimulations. Il y a d'ailleurs, dans l'expérience du gorille, une double illusion. D'une part, nous ne voyons pas le gorille, ce qui est déjà assez grave (d'autres expériences montrent que l'inattention peut nous conduire à griller un stop ou à écraser un piéton – ne téléphonez jamais au volant !). Mais il y a pire : nous sommes inconscients de notre propre inconscience, et cela nous conduit à surestimer constamment ce que nous percevons du monde extérieur. La plupart des gens qui font, pour la première fois, l'expérience du gorille invisible ne parviennent pas à croire à leur propre cécité. Ils pensent qu'on leur a joué un tour, en changeant de film par exemple. Car, disent-ils, s'il y avait un gorille dans la vidéo, je l'aurais vu ! Eh bien c'est faux : notre attention est extrêmement limitée, et malgré toute notre bonne volonté, lorsque nous sommes concentrés sur un objet de pensée, d'autres objets, même saillants, amusants ou importants peuvent nous échapper totalement et demeurer invisibles à nos yeux.

L'expérience du gorille mérite vraiment d'être connue de tous, et particulièrement des parents et des enseignants. En effet, quand nous enseignons, nous avons tendance à oublier ce que c'est que d'être ignorant. Nous pensons que ce que nous voyons, tout le monde peut le voir. Et nous ne comprenons donc pas qu'un enfant puisse, sans aucune mauvaise volonté, ne pas *voir*, au sens le plus littéral du terme, ce qu'on cherche à lui enseigner. Or l'expérience est claire : s'il ne comprend pas à quoi il doit faire attention, il ne le voit pas, et ce qu'il ne voit pas, il ne peut pas l'apprendre<sup>206</sup>.

Prenez, par exemple, une expérience récente du psychologue américain Bruce McCandliss sur l'importance de l'attention dans l'apprentissage de la

lecture<sup>207</sup>. Vaut-il mieux faire attention aux lettres, ou à la forme globale du mot ?



**Figure 28.** L'orientation sélective de l'attention module massivement l'apprentissage. Dans cette expérience, on simule l'apprentissage de la lecture selon la méthode globale ou la méthode phonique. Un groupe de personnes prête attention à la forme globale des mots. Après 300 essais, elles ne s'aperçoivent toujours pas que les mots sont formés de lettres. L'attention globale oriente la lecture vers un circuit inapproprié de l'hémisphère droit et les empêche de généraliser l'apprentissage.

entissage à des mots nouveaux. C'est seulement lorsqu'on attire leur attention sur le niveau des lettres (apprentissage local et phonique) que les personnes parviennent à déchiffrer l'alphabet et à lire des mots nouveaux, en utilisant le circuit approprié de l'hémisphère gauche.

Pour le savoir, les chercheurs ont demandé à des adultes d'apprendre un système d'écriture inhabituel, constitué d'élégantes courbes. Les sujets recevaient d'abord un entraînement avec 16 mots, puis on scannait leur cerveau tandis qu'ils tentaient de lire ces 16 mots anciens et 16 autres mots nouveaux. Cependant, à leur insu, on manipulait également leur attention ([figure 28](#)). À la moitié d'entre eux, on disait qu'il fallait faire attention aux courbes dans leur ensemble, car chacune d'elles, tel un caractère chinois, correspondait à un mot. À l'autre groupe, on expliquait qu'en réalité, les pelotes de courbes étaient formées de lettres superposées : chaque mot était fait de trois lettres et ils devaient prêter attention à chacune d'entre elles. Ainsi, le premier groupe faisait attention au niveau global et le second au niveau local, celui qui avait réellement servi à écrire les mots.

Résultat ? Les deux groupes parvenaient à mémoriser les 16 premiers mots, mais l'attention modifiait radicalement leur capacité de lire des mots nouveaux. Le second groupe, focalisé sur les lettres, réussissait à découvrir les correspondances entre les lettres et les sons, et parvenait à lire 78 % des mots nouveaux – et l'examen de leur cerveau montrait qu'ils activaient des régions latéralisées à l'hémisphère gauche, dans le circuit normal de la lecture. À l'inverse, dans le premier groupe, le fait de prêter attention à la forme globale des mots bloquait totalement l'apprentissage : ces volontaires ne parvenaient absolument pas à lire des mots nouveaux, et ils activaient un circuit tout à fait inapproprié, situé dans les aires visuelles de l'hémisphère droit.

Le message est clair : l'attention change radicalement l'activité cérébrale. Prêter attention à la forme globale du mot empêche de découvrir le code alphabétique et oriente l'activité cérébrale en direction d'un circuit inadéquat. Pour apprendre à lire, seul l'entraînement phonique, qui attire l'attention sur les correspondances entre les lettres et les sons, active le circuit de la lecture et permet l'apprentissage. Tous les enseignants de CP devraient connaître ces

données ! Ils mettraient alors tout en œuvre pour que l'enfant débutant comprenne qu'il doit orienter son attention vers les correspondances graphèmes-phonèmes. Lorsqu'un enfant apprend à lire en suivant chaque lettre du doigt et du regard, de la gauche vers la droite, l'apprentissage s'en trouve massivement facilité. Si, par contre, on ne donne pas à l'enfant les bonnes clés, et qu'il se contente d'examiner le mot écrit comme un objet dépourvu de toute structure interne, rien ne se passe, car il ne voit littéralement pas ce qu'on cherche à lui inculquer.

Bien enseigner, c'est donc avant tout choisir avec soin le niveau auquel on souhaite que l'enfant fasse attention, parce que c'est lui et lui seul qui sera représenté dans son cerveau avec toute la force nécessaire pour un apprentissage efficace. Les autres stimuli, les perdants de la compétition attentionnelle, ne seront pas appris.

## *Contrôle exécutif : l'aiguilleur du cerveau*

Le troisième et dernier système attentionnel, enfin, détermine *comment* les informations sélectionnées sont traitées. On le désigne sous le terme de « contrôle exécutif ». C'est un ensemble complexe de processus mentaux qui nous permettent de choisir un plan d'action et de nous y tenir<sup>208</sup>. Il implique toute une hiérarchie de régions situées principalement dans le cortex préfrontal – cette masse de cortex qui occupe un tiers de nos hémisphères cérébraux. Comparés aux autres primates, nos lobes frontaux sont gigantesques, mieux connectés, et abritent des neurones aux arborescences bien plus larges et complexes<sup>209</sup>. Ainsi, notre dispositif de supervision cortical est bien plus développé que chez les autres primates, notamment au plus haut niveau de la hiérarchie, celui qui nous permet d'organiser nos opérations mentales et de prendre conscience de nos erreurs : le contrôle exécutif<sup>210</sup>.

Imaginez que vous deviez calculer, mentalement, le produit de 23 par 8. C'est votre système de contrôle exécutif qui assure le bon déroulement de toute la série d'opérations mentales pertinentes, du début jusqu'à la fin : faire attention au chiffre de droite (3) et le multiplier par 8, puis stocker le résultat en mémoire (24), faire attention au chiffre de gauche (2), le multiplier aussi par 8 pour obtenir 16, se souvenir qu'il s'agit en réalité de la colonne des dizaines, donc de 160 et, enfin, additionner 160 et 24 pour obtenir le résultat final : 184.

Le contrôle exécutif agit comme un aiguilleur du cerveau : il pilote, dirige, gouverne, comme un chef de gare qui, en orientant l'aiguillage, parvient à mener un train sur la voie désirée. C'est encore un système de sélection mentale, mais parmi les opérations mentales disponibles plutôt que parmi les stimuli qui nous parviennent. Ainsi, l'attention spatiale et l'attention exécutive se complètent. L'attention spatiale parcourt la page du manuel de mathématiques et donne un coup de projecteur au problème  $23 \times 8$  – mais c'est l'attention exécutive qui guide le projecteur. C'est elle qui va le focaliser d'abord sur le 3 et le 8, puis envoyer ces chiffres vers les circuits de la multiplication, et ainsi de suite. Elle active les opérations pertinentes et inhibe les actions inappropriées. Elle surveille en permanence le bon déroulement du programme et décide quand changer de stratégie. C'est également elle qui, dans des circuits spécialisés du cortex cingulaire, repère que nous sommes en train de nous tromper ou de dévier du but, et corrige aussitôt le tir.

Il y a un lien étroit entre le contrôle exécutif et ce que l'on appelle la mémoire de travail. Suivre un algorithme mental et contrôler son exécution impliquent de conserver une mémoire de tous les éléments du programme en cours : résultats intermédiaires, étapes déjà effectuées, opérations restant à effectuer... En particulier, l'attention exécutive contrôle les entrées et les sorties de ce que j'ai appelé l'espace de travail neuronal global : une mémoire temporaire, consciente, au sein de laquelle nous pouvons faire entrer pratiquement n'importe quelle information qui nous paraît pertinente<sup>211</sup>. C'est le routeur du cerveau, l'aiguilleur qui décide comment, et dans quel ordre, envoyer les informations aux différents processeurs que comprend notre cerveau. À ce niveau, les opérations mentales sont lentes et organisées en série : ce système

traite les informations une par une, il n'est pas capable de faire deux opérations à la fois. Les psychologues parlent d'un goulot d'étranglement central.

Sommes-nous vraiment incapables de faire deux choses à la fois ? Nous avons parfois l'impression que notre pensée peut se diviser et suivre, simultanément, deux programmes distincts – mais c'est, là encore, une pure illusion. L'expérience le montre : donnez à une personne deux tâches très simples, par exemple appuyer avec la main gauche si elle entend un son aigu et appuyer avec la main droite si elle voit la lettre Y. Lorsque les deux cibles surviennent simultanément, la personne effectue la première tâche à vitesse normale, mais la seconde est ralentie en proportion directe du temps passé à prendre la première décision<sup>212</sup>. Le ralentissement se chiffre en centaines de millisecondes, et il se peut même que l'on rate totalement la deuxième cible. Mais nous n'en sommes pas conscients – car, par définition, nous ne pouvons prendre conscience que des informations qui pénètrent dans notre espace de travail conscient ! Le second stimulus attend devant la porte, jusqu'à ce que l'espace de travail soit libre – mais nous n'avons aucune introspection de ce temps d'attente, et si on nous le demande, nous pensons que le stimulus vient seulement d'apparaître<sup>213</sup>.

Ce n'est qu'avec un entraînement intense que nous parvenons, avec plus ou moins de bonheur, à faire deux choses complexes à la fois. L'automatisation libère l'espace de travail conscient. À force de pratique acharnée, un pianiste professionnel parviendra peut-être à jouer tout en parlant, ou une dactylo à saisir un texte tout en écoutant la radio. Toutefois, ce sont des exceptions rares et qui restent débattues en psychologie, car il se peut que l'attention exécutive fasse des sauts rapides et quasi indétectables d'une opération à l'autre<sup>214</sup>. La règle de base, c'est qu'en situation de « double tâche », c'est-à-dire lorsqu'on demande de réaliser plusieurs opérations cognitives sous le contrôle de l'attention, l'une des opérations au moins est ralentie ou abolie.

En bref, apprendre à se concentrer est un ingrédient essentiel de l'apprentissage. On ne peut pas exiger d'un enfant ou d'un adulte qu'il apprenne simultanément deux choses à la fois. Apprendre requiert d'accorder la priorité à une tâche précise. Toute distraction ralentit ou anéantit nos efforts : nous

perdons le fil de nos pensées. Les sciences cognitives rejoignent ici l'expérimentation dans les salles de classe – il a été montré, par exemple, qu'une classe trop décorée distrait l'enfant et l'empêche de se concentrer<sup>215</sup> .

## *Apprendre à faire attention*

L'attention exécutive correspond approximativement à ce que nous appelons la concentration ou le contrôle de soi. Or ce système ne se met pas en place immédiatement chez l'enfant. Le contrôle exécutif se développe lentement, tout au long de l'enfance et de l'adolescence. Nous apprenons progressivement à nous contrôler, c'est-à-dire à privilégier les stratégies appropriées et à inhiber les stratégies inadéquates, en évitant toute distraction.

La psychologie cognitive regorge de tâches où l'on voit l'enfant surmonter progressivement ses erreurs, à mesure qu'il augmente sa capacité à se concentrer et à inhiber les stratégies inappropriées. Le psychologue Jean Piaget avait été le premier à le remarquer : les très jeunes enfants font parfois des erreurs stupides. Si, par exemple, on cache plusieurs fois un objet à l'endroit A, puis qu'on le cache à l'endroit B, un bébé de moins de 1 an continuera d'aller le chercher au point A – c'est la fameuse « erreur A, non B ». Piaget en avait conclu que le bébé ne maîtrise pas la permanence de l'objet (le fait de savoir que l'objet continue d'exister lorsqu'il est caché), mais nous savons aujourd'hui que ce n'est pas la bonne interprétation du comportement de l'enfant. L'examen de son regard montre que bébé sait parfaitement représenter les objets cachés, au moins depuis l'âge de quelques mois. Par contre, il a du mal à résoudre les conflits mentaux : dans la tâche « A, non B », la routine des essais précédents lui dit d'aller chercher au point A, tandis que sa mémoire de travail, plus récente, lui dit d'aller chercher au point B – et avant 10 mois, c'est souvent l'habitude qui prime. À cet âge, c'est donc le contrôle exécutif qui manque, pas la connaissance. L'erreur « A, non B » disparaît vers 12 mois, en lien direct avec le développement du cortex préfrontal<sup>216</sup> .



Autre erreur de l'enfant : confondre le nombre et la taille. Là encore, Jean Piaget a fait une découverte essentielle, mais s'est trompé sur son interprétation. Il a constaté que les jeunes enfants, avant environ 3 ans, ont du mal à juger le nombre d'un ensemble d'objets : si on leur montre deux rangées égales de billes, puis qu'on se contente d'espacer l'une d'entre elles,

oooooooo oooooo → o o o o o o oooooo

les enfants disent que le nombre a changé, et que la rangée la plus longue comprend le plus grand nombre d'objets. Mais, contrairement à ce que Piaget pensait, cela ne signifie pas que les enfants de cet âge ne maîtrisent pas la « conservation du nombre » : comme nous l'avons vu, même des bébés nouveaux-nés possèdent déjà un sens abstrait du nombre, indépendamment de l'espacement ou même de la modalité sensorielle. Non, la difficulté vient, une fois de plus, du contrôle exécutif. L'enfant doit apprendre à inhiber un trait saillant (la taille) et à amplifier un trait moins évident (le nombre). L'expérience montre que même les adultes se trompent : il n'est pas facile de dire qu'un nuage de points comprend moins d'objets, lorsque ces objets sont particulièrement grands et étalés dans l'espace, et nous avons beaucoup de mal à choisir le plus grand nombre parmi 7 et 9. Ce qui se développe avec l'âge et l'éducation, ce n'est pas tant la précision intrinsèque du système numérique que la capacité à l'utiliser avec efficacité et concentration, en filtrant les indices qui ne sont pas pertinents<sup>217</sup> – et l'entraînement, une fois de plus, corrèle avec le développement des réponses neuronales dans le cortex préfrontal<sup>218</sup>.

Je pourrais multiplier les exemples : à tous les âges de la vie, dans tous les domaines du savoir (cognitifs ou émotionnels), c'est notre capacité de contrôle exécutif qui, en se développant, nous permet d'éviter les erreurs<sup>219</sup>. Nommez, sans vous tromper, *la couleur* des mots qui suivent :

chien maison bien car sofa trop rouge noir vert bleu gris...

Votre ralentissement, dans la seconde moitié de la liste, traduit l'entrée en scène du contrôle exécutif, qui doit apprendre à inhiber la lecture des mots pour se concentrer sur leur couleur.

Maintenant, résolvez le problème suivant : « Marie possède vingt-cinq billes, c'est cinq de plus que Grégoire. Combien de billes Grégoire possède-t-il ? » Avez-vous ressenti l'envie d'additionner les deux nombres ? L'énoncé du problème emploie le mot « plus », alors qu'il faut soustraire – c'est un piège dans lequel tombent bien des enfants, avant de parvenir à se contrôler et à se forcer à réfléchir au sens profond d'un problème de maths ou de logique.

L'attention et le contrôle exécutif se développent spontanément avec la maturation progressive du cortex préfrontal, qui s'étend sur les deux premières décennies de notre vie. Mais ce circuit, comme tous les autres, est plastique, et de nombreuses recherches montrent que son développement peut être accéléré par l'entraînement et l'éducation<sup>220</sup>. Comme c'est un système partagé entre des tâches cognitives très diverses, de nombreuses activités, y compris les plus ludiques, développent le contrôle exécutif. Michael Posner a été l'un des premiers à concevoir des jeux pour ordinateurs qui améliorent la capacité de concentration des très jeunes enfants, par exemple en les forçant à suivre la direction d'un objet situé au centre de l'écran, sans se laisser distraire par ses voisins :



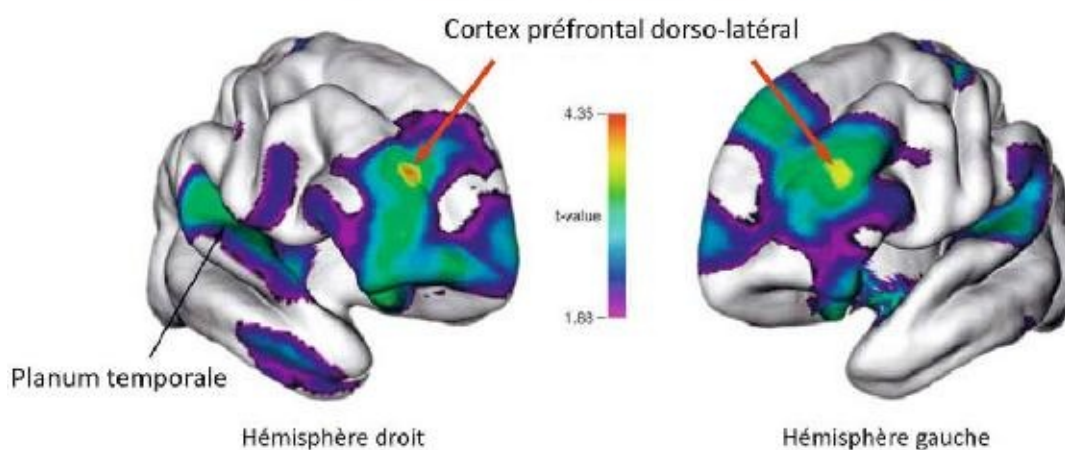
Bien avant l'invention de l'ordinateur, la médecin et pédagogue italienne Maria Montessori avait déjà noté à quel point les activités pratiques peuvent développer la concentration. Dans les écoles Montessori d'aujourd'hui, on demande par exemple à un enfant de marcher sur une ellipse tracée au sol sans

jamais quitter cette ligne du pied. Lorsque l'enfant y parvient, on corse la difficulté en le faisant recommencer avec une cuillère dans la bouche, puis avec une balle de ping-pong dans la cuiller, et ainsi de suite. Cette « méthode Montessori » a démontré ses mérites<sup>221</sup>. D'autres études montrent les bénéfices des jeux vidéo, de la méditation, de la pratique d'un instrument de musique... Il est si difficile, pour un enfant, de contrôler tout son corps, son regard, sa respiration, en coordonnant ses gestes avec ceux des autres – c'est pourquoi la pratique d'un instrument de musique, dès le plus jeune âge, a des effets importants sur les circuits attentionnels du cerveau, y compris une augmentation importante et bilatérale de l'épaisseur du cortex préfrontal (figure 29)<sup>222</sup>.

L'entraînement du contrôle exécutif modifie même le quotient intellectuel. Ce qu'on appelle l'intelligence fluide – la capacité à raisonner et à résoudre des problèmes nouveaux – fait massivement appel aux systèmes cérébraux de contrôle exécutif : tous deux mobilisent un réseau similaire d'aires cérébrales, notamment le cortex préfrontal dorso-latéral<sup>223</sup>. Pas étonnant, d'ailleurs, puisque les tests standardisés qui mesurent l'intelligence fluide ressemblent à ceux qui évaluent le contrôle exécutif : tous deux privilégient l'attention, la concentration, la capacité de passer rapidement d'une activité à l'autre, sans perdre de vue le but global de la tâche. Et de fait, l'entraînement de la mémoire de travail et du contrôle exécutif fait légèrement bouger les mesures de l'intelligence fluide<sup>224</sup>. Ces résultats concordent avec le fait que, même si l'intelligence n'est pas dépourvue de déterminants génétiques, elle varie également en fonction de facteurs environnementaux, et notamment l'éducation. Et ces effets sont massifs : le QI peut augmenter de 20 points lorsqu'un enfant est adopté dans une famille de haut niveau socio-économique, et on estime que le gain lié à l'éducation atteint entre 1 et 5 points de QI par année d'éducation supplémentaire<sup>225</sup>.

La frontière actuelle de la recherche consiste à optimiser ces effets de l'entraînement cognitif et à en préciser les limites. Comment augmenter la durée des effets observés ? Et surtout, comment faire en sorte qu'ils ne soient pas limités à la tâche spécifique qu'on a entraînée, mais qu'ils s'étendent bien au-delà, et ce tout au long de la vie ? Tel est le défi, car il est certain que, par défaut, le cerveau tend à développer des astuces spécifiques à chaque tâche, au cas par

cas. Il faut donc diversifier les apprentissages. Les meilleurs résultats semblent être obtenus par les programmes éducatifs qui, par des activités variées, stimulent le noyau de compétences cognitives que constituent la mémoire de travail et l'attention exécutive.



**Figure 29.** L'attention exécutive, c'est-à-dire la capacité de se concentrer et de se contrôler, se développe avec l'âge et l'éducation. La pratique d'un instrument de musique est l'une des nombreuses manières d'apprendre à réguler, dès le plus jeune âge, la concentration et le contrôle de

soi. Par rapport aux non-musiciens, toutes choses égales par ailleurs, les musiciens présentent des augmentations de l'épaisseur du cortex préfrontal dorso-latéral, qui joue un rôle important dans le contrôle exécutif. L'échelle de couleur indique l'ampleur de la différence entre les musiciens et les non-musiciens.

Certains résultats rendent optimiste. L'entraînement précoce de la mémoire de travail, particulièrement s'il démarre dès la maternelle, semble avoir des effets positifs sur la concentration et la réussite dans de nombreux domaines, y compris ceux qui sont les plus directement pertinents pour l'école : la lecture et les mathématiques<sup>226</sup>. Ce n'est d'ailleurs pas étonnant, puisque l'on sait depuis des années que la mémoire de travail est l'un des meilleurs prédicteurs de la réussite en arithmétique des années plus tard<sup>227</sup>. Les effets de ces entraînements sont démultipliés si l'on combine un travail sur la mémoire avec un enseignement plus direct du concept de « ligne numérique » – l'idée essentielle que les nombres s'organisent sur un axe linéaire, et qu'additionner ou soustraire consiste à se déplacer sur cette ligne<sup>228</sup>. Toutes ces interventions éducatives semblent bénéficier le plus aux enfants de milieu défavorisé. Pour les familles de bas niveau socio-économique, intervenir le plus tôt possible, dès la maternelle, et enseigner les fondamentaux de l'apprentissage et de l'attention peuvent être l'un des meilleurs investissements éducatifs.

## *Je fais attention si tu fais attention*

« ὁ ἄνθρωπος φύσει πολιτικὸν ζῷον »

« L'homme est par nature un animal social. »

ARISTOTE (384-322 av. J.-C.).

Toutes les espèces de mammifères, et bien sûr tous les primates, possèdent un système attentionnel. Mais l'attention, chez les êtres humains, présente une caractéristique singulière qui accélère encore les apprentissages : la prise en

compte du contexte social. Chez *Homo sapiens*, plus que chez tout autre primate, l'attention et l'apprentissage dépendent de signaux sociaux : je fais attention là où tu fais attention, et j'apprends ce que tu m'enseignes.

Dès le plus jeune âge, l'enfant suit l'adulte du regard. Il prête une attention extrême aux yeux des personnes avec qui il interagit. Dès qu'on lui dit quelque chose, son premier réflexe n'est pas d'explorer la scène, mais de croiser le regard de la personne qui lui parle. Ce n'est qu'une fois le regard accroché que l'enfant se tourne dans la direction qu'examine l'adulte. Cette capacité de partage social, qu'on appelle l'attention partagée, détermine ce que l'enfant apprend.

Je vous ai déjà parlé des expériences où l'on fait apprendre un mot nouveau comme « wog » à un bébé : si l'enfant interagit avec la personne qui parle et parvient à suivre son regard, il n'a aucune difficulté à en apprendre le sens, en seulement quelques essais – mais si le même mot tombe d'un haut-parleur, aucun apprentissage ne se produit<sup>229</sup>. Même chose pour l'apprentissage des catégories phonétiques : un enfant américain qui, vers l'âge de 9 mois, interagit avec une nounou chinoise pendant quelques semaines apprend les phonèmes du chinois – mais s'il reçoit exactement la même quantité de stimulation linguistique d'une vidéo de très haute qualité, aucun apprentissage ne se produit<sup>230</sup>.

Les psychologues hongrois Gergely Csibra et György Gergely postulent que l'apprentissage en société constitue une adaptation évolutive fondamentale de l'espèce humaine<sup>231</sup>. Animal social, *Homo sapiens* dispose d'un « module pédagogique » qui se déclenche dès que nous faisons attention à ce que les autres cherchent à nous enseigner. Notre succès planétaire, nous le devons, au moins en partie, à une évolution spécifique : notre capacité à partager l'attention avec les autres. La plupart des informations que nous apprenons, nous les devons aux autres plutôt qu'à notre expérience personnelle. Ainsi, la culture collective de l'espèce humaine s'élève bien au-delà de ce que chacun peut découvrir seul. C'est ce que le psychologue Michael Tomasello appelle l'effet de « cliquet culturel » : comme un cliquet sur un ascenseur, le partage des connaissances empêche la culture de redescendre. Lorsqu'une personne fait une découverte,

tout le groupe en bénéficie rapidement. Grâce à l'apprentissage social, il est bien rare que l'ascenseur culturel redescende et qu'une invention majeure sombre dans l'oubli.

Notre système attentionnel s'est adapté à ce contexte culturel. Les recherches de Gergely et Csibra montrent que, dès le plus jeune âge, l'orientation de l'attention de l'enfant tient compte des signaux que lui envoient les adultes. La présence d'un tuteur humain module massivement l'apprentissage. Non seulement elle attire l'attention de l'enfant, mais elle lui signale que l'adulte a l'intention de lui enseigner un point important. Et le bébé y est immédiatement sensible : le contact oculaire induit en lui une « posture pédagogique » qui l'induit à penser que l'information est importante et généralisable.

Prenons un exemple : une jeune femme se tourne vers un objet A en faisant la grimace, puis vers un objet B avec un grand sourire. Un bébé de 18 mois regarde la scène. Quelle conclusion va-t-il en tirer ? Tout dépend des signaux que l'enfant et l'adulte ont échangés. Si leurs regards ne se sont pas croisés, alors l'enfant se souvient juste d'une information ponctuelle : cette personne n'aime pas l'objet A, elle préfère l'objet B. Mais s'il y a eu contact oculaire, alors l'enfant en déduit beaucoup plus : il retient que l'objet A est mauvais, et que l'objet B est agréable, pas seulement pour cette personne-ci, mais pour tous (figure 30 ). Le cerveau de l'enfant prête une attention extrême aux indices ostensibles de communication volontaire. Si quelqu'un le regarde, il en conclut que cette personne cherche à lui enseigner des informations cruciales, pas juste des préférences anecdotiques.

Il n'y a d'ailleurs pas que le regard : les enfants comprennent également très vite ce que signifie pointer du doigt (alors que les chimpanzés ne comprennent jamais vraiment ce geste). Même un bébé se rend compte si quelqu'un cherche à attirer son attention pour lui transmettre une information importante. Lorsqu'un bébé de 9 mois voit quelqu'un qui attire son attention, puis qui pointe vers un objet, il se souvient de l'identité de cet objet, parce qu'il comprend que c'est une information que son interlocuteur juge importante – alors que s'il voit la même

personne tendre le bras vers l'objet, sans regarder l'enfant, celui-ci ne se souvient que de la position, pas de l'identité de l'objet<sup>232</sup> .

Parents, enseignants, gardez constamment à l'esprit ce fait capital : votre attitude et votre regard changent tout pour l'enfant. S'attacher l'attention de l'enfant par le contact visuel et verbal, c'est garantir que celui-ci partage votre attention, et multiplier d'autant les chances qu'il retienne l'information que vous cherchez à lui transmettre.



### Attention partagée

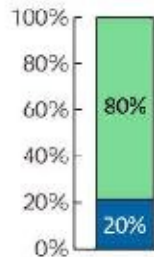


### Compréhension des intentions

Un adulte, les mains **occupées**, fait une action étrange avec la tête.



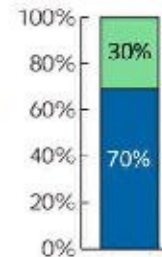
80 % des enfants imitent intelligemment de la **main**.



Un adulte, les mains **libres**, fait une action étrange avec la tête.



70 % des enfants imitent servilement de la **tête**.



**Figure 30.** Dans l'espèce humaine, l'apprentissage est social : il dépend fortement de l'attention et de la compréhension des intentions des autres. Même un bébé de 18 mois comprend que, si on le regarde dans les yeux, c'est qu'on cherche à lui transmettre une information importante. Il apprend alors plus efficacement et parvient à généraliser à une autre personne, alors qu'il n'y parvient pas si on ne l'a pas regardé (en haut). Dès 14 mois, un bébé interprète déjà les intentions de son entourage : après avoir vu une personne allumer une lumière en l'appuyant de la tête, il va imiter son geste en tout point, sauf si la personne avait les mains occupées, auquel cas il comprend que lui peut simplement appuyer de la main (en bas).

## *Enseigner, c'est faire attention à l'attention de l'autre*

Aucune autre espèce animale ne sait enseigner comme nous le faisons. La raison en est simple : nous sommes les seuls à posséder une « théorie » de l'esprit des autres, la capacité de faire attention à leur attention, d'imaginer leurs pensées – y compris ce qu'ils pensent que les autres pensent, et ainsi de suite, dans une boucle infinie. Ce type de représentations récursives, enchâssées les unes dans les autres, est typique du cerveau humain et joue un rôle essentiel dans la relation pédagogique. L'adulte qui enseigne doit constamment penser à ce que l'enfant ne sait pas : il adapte ses mots et choisit ses exemples afin de transformer, au plus vite, la connaissance de l'enfant. Inversement, l'enfant sait que l'adulte sait qu'il ne sait pas : une fois qu'il est engagé dans la « posture pédagogique », il interprète chaque acte de l'enseignant comme une tentative de lui transférer des connaissances. Et la boucle continue à l'infini : l'adulte sait que l'enfant sait que l'adulte sait qu'il ne sait pas... ce qui permet à l'adulte de choisir ses exemples en sachant que l'enfant va tenter de les généraliser.

Cette relation pédagogique n'existe chez aucune autre espèce que la nôtre. Dans un article retentissant<sup>233</sup>, publié dans *Science* en 2006, certains chercheurs en éthologie ont bien décrit une forme d'enseignement chez le suricate, un petit mammifère de la famille des mangoustes qui vit en Afrique du Sud – mais à mon sens, c'est la définition même d'enseignement qu'ils ont dévoyée. De quoi s'agit-il ? De la grande affaire des familles : apprendre à préparer à manger ! La difficulté à laquelle les suricates sont confrontés, c'est qu'ils consomment des proies extrêmement dangereuses : des scorpions au dard mortel. Or l'article démontre que les suricates adultes aident leurs petits à apprendre en leur proposant de la nourriture « préparée », en l'occurrence des scorpions dont l'adulte a ôté le dard. À mesure que le petit grandit, l'adulte lui fournit une proportion de plus en plus grande de scorpions vivants, et cela aide visiblement le petit à devenir un chasseur autonome. Ainsi, selon les auteurs, trois critères de l'enseignement sont remplis : l'adulte effectue, en présence des petits, un

comportement spécifique ; ce comportement présente un coût pour l'adulte ; et le petit en bénéficie en acquérant des connaissances plus vite que si l'adulte n'était pas intervenu.

Le cas des suricates est étonnant : l'évolution a mis en place un mécanisme singulier qui, de toute évidence, facilite la survie des mangoustes. Mais est-ce vraiment de l'enseignement ? À mon sens, non. Ces données ne permettent pas de conclure que les suricates enseignent vraiment à leurs petits, car il y manque un ingrédient crucial : l'attention partagée. Il n'y a aucune preuve que les adultes font attention à ce que savent les petits ou, inversement, que les petits tiennent compte de la posture pédagogique des adultes. Les mangoustes adultes se contentent de présenter aux enfants, en fonction de leur âge, des animaux plus ou moins dangereux, mais rien ne prouve qu'elles prêtent attention au savoir de l'enfant. Pour autant que l'on sache, il pourrait s'agir d'un comportement entièrement précâblé, spécifique à l'apprentissage de la consommation des scorpions – une conduite complexe, mais bornée, comparable à la fameuse danse des abeilles ou à la parade nuptiale des flamants roses.

Mangoustes et scorpions nous enseignent, comme en miroir, tout ce que l'enseignement possède d'unique et de précieux dans notre espèce. Toute véritable relation pédagogique implique un lien mental fort entre l'enseignant et l'élève. Le bon pédagogue se construit un modèle mental de son élève, de ses compétences et de ses erreurs, et il fait de son mieux pour le faire progresser. Cette définition idéale exclut l'enseignant (humain ou ordinateur) qui se contente de délivrer mécaniquement une leçon stéréotypée, sans l'avoir optimisée pour tenir compte des attentes et des savoirs de son auditoire – cela ne fonctionne pas. L'élève, pour sa part, sait que l'enseignant sait, et qu'il fait de son mieux pour transmettre son savoir. Toute relation pédagogique saine doit se fonder sur l'attention, l'écoute, le respect et la confiance, dans les deux sens.

La modeste pédagogie des mangoustes ne rend également pas justice au rôle que joue l'éducation dans les sociétés humaines. « Chaque homme est une humanité, une histoire universelle », affirme Jules Michelet. Grâce à l'éducation, nous emportons en nous le meilleur des pensées des générations humaines qui nous précèdent. Chaque mot, chaque concept que nous apprenons est une petite

conquête que nous transmettent nos ancêtres. Sans langage, sans transmission culturelle, sans éducation en collectivité, jamais nous n'aurions pu découvrir seuls tous les outils mentaux dont nous sommes aujourd'hui dotés. Mais cette véritable dépendance d'*Homo sapiens* envers la communication sociale est une malédiction autant qu'un don. Revers de la médaille, c'est par sa faute que se propagent les utopies des gourous, les mythes des religions, les *fake news* des hommes politiques et les commérages des concierges. Dès le plus jeune âge, notre cerveau absorbe aveuglément les contes de notre entourage, qu'ils soient vrais ou faux. Dans un contexte social, il baisse la garde et cesse de se comporter comme un scientifique en herbe pour devenir un mouton de Panurge.

Une célèbre expérience révèle à quel point, dès le plus jeune âge, l'apprentissage peut nous rendre esclaves de notre environnement social. Dès 14 mois, les bébés imitent servilement les actions d'une personne, même si elles n'ont aucun sens<sup>234</sup>. S'ils voient quelqu'un, les mains occupées, appuyer sur un bouton avec la tête, la raison leur dicte qu'il suffit d'appuyer avec la main – et c'est ce qu'ils font. Mais, s'ils voient la même personne appuyer de la tête sans raison particulière, puisque ses mains sont libres et parfaitement visibles, alors ils abdiquent tout raisonnement et font simplement confiance à l'adulte – et ils se mettent à imiter fidèlement l'action qu'ils ont vue, même si elle est dépourvue de sens (figure 30). À l'âge adulte, ce conformisme social persiste et s'amplifie : si nos congénères portent un jugement différent du nôtre, même dans des situations perceptives les plus banales, nous révisons fréquemment notre jugement afin de l'aligner avec le leur, même si c'est au mépris de toute plausibilité<sup>235</sup>.

En bref, notre cerveau d'*Homo sapiens* dispose de deux modes d'apprentissage : un mode actif, dans lequel nous mettons à l'épreuve, en bons scientifiques, des hypothèses sur le monde extérieur ; et un mode réceptif, dans lequel nous absorbons ce que d'autres nous transmettent sans l'avoir personnellement vérifié. C'est le second mode qui, par accumulation culturelle, a permis l'expansion des sociétés humaines depuis cinquante mille ans. Mais sans l'esprit critique qui caractérise le premier mode, le second est vulnérable à la propagation de toutes les *fake news*. La vérification active des connaissances, le rejet du simple ouï-dire, la construction personnelle du sens sont des filtres

essentiels pour nous protéger des légendes et des gourous. Il nous faut donc trouver un compromis entre les deux modes : des étudiants attentifs, confiants dans les autres, mais également autonomes, capables d'esprit critique, acteurs de leur propre apprentissage.

C'est de ce deuxième pilier de l'apprentissage que nous allons discuter à présent : l'engagement actif.

# CHAPITRE 8

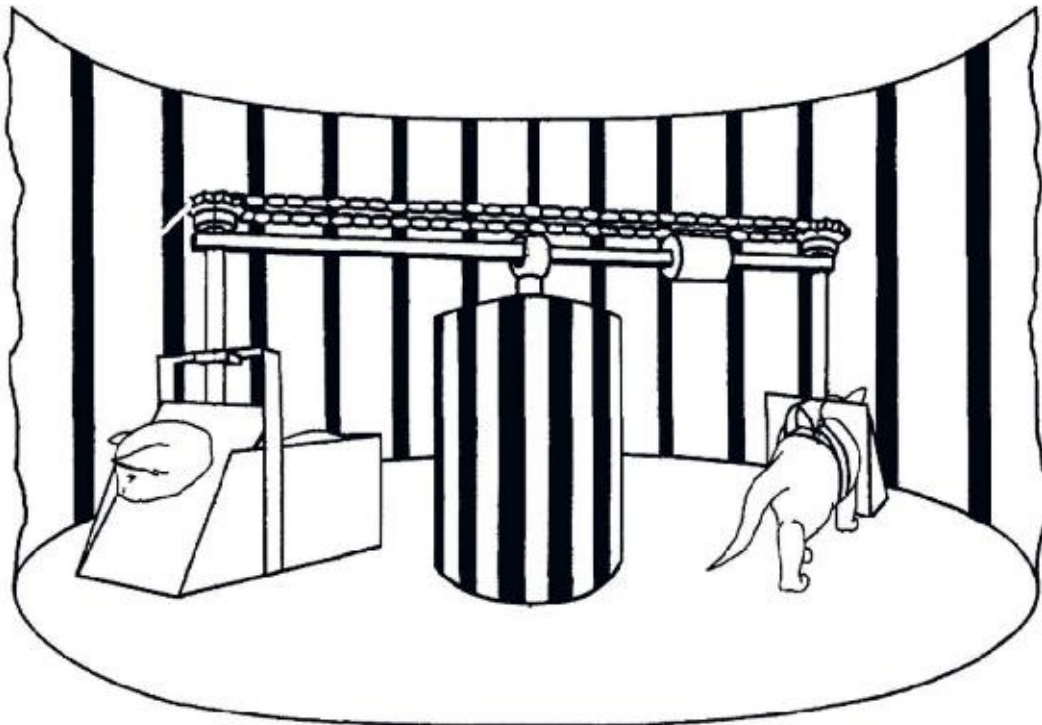
## L'engagement actif

---

Prenez deux chatons. Donnez au premier un collier. Quant au second, placez-le dans une nacelle. Enfin, reliez-les à un manège (figure 31). Ce dispositif assure que les mouvements des deux chatons sont rigoureusement identiques, l'un étant actif et l'autre passif. Le premier explore l'environnement de son propre chef, tandis que le second bouge exactement de la même manière, mais sans contrôle.

Telle est l'expérience classique du manège qu'ont menée Richard Held et Alan Hein en 1963 – à une époque où l'éthique de l'expérimentation animale n'était clairement pas aussi développée qu'aujourd'hui ! Pendant quelques semaines, les deux chatons vécurent, trois heures par jour, dans un grand cylindre tapissé uniquement de barres verticales. Tous deux y recevaient exactement les mêmes entrées visuelles. Et pourtant, leur développement fut très différent, et cette expérience simple a conduit à une découverte fondamentale : l'exploration active du monde est essentielle au bon développement de la vision<sup>236</sup>. Malgré la pauvreté de son environnement, constitué uniquement de barres verticales, le chaton actif développe rapidement une vision normale. Le chaton passif, par contre, perd ses capacités visuelles et, à la fin de l'expérience,

échoue dans des tests élémentaires d'exploration visuelle. Dans le test de la falaise, par exemple, on place l'animal sur un pont qu'il peut quitter soit d'un côté où se trouve une haute falaise, soit d'un côté peu profond. L'animal normal n'hésite pas une seconde et saute du côté facile. L'animal élevé dans la passivité, par contre, choisit au hasard. D'autres tests montrent qu'il ne perçoit pas bien l'espace et ne l'explore pas de la patte, comme le fait n'importe quel félin.



**Figure 31.** L'engagement actif est le second pilier de l'apprentissage : un organisme passif n'apprend pas. Dans l'expérience classique de Held et Hein (1963), un chaton explore activement le monde tandis que l'autre ne reçoit que l'empreinte passive des sens. Après quelques dizaines d'heures, l'animal passif présente des troubles prononcés de l'exploration visuelle. Apprendre efficacement, c'est refuser la passivité, s'engager, explorer avec curiosité, générer activement des hypothèses et les mettre à l'épreuve.

*Un organisme passif n'apprend pas*

L'expérience du manège de Held et Hein est la métaphore de notre deuxième pilier de l'apprentissage : l'engagement actif. De très nombreux résultats, dans les domaines les plus divers, suggèrent qu'un organisme passif n'apprend pas ou très peu. Apprendre efficacement, c'est refuser la passivité, s'engager, explorer, générer activement des hypothèses.

Pour apprendre, notre cerveau doit d'abord se former un modèle mental hypothétique du monde extérieur et, ensuite seulement, le projeter sur son environnement et le mettre à l'épreuve en comparant ses prédictions avec la réalité sensorielle. Cela implique une posture active, engagée, attentive. La motivation est essentielle : on n'apprend bien que si on a une idée claire du but à atteindre, et qu'on adhère pleinement à cet objectif.

Ne vous méprenez pas : l'engagement actif ne signifie pas que l'enfant doive gigoter en classe ! J'ai visité un jour un centre de formation où le directeur, avec une certaine fierté, m'a expliqué comment il appliquait mes idées : il avait équipé tous les pupitres de pédaliers afin que ses élèves restent actifs pendant le cours de maths... Ce n'est pas du tout ce que je veux dire – et nous touchons la limite de la métaphore de l'expérience du manège. Actif ne veut pas dire que l'organisme doit bouger. C'est dans la tête, et non dans les pieds, que l'engagement actif fait son œuvre. Le cerveau n'apprend bien que s'il est attentif, concentré et en pleine activité de génération de modèles mentaux. Pour mieux digérer les faits à apprendre, un étudiant actif les reformule sans cesse en mots ou en pensées qui font sens pour lui. Un étudiant passif ou, pire, distrait, ne profite en rien d'une séance d'apprentissage, car son cerveau ne met pas à jour ses modèles mentaux du monde. Cet engagement n'a rien à voir avec un mouvement visible de l'extérieur : les deux étudiants peuvent très bien rester immobiles, mais c'est le mouvement de la pensée qui les distingue. L'un suit activement le cours, l'autre décroche.

L'expérience montre que rares sont les apprentissages où le cerveau se contente d'accumuler passivement les statistiques de ses entrées. Cela peut arriver, mais seulement au plus bas niveau des systèmes sensoriels ou moteurs. Je vous ai parlé, par exemple, de ces expériences où un enfant entend des centaines de syllabes et finit par y détecter la présence de mots comme « bi-be-



ron » – ce type d'apprentissage, implicite, semble subsister durant le sommeil<sup>237</sup>. Mais c'est l'exception qui confirme la règle : dans l'immense majorité des cas, et dès que l'apprentissage concerne des propriétés de haut niveau cognitif, telles que la mémoire explicite du sens des mots et pas seulement de leur forme, l'apprentissage ne survient que si l'apprenant fait attention, réfléchit, anticipe, avance des hypothèses, au risque de se tromper. Sans attention, sans effort, sans profondeur de la réflexion, la leçon s'évanouit sans laisser de traces dans le cerveau.

## *Approfondir pour mieux apprendre*

Prenons un exemple : l'effet de la profondeur de traitement des mots. Imaginez que je présente soixante mots à trois groupes d'étudiants. Aux uns, je demande de juger si ces mots sont écrits en minuscules ou en majuscules ; aux deuxièmes, s'ils riment avec « chaise » ; et aux troisièmes, s'il s'agit de noms d'animaux ou pas. Ensuite, à l'impromptu, je leur fais passer un test de mémoire. Résultat : la mémoire des mots est bien meilleure dans le troisième groupe, qui traitait les mots en profondeur, au niveau du sens (75 % de réussite), que dans les deux autres groupes, qui traitaient soit la surface du mot écrit (33 % de réussite), soit celle du mot parlé (52 % de réussite)<sup>238</sup>. On trouve certes une faible trace implicite des soixante mots dans tous les groupes – un peu d'apprentissage a laissé son empreinte subliminale dans les systèmes orthographiques et phonologiques. Cependant, seul le travail en profondeur induit une mémoire explicite, détaillée, des mots perçus. Le même phénomène survient au niveau des phrases : faire l'effort de les comprendre soi-même, sans que l'enseignant donne la solution, entraîne une bien meilleure rétention des informations en mémoire<sup>239</sup>. C'est une règle générale, que le psychologue américain Henry Roediger énonce ainsi<sup>240</sup> : « Rendre les conditions d'apprentissage plus difficiles, ce qui oblige les étudiants à un surcroît d'engagement et d'effort cognitif, conduit souvent à une meilleure rétention. »

L'imagerie cérébrale permet de comprendre pourquoi un traitement actif de mots, en profondeur, les imprime mieux en mémoire<sup>241</sup> : il oblige à activer les aires du cortex préfrontal, qui sont associées au traitement conscient des mots – et ces régions forment des boucles puissantes avec l'hippocampe, qui stocke les informations sous la forme de souvenirs épisodiques et explicites.

Dans un court métrage devenu un film culte, *La Jetée* (1962) du réalisateur français Chris Marker, une voix off énonce un aphorisme qui sonne comme une vérité profonde : « Rien ne distingue les souvenirs des autres moments : ce n'est que plus tard qu'ils se font reconnaître, à leurs cicatrices. » Bel adage... mais faux proverbe car, dès l'encodage, les événements de notre vie qui resteront gravés dans notre mémoire sont déjà différents de ceux qui ne laisseront aucune trace : ils font l'objet d'un traitement plus actif et plus profond. Lorsqu'on scanne une personne alors qu'elle est simplement exposée à une liste de mots ou d'images, on peut prédire lesquels seront retenus une demi-heure plus tard : ce sont ceux qui ont induit un maximum d'activité dans le cortex frontal, l'hippocampe et les régions voisines du cortex parahippocampique<sup>242</sup> . L'engagement actif de ces régions, reflet direct de la profondeur de pénétration des mots et des images dans le cerveau, prédit la force du souvenir. Une image inconsciente entre dans les aires sensorielles, mais elle ne crée qu'une toute petite vague d'activité dans la profondeur du cortex préfrontal. Seuls l'attention, la concentration, la prise de conscience et le traitement en profondeur maximisent la mémorisation<sup>243</sup> .

Prenons un autre exemple : l'apprentissage de la physique. Des étudiants de licence doivent apprendre les concepts abstraits de moment cinétique angulaire et de couple moteur. Séparons-les en deux groupes : aux uns, on donne dix minutes d'expérimentation avec une roue de bicyclette, et aux autres, dix minutes d'explications verbales et d'observations des autres étudiants. Résultat : l'apprentissage est bien meilleur dans le groupe qui bénéficie d'une interaction active avec les objets physiques<sup>244</sup> . Rendre l'enseignement plus profond, plus engageant est un gage de succès.

Une revue récente de plus de deux cents études pédagogiques le confirme : le cours magistral, où l'élève reste passif tandis que l'enseignant péroré pendant

cinquante minutes, a fait son temps<sup>245</sup>. Lorsqu'on compare ce cours magistral traditionnel aux pédagogies qui promeuvent l'engagement actif, l'effet est manifeste : dans toutes les disciplines, des maths à la psychologie en passant par la biologie ou l'informatique, un étudiant actif réussit mieux. Avec l'engagement actif, les résultats progressent d'un demi-écart-type, ce qui est considérable, et le taux d'échec diminue de près de 10 %. Mais quelles sont les stratégies qui sollicitent le plus les élèves ? Point ici de méthode unique et miraculeuse, mais toute une gamme d'activités qui obligent les étudiants à réfléchir par eux-mêmes : activités pratiques, discussions auxquelles chacun prend part, travaux en petits groupes, interruption d'un exposé par des questions difficiles auxquelles on laisse les étudiants réfléchir longuement... Toutes les solutions sont bonnes à condition de forcer les élèves à renoncer au confort de la passivité.

## *L'échec des pédagogies de la découverte*

Rien de tout cela n'est bien nouveau, me direz-vous, et de nombreux enseignants appliquent déjà spontanément ces idées. Toutefois, l'intuition ne suffit pas : encore faut-il vérifier scientifiquement ce qui marche et ce qui ne marche pas. Et c'est l'occasion pour moi de clarifier un point très important : l'idée que l'enfant doit être attentif, actif, engagé, acteur de son propre apprentissage – idée fondamentalement juste – ne doit pas être confondue avec le constructivisme ou les pédagogies de la découverte – une vision généreuse, séduisante, mais dont l'inefficacité a été cent fois démontrée. La distinction est fondamentale, mais rarement comprise, en partie du fait que l'on connaît également ces dernières sous le nom de « pédagogies actives », ce qui est source de nombreuses confusions.

Lorsqu'on parle de pédagogies de la découverte, de quoi s'agit-il ? D'une nébuleuse d'idées qui remontent à Jean-Jacques Rousseau et qui nous sont

parvenues par l'entremise de pédagogues tels que John Dewey, Ovide Decroly, Célestin Freinet, Maria Montessori ou, plus récemment, Jean Piaget et Seymour Papert. « Oserais-je exposer ici, dit Jean-Jacques Rousseau dans l'*Émile*, la plus grande, la plus importante, la plus utile règle de toute l'éducation ? Ce n'est pas de gagner du temps, c'est d'en perdre. » Pour Rousseau et ses successeurs, mieux vaut toujours laisser l'enfant découvrir par lui-même et construire son propre savoir, quitte à perdre des heures à tâtonner, explorer... Ce temps n'est pas perdu, estiment-ils : il est formateur d'un esprit autonome, qui pense de lui-même plutôt que de tout gober sans réfléchir, et qui sait résoudre de vrais problèmes plutôt que de réciter des solutions toutes faites. Écoutons encore Rousseau : « Rendez votre élève attentif aux phénomènes de la nature, bientôt vous le rendrez curieux ; mais pour nourrir sa curiosité, ne vous pressez jamais de la satisfaire. Mettez les questions à sa portée, et laissez-les-lui résoudre. »

La théorie est séduisante... Mais hélas, toute une série d'études, étagées sur plusieurs décennies, démontrent que la plus-value pédagogique de ces idées est proche de zéro. Laissés à eux-mêmes, les enfants éprouvent les plus grandes difficultés à découvrir les règles qui gouvernent un domaine, et ils apprennent beaucoup moins bien, voire pas du tout. Comment s'en étonner ? Ce que l'humanité a mis des siècles à découvrir, comment penser que l'enfant puisse le redécouvrir en quelques heures, sans aide extérieure ? Les échecs sont retentissants dans tous les domaines<sup>246</sup> :

- En lecture : la simple exposition aux mots écrits ne conduit à rien si l'on n'explique pas aux enfants la présence des lettres. Bien rares sont les enfants qui parviennent seuls à corréler l'écrit et l'oral et à découvrir que tous les mots qui commencent par le son *R* comprennent, à leur gauche, le symbole *R* ou *r*. Découvrir seul le principe alphabétique... La tâche serait démesurée si l'enseignant ne prenait pas la peine de choisir avec soin des exemples, des mots simples ou des lettres isolées qui facilitent la découverte.
- En mathématiques : on raconte qu'à l'âge de 7 ans, le génial mathématicien Carl Gauss découvrit, seul, comment additionner

rapidement tous les nombres de 1 à 100 (la solution est en note<sup>247</sup> ). Mais ce qui vaut peut-être pour Gauss ne s'applique pas à tous les autres enfants. La recherche est claire : en mathématiques, l'apprentissage marche mieux lorsque l'enseignant explique en détail un exemple, avant de laisser l'enfant se frotter à d'autres. Les enfants à qui on commence par expliquer comment résoudre un problème parviennent mieux à en résoudre d'autres que ceux qui ont découvert la solution par eux-mêmes.

- En informatique : dans son livre *MindStorms*, l'informaticien Seymour Papert explique pourquoi il a inventé le langage Logo (une sorte de tortue qui dessine et à qui le programme donne des ordres). Son idée était de laisser les enfants explorer l'informatique par eux-mêmes, sans instruction, en mettant d'emblée la main à la pâte. L'expérience fut un échec : au bout de quelques mois, les enfants ne savaient écrire que de petits programmes simples. Tous les concepts abstraits leur échappaient, et dans un test de résolution de problèmes, ils ne faisaient pas mieux que d'autres enfants non entraînés : le peu d'informatique qu'ils avaient appris ne se généralisait pas à d'autres domaines. Au contraire, un enseignement explicite, qui alternait des explications claires et leur mise à l'épreuve dans le langage Logo, leur permettait d'aller bien plus loin en informatique.

J'ai vécu moi-même l'arrivée du premier ordinateur personnel à la maison, alors que j'avais 15 ans. Je fais donc partie de cette génération qui a appris à programmer en langage Basic, sans professeur ni cours – quoique pas sans aide : mon frère et moi dévorions toutes les revues, manuels, livres et exemples sur lesquels nous parvenions à mettre la main. Je suis devenu un programmeur raisonnablement efficace... Mais quand je suis entré à l'université, en maîtrise d'informatique, j'ai pris conscience de l'énormité de mes lacunes : je bricolais sans avoir compris la structure logique profonde d'un programme, ni les bonnes pratiques qui le rendent clair et lisible. Et tel est sans doute le pire dans la pédagogie de la découverte : elle laisse l'étudiant dans l'illusion qu'il maîtrise son sujet, sans jamais lui donner les moyens d'accéder aux concepts profonds de sa discipline.

En résumé, le principe d'engagement actif énonce qu'il est crucial que l'étudiant soit motivé, actif, engagé dans l'apprentissage – mais cela ne signifie en rien qu'il doive être laissé à lui-même. L'échec du constructivisme le montre bien : il est tout aussi crucial que l'enseignant réponde à sa demande en lui fournissant un environnement d'apprentissage progressif, structuré, explicite, conçu pour le guider le plus vite possible vers les sommets. Le mieux est une pédagogie qui rende l'étudiant actif, mais qui soit étroitement guidée par l'enseignant – un enseignement structuré, doté d'une progression claire et rigoureuse, qui commence par les fondamentaux, vérifie leur maîtrise, et s'appuie sur eux pour construire une pyramide de sens. Et c'est d'ailleurs ce que font, aujourd'hui, la grande majorité des écoles Montessori : elles ne laissent pas l'enfant mariner sans rien faire, mais elles lui proposent toute une série d'activités rationnelles, hiérarchisées, qui font l'objet de démonstrations précises par l'enseignant avant d'être réalisées en autonomie par les enfants. Engagement, plaisir, autonomie, avec une pédagogie explicite appuyée sur un matériel stimulant : c'est la recette d'un cocktail gagnant dont l'efficacité a été démontrée<sup>248</sup> .

La pure pédagogie de la découverte, cette idée que l'enfant peut s'autoéduquer, fait partie des mythes éducatifs maintes fois dénoncés, mais qui demeurent curieusement populaires. Ce n'est pas le seul – au moins deux autres idées fausses lui sont également attachées<sup>249</sup> :

- Les enfants de la nouvelle génération, contrairement à leurs parents, seraient des champions du monde numérique. Baignant dans l'informatique depuis la plus tendre enfance, ils seraient devenus des *digital natives* , des *Homo zappiens* pour qui bits et octets n'ont plus de secrets. C'est faux : leur connaissance de la technologie est souvent superficielle. Seuls les plus éduqués savent critiquer les sources d'information sur Internet, et ils ne sont en aucun cas meilleurs pour réaliser plusieurs tâches à la fois (comme nous l'avons vu, le goulot d'étranglement central qui nous empêche de faire deux choses à la fois est une propriété fondamentale de l'architecture de notre cerveau).

- Chaque enfant possède son propre style d'apprentissage, il serait donc naturel de le laisser apprendre à sa guise. C'est faux : il n'existe aucune recherche qui démontre que certains enfants seraient plutôt visuels et d'autres auditifs ou tactiles (« sauf les aveugles, bien entendu ! », dirait Georges Brassens). Ce qui est vrai, c'est que certaines stratégies marchent mieux que d'autres – par exemple, il est plus facile de mémoriser une image qu'un mot. Cependant, c'est le cas pour tous les enfants. Il n'existe tout simplement aucune preuve que les enfants de type A apprendraient mieux avec la stratégie A, et les enfants de type B avec la stratégie B<sup>250</sup>.

## *Savoir piquer la curiosité*

« L'homme a naturellement passion de connaître. »

ARISTOTE , *Métaphysique* (vers – 335).

L'un des fondements de l'engagement actif, c'est la curiosité, l'envie d'apprendre, la soif de savoir. Piquer la curiosité des enfants, c'est avoir la partie à moitié gagnée : leur attention est mobilisée, leur esprit est déjà en recherche d'une explication – il ne reste plus qu'à les guider. Toutes choses égales par ailleurs, dès le jardin d'enfants, les élèves les plus curieux réussissent mieux en lecture ou en mathématiques<sup>251</sup>. Maintenir en éveil de la curiosité des enfants est donc l'un des facteurs clés d'une éducation réussie. Mais qu'est-ce au juste que la curiosité ? À quelle nécessité darwinienne répond-elle, à quelle sorte d'algorithme correspond-elle ?

Jean-Jacques Rousseau écrivait dans *Émile ou De l'éducation* « on n'est curieux qu'à proportion qu'on est instruit », et il se préoccupait de « rendre curieux » son élève. Là encore, il se trompait : la curiosité n'est pas un effet de l'instruction, une fonction que nous devons acquérir. Elle est présente dès le plus jeune âge et fait partie intégrante de notre biologie d'homme neuronal. C'est l'un des ingrédients de notre algorithme d'apprentissage. Nous n'attendons pas

simplement, de façon passive, que des informations nouvelles nous parviennent – comme le font encore, bêtement, tous les réseaux de neurones actuels, de simples fonctions d'entrée-sortie soumis à leur environnement. Comme le notait déjà Aristote, nous avons naturellement la passion de connaître : nous recherchons activement la nouveauté en explorant notre environnement pour y découvrir des choses à apprendre.

La curiosité est un ressort fondamental de l'organisme, un *drive*, disent les Américains : une force propulsive qui nous pousse à agir, au même titre que la faim, la soif, le besoin de sécurité ou l'envie de se reproduire. Quel rôle joue-t-elle dans la survie ? Les espèces comme la nôtre (mais aussi la plupart des mammifères et de nombreux oiseaux et poissons) ont intérêt à explorer leur environnement, afin de mieux le maîtriser. Il serait dangereux d'installer son nid ou sa litière sans avoir auparavant dressé une carte des alentours. Dans un univers instable, peuplé de prédateurs, faire régulièrement le tour du propriétaire, vérifier si l'environnement a changé, contrôler le pourquoi d'un bruit inattendu peuvent faire la différence entre la vie et la mort. La curiosité est cette détermination qui pousse les animaux à sortir de leur zone de confort pour acquérir du savoir. Dans un monde incertain, toute information possède de la valeur dans la monnaie même de Darwin : la survie.

La curiosité est donc une force qui nous incite à explorer. En cela, elle ressemble à la recherche de nourriture ou de partenaires sexuels, mais elle est motivée par une valeur immatérielle : l'acquisition d'informations nouvelles. La découverte d'une information inconnue porte en elle-même sa propre récompense : elle active le circuit de la dopamine. Souvenez-vous : c'est lui qui décharge en réponse à la nourriture, à la drogue ou au sexe. Chez les primates, et sans doute chez tous les mammifères, ce circuit ne répond plus seulement aux récompenses matérielles, mais également aux informations nouvelles. Certains neurones dopaminergiques signalent la quantité d'information à venir, comme si un surcroît d'informations nouvelles apportait sa propre gratification<sup>252</sup>. C'est ainsi qu'un rat peut être conditionné, non seulement à la drogue, mais également à la nouveauté : on peut l'amener à préférer un lieu simplement parce que des objets nouveaux y sont souvent présents, en sorte qu'il y trouvera matière à



satisfaire sa curiosité<sup>253</sup> . Nous n'agissons pas autrement lorsque nous déménageons pour une grande ville afin d'y trouver de nouvelles distractions ou lorsque, avides du dernier potin, nous nous précipitons sur Facebook ou Twitter.

Chez l'homme aussi, l'appétit de savoir passe par le circuit de la dopamine – même lorsqu'il s'agit d'une curiosité purement intellectuelle. Imaginez que vous soyez allongé dans une IRM, et qu'on vous pose une question du jeu Trivial Pursuit, par exemple : « Quand Marianne est-elle devenue le symbole de la République<sup>254</sup> ? » Avant de satisfaire votre curiosité, on vous demande à quel point vous êtes avide de connaître la réponse. Résultat ? Le degré de curiosité que vous manifestez correspond étroitement à l'activité du noyau accumbens et de l'aire tegmentale ventrale, deux régions essentielles du circuit cérébral de la dopamine. Plus vous êtes curieux, plus ces régions déchargent. Mieux, leurs signaux se déclenchent par anticipation : avant même de connaître la réponse, le simple fait de savoir que vous allez savoir excite vos circuits dopaminergiques et apporte sa propre récompense. Et toute cette curiosité est utile : elle prédit l'apprentissage. Vous retiendrez mieux les faits pour lesquels vous étiez le plus curieux. Vous vous souviendrez même d'autres éléments anecdotiques, par exemple le visage de la personne qui était présente ou qui vous a appris cette information que vous souhaitiez tellement connaître. L'avidité de savoir détermine la profondeur de la mémoire.

Ainsi, le simple fait de satisfaire notre appétit d'apprendre, ou même d'anticiper que nous allons pouvoir le faire, constitue une récompense en soi. Apprendre possède une valeur intrinsèque pour le système nerveux. Ce que nous appelons la curiosité n'est rien d'autre que l'exploitation de cette valeur.

À ce titre, notre espèce est spéciale, car elle possède, plus que toute autre, la capacité d'apprendre. Au fil de l'hominisation, notre capacité de représentation du monde a changé. Nous sommes probablement les seuls animaux à formuler des théories formelles du monde dans un langage de l'esprit. La science est devenue notre niche écologique : seule espèce à ne pas avoir d'habitat spécifique, nous parvenons à nous adapter, grâce à l'apprentissage, à pratiquement n'importe quel environnement. En miroir de cette extraordinaire faculté d'apprentissage, notre curiosité s'est décuplée. Au cours de notre

évolution, nous avons acquis une forme étendue de curiosité qu'on appelle *curiosité épistémique* : la pure envie de savoir dans tous les domaines, y compris les plus abstraits.

Comme d'autres mammifères, nous jouons et nous explorons – mais plus seulement par le mouvement : par la pensée. Là où d'autres animaux visitent l'espace qui les entoure, nous auscultons les concepts. Et notre espèce éprouve des émotions spécifiques, épistémiques, c'est-à-dire destinées à guider notre soif de savoir. Nous nous réjouissons, par exemple, devant la symétrie et la beauté pure des structures mathématiques – un magnifique théorème peut nous émouvoir bien plus qu'un morceau de chocolat. Le rire semble également être l'une de ces émotions sociales qui guident l'apprentissage. En effet, il se déclenche lorsque nous découvrons brutalement que l'une de nos hypothèses implicites se révèle fausse, ce qui nous oblige à réviser notre modèle du monde. L'hilarité serait, selon le philosophe Dan Dennett, un signal social que nous partageons avec les autres afin d'attirer l'attention sur une chose que nous avons mal comprise, et dont nous réalisons soudain qu'elle exige que nous révisions nos croyances<sup>255</sup>. Et, effectivement, toutes choses égales par ailleurs, le simple fait de rire augmente la curiosité et maximise l'apprentissage<sup>256</sup>.

## *Vouloir savoir : le moteur de la motivation*

Plusieurs psychologues ont essayé de définir, avec précision, à quel algorithme correspond la curiosité humaine. En effet, si l'on parvenait à le comprendre, non seulement on maîtriserait mieux cette variable essentielle à l'apprentissage, mais on pourrait également la reproduire dans une machine qui imiterait les performances de l'espèce humaine : un robot curieux.

Cette approche algorithmique commence à porter ses fruits. Les plus grands psychologues, depuis William James jusqu'à Jean Piaget ou Donald Hebb, ont

spéculé sur la nature des opérations mentales qui sous-tendent la curiosité. Selon eux, la curiosité est la manifestation directe de la motivation des enfants à comprendre le monde, à essayer d'en construire un modèle<sup>257</sup>. Elle se déclencherait chaque fois que notre cerveau détecte un décalage entre ce que nous connaissons déjà et ce que nous pourrions savoir – une zone d'apprentissage potentiel. À chaque instant, nous choisirions, parmi les actions qui nous sont accessibles, celles qui ont le plus de chances de nous permettre d'acquérir des informations utiles. La curiosité fonctionnerait donc comme un petit système cybernétique, semblable au fameux régulateur de Watt qui, sur les locomotives à vapeur, ouvre ou ferme une valve d'échappement afin de conserver une pression de vapeur constante. Dans le cas de la curiosité, il s'agirait d'un régulateur qui cherche à conserver constamment une certaine pression d'apprentissage. La curiosité s'oriente vers ce qu'il nous paraît utile d'apprendre, tandis qu'elle se détourne de ce que nous savons déjà, ou dont nous savons, d'expérience, que nous n'avons plus rien à apprendre.

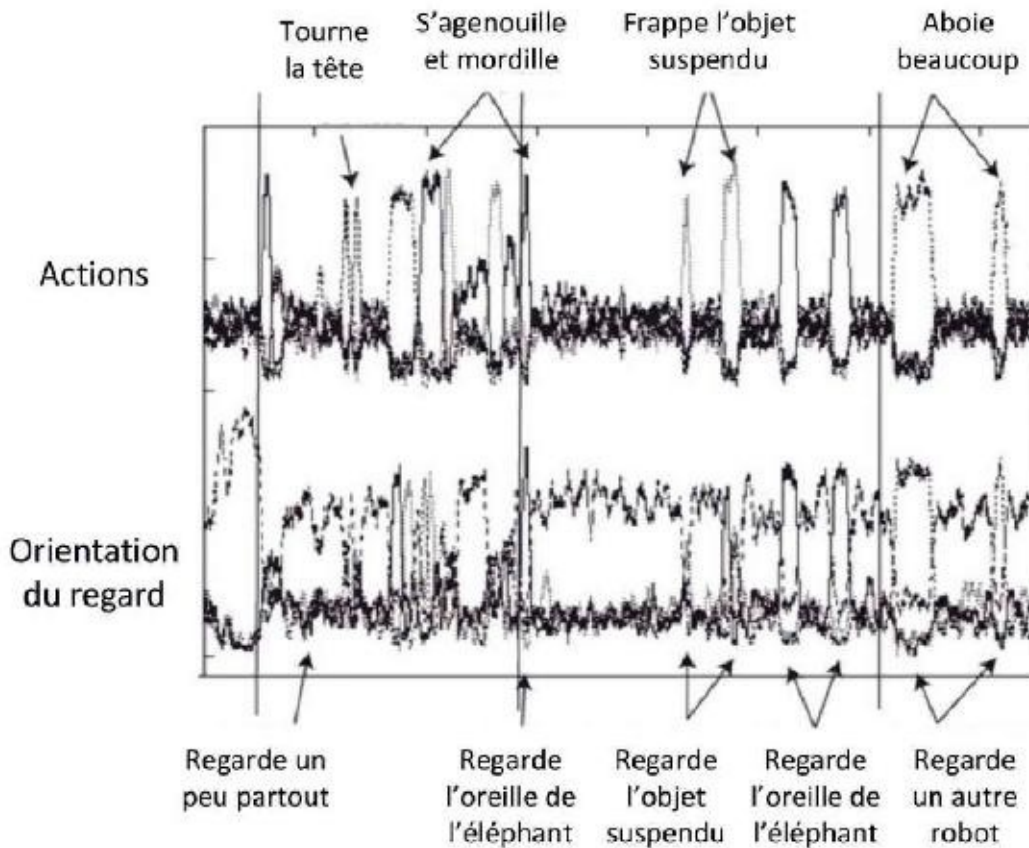
Ainsi, la curiosité n'est pas directement liée au degré de surprise ou de nouveauté, mais elle suit une courbe en cloche<sup>258</sup>. Nous n'éprouvons aucune curiosité pour les choses que nous avons déjà vues mille fois – leur familiarité est rébarbative. Nous n'éprouvons pas non plus d'attrance pour les choses trop nouvelles, tellement confuses que leur structure nous échappe – leur complexité même nous rebute. Entre la monotonie du trop simple et l'écueil du trop complexe, notre curiosité nous oriente naturellement vers les domaines nouveaux et accessibles. Mais cette attrance fluctue sans cesse. À mesure que nous les maîtrisons, les objets qui nous semblaient attirants perdent leur attrait, et nous réorientons notre curiosité vers d'autres défis nouveaux. C'est ainsi qu'un bébé se passionne initialement pour les choses les plus triviales (jouer avec ses doigts de pied, fermer les yeux, se cacher derrière sa main...), parce que tout est nouveau pour lui. Une fois qu'il a tiré tout le savoir de ces expériences, elles cessent de l'attirer, exactement pour la même raison que plus aucun scientifique ne reproduit les expériences de Galilée : le connu devient rébarbatif.

Mais le même algorithme explique également qu'on se détourne parfois d'un domaine qui nous attirait, mais qui s'avère trop difficile pour nous : notre

cerveau évalue la vitesse d'apprentissage, et la curiosité s'éteint s'il détecte que nous ne progressons pas suffisamment en regard de nos attentes. Ainsi un enfant peut-il rentrer d'un concert avec une passion pour le violon... pour s'en détourner quelques semaines plus tard lorsqu'il se rend compte qu'apprendre à en jouer ne vient pas en quelques jours.

Deux roboticiens français, Frédéric Kaplan et Pierre-Yves Oudeyer, sont parvenus à concrétiser ces idées dans un robot<sup>259</sup>. Leur algorithme comprend plusieurs modules. Le premier est un système classique d'intelligence artificielle, qui essaie en permanence de prédire l'état du monde extérieur. Le second, plus innovant, évalue la performance du premier : il mesure la vitesse récente d'apprentissage et l'utilise pour prédire dans quels domaines le robot apprend le plus d'informations nouvelles par seconde. Enfin, le circuit de la récompense accorde une valeur plus importante aux actions qui sont censées conduire à un meilleur apprentissage. Résultat : le système s'oriente naturellement vers les domaines où il estime qu'il apprendra le plus – c'est la définition même de la curiosité selon Kaplan et Oudeyer.

Lorsque le robot, muni de cet algorithme, est placé sur un tapis d'éveil, il se comporte exactement comme un jeune enfant (figure 32). Pendant quelques minutes, il se passionne pour un élément bien particulier, et passe par exemple tout son temps à soulever l'oreille d'un éléphant en tissu. Mais à mesure qu'il apprend à s'en servir, sa curiosité diminue. Au bout d'un moment, il s'en détourne et recherche activement d'autres sources de stimulation. Après une heure, il cesse d'explorer : tout ce qui pouvait être appris est désormais connu, et tout est devenu ennuyeux.



**Figure 32.** La curiosité est un ingrédient essentiel de l'algorithme d'apprentissage qu'emploie notre cerveau. Elle commence à être reproduite dans des machines. Ici, un petit robot explore un tapis d'éveil pour bébés. Il est guidé par une fonction de récompense qui lui fait choisir, à chaque instant, l'action qui maximise l'apprentissage. Dès que le robot maîtrise un aspect du monde, celui-ci perd de son intérêt et le robot réoriente son attention ailleurs.

L'analogie avec un petit enfant est frappante. Même les bébés de quelques mois s'orientent déjà vers les stimuli de complexité intermédiaire, ni trop simples, ni trop complexes, mais juste assez structurés pour qu'ils puissent les apprendre (ce qui a été décrit comme l'effet Boucle d'Or)<sup>260</sup>. Pour maximiser leurs apprentissages, il faut donc enrichir en permanence leur environnement d'objets nouveaux et juste assez stimulants pour ne pas être décourageants – une hiérarchie pédagogique bien conçue pour les entraîner progressivement vers les sommets, en piquant constamment leur appétit de savoir et de nouveauté.

Cette vision de la curiosité conduit à une prédiction intéressante. Elle dit qu'un enfant, pour être curieux, doit être conscient de ce qu'il ne sait pas, et peut-être même de sa vitesse d'apprentissage. Autrement dit, il doit posséder, dès le plus jeune âge, des facultés *métacognitives*. La métacognition, c'est la cognition sur la cognition, c'est-à-dire l'ensemble des systèmes qui supervisent les apprentissages et évaluent, en permanence, ce que nous savons et ce que nous ne savons pas, si nous nous trompons ou pas, si nous sommes rapides ou lents, et ainsi de suite – tout ce que nous connaissons sur notre propre esprit.

La métacognition joue un rôle déterminant dans la curiosité. En effet, être curieux, c'est vouloir savoir, et cela suppose de savoir ce qu'on ignore. Et, là encore, des expériences très récentes<sup>261</sup> confirment que, dès l'âge de 1 an et peut-être même plus tôt encore, un enfant comprend qu'il y a des choses qu'il ne sait pas. En effet, il est capable de se tourner vers sa maman lorsqu'il ne parvient pas à résoudre un problème seul. Le fait de savoir qu'il ne sait pas le pousse à solliciter des informations supplémentaires. C'est, déjà, la manifestation d'une forme de curiosité épistémique : l'envie irrésistible de savoir.

## *Les trois façons dont l'école peut tuer la curiosité*

Tous les parents ont la nostalgie de la petite enfance, où l'esprit de leur bambin pétillait de curiosité. Entre 2 et 5 ans, l'enfant est curieux de tout. Son mot favori est « pourquoi » : il ne cesse d'expérimenter sur le monde et d'interroger les adultes afin d'éteindre sa soif de savoir. Cet appétit qui paraissait insatiable finit pourtant par s'évanouir, souvent après quelques années d'école. Certains enfants restent curieux de tout, mais d'autres se renferment. Leur engagement actif se mue en morne passivité. La science de la curiosité peut-elle expliquer pourquoi ? Nous n'avons pas toutes les réponses, mais je voudrais proposer ici quelques éléments de spéculation informée.

Première hypothèse : le manque de stimulation appropriée au niveau de l'enfant. Selon l'algorithme que nous venons de décrire, il est normal que la curiosité diminue au fil des apprentissages : mieux nous maîtrisons un domaine, plus nous arrivons aux limites de ce qu'il nous offre à apprendre et moins il nous intéresse. Pour maintenir la curiosité, il faut donc que l'école continue de fournir, au superordinateur qu'est le cerveau, des stimulants à la hauteur de l'intelligence des enfants. Or ce n'est pas toujours le cas. Les élèves les plus avancés peuvent manquer de stimulation : après quelques mois, leur curiosité se fane et ils n'attendent plus grand-chose de l'école, parce que leur système métacognitif leur prédit que, malheureusement, ils n'y apprendraient guère. Une seule solution : piquer leur curiosité en leur donnant du nouveau grain à moudre – des langues à déchiffrer, des casse-tête à résoudre, de vrais défis à leur intelligence.

À l'autre bout de la courbe, les élèves en difficulté, eux, peuvent s'étioler pour la raison inverse. C'est toujours la métacognition qui est en cause : ils n'ont plus de raison d'être curieux, car ils ont appris... qu'ils ne parviendront pas à apprendre. Leur expérience passée a gravé, au plus profond de leurs circuits métacognitifs, une règle simple (quoique fautive) : tu es incapable d'apprendre tel ou tel domaine. Ce n'est pas rare : les petites filles se convainquent que les mathématiques ne sont pas pour elles<sup>262</sup>, certains enfants de banlieue que l'école leur est hostile et qu'ils n'y apprendront rien d'utile à leur avenir, et ainsi de suite. De tels jugements métacognitifs sont désastreux, car ils démotivent et tuent la curiosité dans l'œuf. Seule solution : remettre ces enfants en confiance

en leur démontrant, par des problèmes adaptés à leur niveau, qu'ils sont parfaitement capables d'apprendre, à leur rythme, et qu'y parvenir porte en soi sa propre récompense. La théorie de la curiosité dit que, lorsqu'un enfant est découragé, qu'il soit brillant ou en difficulté, il faut avant tout lui redonner le goût d'apprendre en lui proposant des stimulations adaptées à ses capacités.

Deuxième cas de figure : la punition de la curiosité. L'appétit de découvertes de l'enfant peut être tué dans l'œuf par une organisation scolaire trop rigide. L'enseignement traditionnel, par le biais du cours magistral, peut dissuader l'enfant d'intervenir ou même de réfléchir. Il peut lui donner l'impression qu'il lui est simplement demandé de se taire, de se tenir calme et d'attendre la fin du cours. L'interprétation neurophysiologique de cette situation est simple : le signal de récompense lié à la curiosité entre en compétition avec les autres récompenses et punitions extérieures au sein du circuit de la dopamine. Il est donc possible de décourager la curiosité en sanctionnant chaque tentative d'exploration par une punition. Imaginez un enfant qui, chaque fois qu'il essaie d'intervenir, est tancé, moqué ou puni : question idiote, tu ferais mieux de te taire, tu resteras une demi-heure de plus à travailler... Son cerveau apprend très vite à ne plus tenter la moindre intervention : la récompense que son cerveau tirait du fait de comprendre est largement compensée par les signaux négatifs qu'il reçoit. La punition répétée entraîne un syndrome d'impuissance acquise (*learned helplessness*), une sorte de paralysie physique et mentale associée au stress et à l'anxiété, dont la recherche animale démontre qu'elle inhibe les apprentissages<sup>263</sup>.

La solution ? Elle est bien connue de la plupart des enseignants. Il s'agit de récompenser la curiosité et non de la punir : encourager les questions, demander aux enfants de faire des exposés sur ce qui les passionne, féliciter chaque élève pour ses initiatives, même maladroites... Les neurosciences de la motivation sont extrêmement claires : pour avoir envie de faire une action, il faut anticiper qu'elle conduira à une récompense, laquelle peut être directe (nourriture, sexe, confort) ou cognitive (un gain d'information). Trop d'enfants ont perdu toute curiosité, toute motivation d'apprendre, parce qu'ils ont appris à n'espérer aucune récompense de l'école.



Troisième effet qui peut décourager la curiosité : la transmission sociale des connaissances. Souvenez-vous que l'espèce humaine dispose de deux modes d'apprentissage : le mode actif, où l'enfant s'interroge comme un scientifique en herbe, et le mode réceptif, pédagogique, où il se contente d'enregistrer ce que d'autres lui enseignent. L'école encourage nettement le second mode – mais il se peut qu'elle décourage progressivement le premier, à mesure que l'enfant découvre que l'enseignant sait toujours tout mieux que lui.

Une expérience récente montre que cet effet est réel : l'attitude de l'enseignant peut tuer la curiosité de l'enfant<sup>264</sup>. On présente à des élèves de maternelle un étrange assemblage de tubes, totalement nouveau pour eux, et qui s'avère être un véritable tiroir à malices : en cherchant un peu, on découvre qu'un miroir, un klaxon, un jeu de lumière et une boîte à musique sont cachés dans les différents tuyaux. Donnez ce gadget à un gamin, et vous déclenchez immédiatement un comportement d'exploration : il fouille, fouine et furète un peu partout. Maintenant, passez en mode pédagogique et donnez-lui l'objet en lui disant : « Regarde, je vais te montrer mon jouet, il fait ça »... et vous déclenchez par exemple la boîte à musique. On pourrait penser que cette phrase stimule la curiosité des enfants... mais l'effet est inverse : l'exploration diminue massivement. L'enfant semble faire l'hypothèse (souvent juste) que l'enseignant cherche à l'aider au maximum, et qu'il lui a donc présenté toutes les fonctions intéressantes de l'appareil. Dans ce contexte, plus besoin de chercher : la curiosité est inhibée.

L'expérience montre que les enfants tiennent compte de l'attitude de l'enseignant. Si celui-ci fait toujours des démonstrations exhaustives, alors les enfants cessent de chercher. Confrontés à un jouet nouveau, dont l'enseignant démontre l'une des fonctions, ils n'ont plus la curiosité d'en explorer toutes les facettes, car ils pensent que l'enseignant leur a tout expliqué. Si, au contraire, le maître leur laisse entendre qu'il ne sait pas tout, alors les enfants continuent de chercher.

Quelle est donc la bonne démarche ? Elle consiste à garder toujours en tête le concept d'engagement actif. Engager maximale l'intelligence de l'enfant, cela signifie le maintenir en alerte par des questions, des remarques qui stimulent

son imagination, et lui donnent envie d'aller plus loin. Pas question pour autant de laisser l'élève tout découvrir par lui-même – ce serait retomber dans l'écueil des pédagogies de la découverte. L'idéal est de proposer une pédagogie structurée, mais qui encourage la créativité de l'enfant en lui laissant entendre qu'il lui reste, s'il le souhaite, mille choses à découvrir. Je me souviens d'un enseignant qui, juste avant les grandes vacances, m'avait dit : « Tu sais, je viens de lire un petit problème de maths dont je n'ai pas trouvé la solution... » Et voilà comment je me suis retrouvé à cogiter tout l'été, pour faire mieux que le prof...

Susciter l'engagement actif des enfants va de pair avec un autre impératif : tolérer l'erreur. Ce sera notre troisième pilier de l'apprentissage.

# CHAPITRE 9

## Le retour sur erreur

---

« Il faudrait apprendre à se tromper de bonne humeur. [...] Penser, c'est aller d'erreur en erreur. »

ALAIN (Émile Chartier),  
*Propos sur l'éducation* (1932).

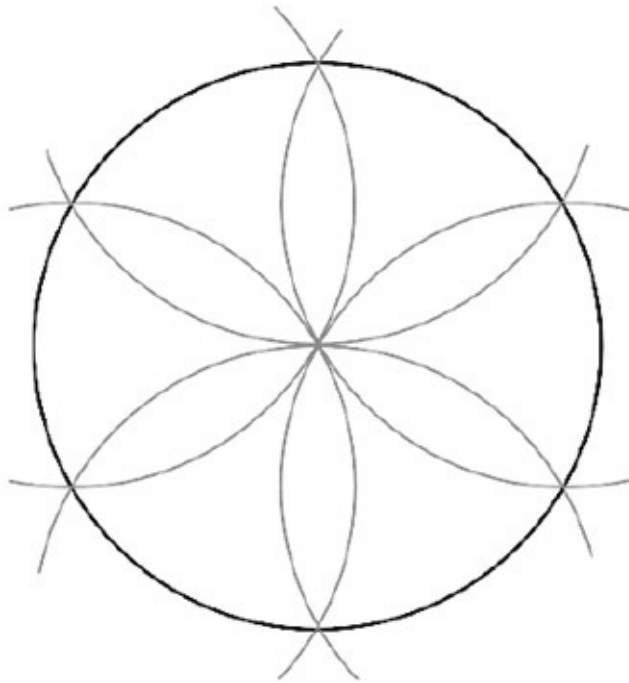
En 1940, le jeune Alexandre Grothendieck n'avait que 11 ou 12 ans. Il ne savait pas qu'il allait devenir l'un des plus influents mathématiciens du  $XX^e$  siècle, qui inspirerait toute une génération (c'est pour lui que fut fondé, en 1958, le célèbre Institut des hautes études scientifiques de Bures-sur-Yvette, qui a fécondé tant de médailles Fields). Mais le jeune Alexandre faisait déjà des mathématiques... avec plus ou moins de bonheur. Écoutons ses Mémoires<sup>265</sup> :

« Vers l'âge de 11 ou 12 ans, alors que j'étais interné au camp de concentration de Rieucros (près de Mende), j'ai découvert les jeux de tracés au compas, enchanté notamment par les rosaces à six branches qu'on obtient en partageant la circonférence en six parties égales à l'aide de l'ouverture du compas reportée sur la circonférence à six reprises, ce

qui fait retomber pile sur le point de départ. Cette constatation expérimentale m'avait convaincu que la longueur de la circonférence était exactement égale à six fois celle du rayon. Quand par la suite [...], j'ai vu dans un livre de classe que la relation était censée être bien plus compliquée, que l'on avait  $L = 2 \pi R$  avec  $\pi = 3,14\dots$ , j'étais persuadé que le livre se trompait, que les auteurs du livre [...] n'avaient jamais dû faire ce tracé très simple, qui montrait à l'évidence que l'on avait tout simplement  $\pi = 3$ .

Cette confiance qu'un enfant peut avoir en ses propres lumières, en se fiant à ses facultés plutôt que de prendre pour argent comptant les choses apprises à l'école ou lues dans les livres, est une chose précieuse. Elle est constamment découragée pourtant par l'entourage.

Beaucoup verront dans l'expérience que je rapporte ici l'exemple d'une présomption enfantine, qui a dû s'incliner devant le savoir reçu – les faits faisant enfin éclater un certain ridicule. Tel que j'ai vécu cet épisode, il n'y avait pourtant nullement le sentiment d'une déconvenue, d'un ridicule, mais bien celle d'une nouvelle découverte [...] : celle d'une erreur. »



Extraordinaire leçon d'humilité que celle d'un grand mathématicien qui avoue avoir longtemps cru que le nombre pi valait trois... Mais sur l'importance de l'erreur, Grothendieck voit juste. Se tromper, c'est déjà apprendre. Les deux termes sont virtuellement synonymes, car chaque erreur est une opportunité d'apprentissage.

Les Shadoks, avec humour, l'érigeaient au rang de principe : « Ce n'est qu'en essayant continuellement que l'on finit par réussir... En d'autres termes, plus ça rate et plus on a de chances que ça marche ! » Sans aller aussi loin, il est vrai qu'il est pratiquement impossible de progresser si l'on ne commence pas par échouer – à condition de recevoir un signal de feed-back, une rétroaction qui nous indique la bonne voie. C'est pourquoi le retour sur erreur est le troisième pilier de l'apprentissage, et l'un des paramètres éducatifs les plus influents : la qualité et la précision du retour que nous recevons déterminent la rapidité avec laquelle nous apprenons<sup>266</sup> .

## *La surprise, moteur de l'apprentissage*

Souvenez-vous des algorithmes d'apprentissage que nous avons discutés dans l'introduction, et qui permettent à un chasseur d'ajuster son tir, ou à un réseau de neurones d'apprendre en se corrigeant. Le principe est simple : il faut essayer, quitte à échouer, car la taille et la direction de l'erreur indiquent comment se corriger. On vise, on tire, on regarde de combien on a manqué la cible, et on utilise cette information pour corriger le tir suivant. C'est ainsi qu'un chasseur apprend à viser – et, à une échelle plus vaste, c'est ainsi qu'un réseau de neurones artificiel parvient à ajuster les millions de paramètres qui définissent son modèle interne du monde extérieur. Mais le cerveau fonctionne-t-il ainsi ?

Dès les années 1970, les données s'accumulent en faveur de cette hypothèse. Deux chercheurs américains, Robert Rescorla et Allan Wagner, font cette proposition : le cerveau n'apprend que s'il perçoit un décalage entre ce qu'il prédit et ce qu'il reçoit. Aucun apprentissage n'est possible en l'absence d'un signal d'erreur : « Les organismes n'apprennent que lorsque les événements violent leurs attentes<sup>267</sup> . » Autrement dit, la surprise est l'un des moteurs fondamentaux de l'apprentissage.

La théorie parvient à expliquer de nombreux aspects de l'apprentissage par conditionnement. Vous connaissez l'expérience du chien de Pavlov : une cloche, initialement neutre et sans effet, finit par déclencher un réflexe conditionné de salivation lorsque le chien comprend que la cloche précède systématiquement la nourriture. Comment les deux chercheurs expliquent-ils cet apprentissage ? La règle de Rescorla-Wagner suppose que le cerveau utilise les entrées sensorielles (les sensations engendrées par la cloche) pour prédire la probabilité qu'un autre stimulus survienne (la nourriture). Elle stipule que :

- Le cerveau se forge une prédiction en pondérant les entrées sensorielles.
- Il calcule ensuite la différence entre la prédiction et le stimulus réellement obtenu : c'est l'*erreur de prédiction*, un concept fondamental de la théorie, qui détermine le degré de surprise associé au stimulus.

- Le cerveau corrige ensuite sa représentation interne, en proportion directe de la force du stimulus et de la valeur de l'erreur de prédiction, afin que sa prochaine prédiction soit plus proche de la réalité.

Ainsi, cette théorie inclut, de façon explicite, les trois piliers de l'apprentissage que nous avons postulés : l'apprentissage ne se produit que si le cerveau amplifie les entrées sensorielles appropriées (attention) ; s'il les utilise pour produire une prédiction (engagement actif) ; et s'il parvient à en déterminer la justesse (retour sur erreur).

L'équation proposée par Rescorla et Wagner est pratiquement identique à celle qui sera massivement utilisée par la suite dans les réseaux de neurones artificiels sous le nom de « règle delta ». C'est un cas particulier de la règle de rétropropagation d'erreurs qu'emploient pratiquement tous les systèmes actuels d'apprentissage supervisé (où l'on enseigne explicitement au réseau une réponse bien précise). De plus, dans le cas de l'apprentissage par récompense (où l'on dit juste au réseau à quel point il se trompe), c'est encore une équation similaire qui est utilisée : le réseau prédit la récompense, et c'est la différence entre la prédiction et la récompense effective qui est utilisée pour mettre à jour la représentation interne. On peut donc dire qu'aujourd'hui, les cerveaux *in silico* effectuent des calculs inspirés du cerveau *in vivo*.

Comme nous l'avons vu plus haut, le cerveau humain va encore plus loin : pour extraire un maximum d'information de chaque épisode d'apprentissage, il utilise vraisemblablement un langage de la pensée et des modèles statistiques bien plus raffinés que les réseaux de neurones actuels. Cependant, l'idée fondamentale de Rescorla et Wagner reste juste : le cerveau essaie d'anticiper sur les entrées qu'il reçoit, et ajuste ces prédictions selon la surprise, l'improbabilité, l'erreur associées aux informations qu'il reçoit. Apprendre, c'est réduire l'imprévisible.

La théorie de Rescorla et Wagner a eu une influence considérable, car elle constituait un progrès majeur par rapport aux théories précédentes, fondées sur le concept d'apprentissage par association. Auparavant, l'idée reçue était que le cerveau apprenait à « associer » la cloche et la nourriture : il se contentait

d'enregistrer, passivement, la coïncidence entre le stimulus et la réponse. Or, même pour le conditionnement pavlovien, cette vision se révèle complètement fausse<sup>268</sup>. Le cerveau n'est pas un organe passif qui absorbe des associations. L'apprentissage est actif et dépend du degré de surprise lié à la violation de nos attentes.

L'une des expériences les plus remarquables, qui battent en brèche la notion d'apprentissage purement associatif, est celle du blocage<sup>269</sup>. On présente à l'animal non pas un, mais deux indices sensoriels – disons une cloche et une lumière – qui tous deux prédisent la nourriture. L'astuce consiste à les présenter l'un après l'autre. On commence avec la lumière, et l'animal apprend qu'elle prédit la nourriture. Ensuite, on présente des essais doubles, où la cloche et la lumière prédisent la nourriture. Enfin, on teste l'effet de la cloche seule. Surprise : elle n'a aucun effet ! L'animal ne salive pas, il semble n'avoir rien retenu de l'association, maintes fois répétée, entre la cloche et la récompense. Que s'est-il passé ? C'est le premier apprentissage (lumière → nourriture) qui bloque le second (cloche → nourriture). Pourquoi ? Parce que la prédiction fondée sur la lumière explique déjà tout. L'animal sait déjà que la lumière prédit la nourriture, son cerveau ne génère donc aucune erreur de prédiction lors des essais cloche + lumière. L'expérience du blocage montre bien que l'apprentissage ne fonctionne pas par association : l'association cloche → nourriture a beau être répétée des centaines de fois, elle ne conduit à aucun apprentissage lorsqu'elle est déjà expliquée par des connaissances antérieures.

Ainsi, en l'absence de surprise, aucun apprentissage ne se produit : l'erreur de prédiction est indispensable à l'apprentissage. Mais attention : nous parlons d'un signal d'erreur interne, pas nécessairement d'une erreur effective. Supposez que je doive découvrir la bonne réponse parmi deux choix possibles. Est-ce que cela signifie que, si je tombe juste du premier coup, je n'apprends rien ? Pas du tout, parce que même si je répons correctement, je le fais au hasard : ma confiance est faible et ma prédiction incertaine. Dans ce cas, le feed-back m'apporte une information nouvelle : la certitude que ma réponse, hasardée au petit bonheur, était juste. Selon l'équation de Rescorla et Wagner, cette information nouvelle constitue un signal d'erreur : elle mesure le décalage entre



ce que je prédisais (50 % de chances d'avoir raison) et ce que je sais maintenant (100 % de certitude de connaître la bonne réponse). Dans mon cerveau, cette information se propage et met à jour mes connaissances. Ce qui compte, pour apprendre, ce n'est donc pas de se tromper à tous les coups, comme les Shadoks ! L'important, c'est la surprise, c'est-à-dire le décalage entre la prédiction et la réalité. C'est cela qu'on appelle un signal d'erreur.

Pas de surprise, peu ou pas d'apprentissage : cette règle semble aujourd'hui validée chez toutes sortes d'organismes – y compris chez le très jeune enfant. Souvenez-vous que la surprise est l'un des révélateurs des compétences précoces des bébés : ceux-ci écarquillent les yeux et regardent plus longtemps lorsque, par magie, on leur présente des événements surprenants qui violent les lois de la physique, de l'arithmétique, des probabilités ou de la psychologie (voir, le [chapitre 3](#) et la [figure 8](#) ). Mais l'enfant ne se contente pas de sourciller : chaque fois qu'il est surpris, il apprend.

La psychologue américaine Lisa Feigenson tire cette conclusion de nombreuses expériences qui montrent que, chaque fois que l'enfant perçoit un événement impossible ou improbable, l'apprentissage est facilité<sup>270</sup> . Ainsi, lorsqu'ils voient un objet traverser mystérieusement un mur, les bébés retiennent mieux le bruit qu'il fait ou le verbe qu'un adulte vient d'employer (« Regarde, je viens de bliquer le jouet »). Lorsqu'on leur donne cet objet en main, ils l'explorent plus longtemps. Leur comportement montre qu'ils essaient de comprendre : en bons scientifiques, ils pratiquent des expériences pour tenter de répliquer ce qu'ils ont cru voir. Ainsi, si l'objet vient de traverser un mur, ils le frappent comme pour en éprouver la solidité ; alors que s'ils l'ont vu violer les lois de la gravité et rester mystérieusement suspendu dans l'air, ils le font tomber de la table comme pour vérifier s'il est vraiment capable de léviter. Autrement dit, la nature de l'erreur que l'enfant vient d'observer détermine la manière dont il ajuste ses hypothèses dans les instants qui suivent. C'est exactement ce que prédit la théorie de la propagation des erreurs : chaque événement imprévu conduit à un ajustement des modèles du monde.

Tous ces phénomènes sont attestés chez le bébé de 11 mois, mais ils sont probablement présents bien plus tôt. L'apprentissage par correction d'erreurs est

universellement répandu dans le monde animal, et il y a tout lieu de penser que la surprise gouverne les apprentissages depuis le tout début de la vie.

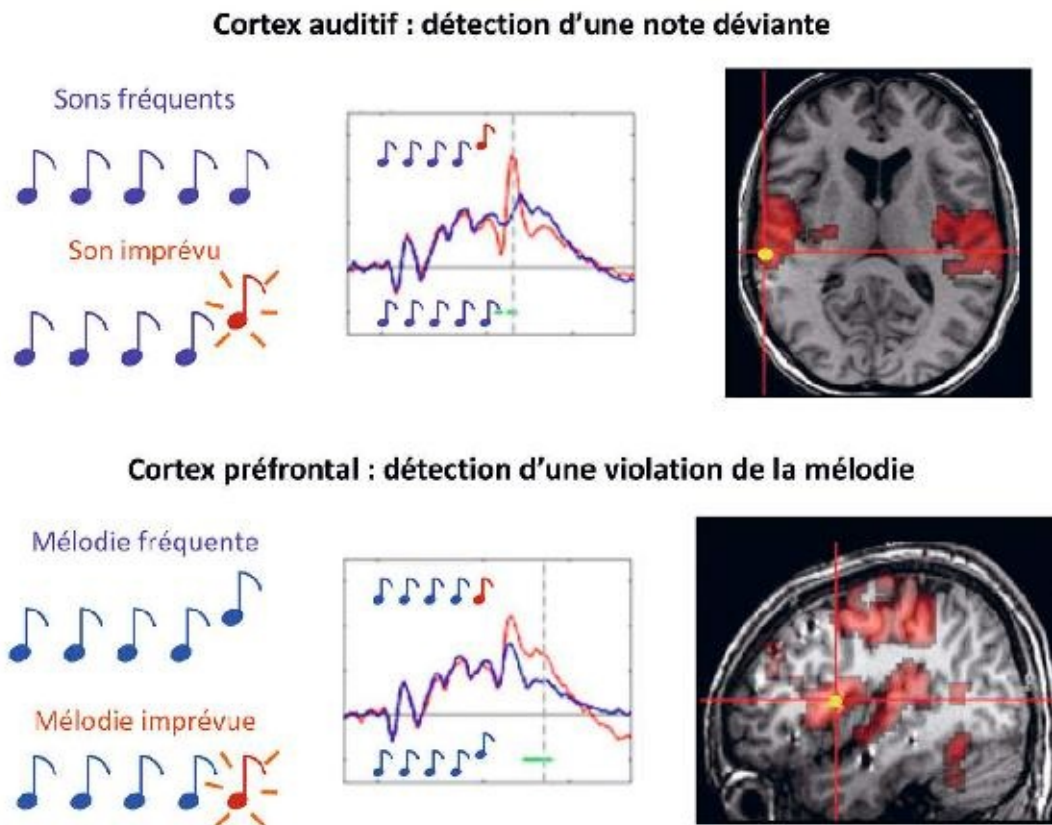
## *Le cerveau fourmille de messages d'erreur*

L'erreur joue un rôle tellement fondamental dans l'apprentissage que pratiquement toutes les aires cérébrales émettent et échangent des messages d'erreur (figure 33)<sup>271</sup>. Prenons un exemple simple : vous entendez une série de notes, toujours les mêmes, *do do do...* À mesure que la même note se répète, les réponses cérébrales diminuent – c'est l'*adaptation*, un phénomène qui montre que votre cerveau apprend à prédire la suite de la série. Soudain, sans prévenir, changez la note : *do do do ré*. Votre aire auditive primaire répond immédiatement par une réaction de surprise : l'adaptation s'évanouit, et les neurones se mettent à décharger vigoureusement au son inattendu.

Notez que ce n'est pas simplement la répétition qui entraîne une adaptation, mais bien le fait que les notes soient prévisibles. En effet, si l'on présente une alternance de notes comme *do ré do ré do ré...*, le cerveau s'habitue à ce balancement systématique, sa réponse diminue, et c'est à présent une répétition inattendue (*do ré do ré do do*) qui déclenche la surprise<sup>272</sup>.

L'aire auditive semble faire un calcul simple : elle utilise le passé récent pour prédire l'avenir. Dès qu'une note ou un groupe de notes se répètent, cette région en conclut qu'elles vont se poursuivre. Cette prédiction est utile, car elle évite de prêter attention aux signaux réguliers. Ceux-ci sont littéralement annulés : comme le signal sensoriel qui entre est pratiquement identique à la prédiction qu'en fait le cerveau, la différence est nulle. La soustraction de la prédiction efface l'information entrante – mais uniquement pour les entrées prévisibles. Tout ce qui viole les attentes, au contraire, est amplifié. Ainsi, l'aire auditive agit comme un filtre : elle ne transmet aux niveaux supérieurs du cortex que les

informations surprenantes, imprévisibles, qu'elle ne parvient pas à expliquer seule.



**Figure 33.** Le retour sur erreur est le troisième pilier de l'apprentissage. En détectant ses erreurs, le cerveau parvient à corriger ses modèles du monde. Pratiquement toutes les régions cérébrales émettent et échangent des signaux d'erreur. Ici, l'erreur est auditive. On fait d'abord écouter plusieurs fois une petite mélodie de 5 notes. Lorsque la séquence change sans prévenir, une réponse de surprise (en rouge) signale l'erreur aux autres régions du cerveau et leur permet de s'amender. Les aires auditives détectent les fausses notes locales (en haut), tandis qu'un réseau étendu, qui va jusqu'au cortex préfrontal, détecte les violations globales de la mélodie (en bas).

Ce qu'une région cérébrale ne parvient pas à expliquer, elle le passe donc au niveau suivant, et celui-ci, à son tour, va tenter de le comprendre. Il existe ainsi toute une hiérarchie de messages d'erreur. Par exemple, comme je viens de l'expliquer, la mélodie *do do do ré* engendre, à bas niveau, un signal d'erreur

dans le cortex auditif au moment du *ré* , qui diffère des sons précédents. Mais, à plus haut niveau, c'est une mélodie connue, le début d'*Au clair de la lune* . La surprise causée par le *ré* n'est donc que périphérique : elle est immédiatement expliquée par une représentation de niveau supérieur, et le signal de surprise s'arrête là : le *ré* , bien que nouveau, n'engendre aucune surprise dans le cortex préfrontal inférieur, qui encode les phrases musicales. Par contre, la répétition *do do do do* a l'effet inverse : monotone, elle n'engendre aucun signal d'erreur dans les aires auditives, mais elle crée la surprise dans les aires de plus haut niveau qui codent pour la mélodie globale : elles prédisaient une montée vers le *ré* , et c'est un nouveau *do* qui survient. Ici, la surprise, c'est qu'il n'y a pas de surprise ! Même les singes macaques présentent, comme nous, ces deux niveaux de traitement auditif : le traitement local des notes dans le cortex auditif, et la représentation globale de la mélodie dans le cortex préfrontal<sup>273</sup> .

De tels signaux d'erreur s'observent dans toutes les régions du cerveau. Partout, les décharges neuronales s'habituent lorsque les événements se répètent et deviennent prévisibles ; et partout, la surprise, c'est-à-dire la survenue soudaine d'un événement inattendu, augmente brutalement la réponse des neurones. La seule chose qui change, d'une région à l'autre, c'est la nature de la prédiction et de la nouveauté détectée. Dans le cortex visuel, c'est la présentation d'une image inattendue qui relance l'activité<sup>274</sup> ; les aires du langage, elles, réagissent à un mot anormal au sein d'une phrase. Lisez, par exemple, la phrase suivante :

« Je préfère manger avec une fourchette et un chameau. »

Votre cerveau vient de générer une onde N400, un signal d'erreur évoqué par un mot ou une image incompatible avec le contexte<sup>275</sup> . Comme son nom l'indique, cette réponse négative survient vers 400 millisecondes après l'anomalie, dans le cortex temporal gauche. Quant au cortex préfrontal inférieur, dans la région dite de Broca, il répond aux erreurs de syntaxe, quand le cerveau

prédit une certaine catégorie de mots et en reçoit une autre<sup>276</sup>, comme dans la phrase suivante :

« N’oubliez pas de prendre votre chaque médicament deux fois par jour. »

Cette fois-ci, vos aires du langage, spécialisées dans la syntaxe, ont émis une onde négative immédiatement suivie d’une onde P600 – un pic positif qui survient vers 600 millisecondes. Celui-ci signale que votre cerveau a détecté l’erreur de grammaire et tente de la réparer.

Le circuit cérébral dans lequel prédictions et erreurs de prédiction ont été le mieux démontrées est celui de la récompense<sup>277</sup>. Le réseau de la dopamine ne se contente pas de répondre aux récompenses : il les anticipe en permanence. Les neurones dopaminergiques situés dans un petit noyau de cellules appelé l’aire tegmentale ventrale ne répondent pas seulement aux plaisirs du sexe, de la nourriture ou de la boisson ; ils déchargent, en réalité, à la différence entre la récompense attendue et celle obtenue, c’est-à-dire à l’erreur de prédiction. Ainsi, si on donne à l’animal une récompense à l’improviste, disons une goutte d’eau sucrée, cette bonne surprise entraîne une décharge neuronale. Mais si cette récompense est précédée d’un signal qui la rend prévisible, alors le même sirop sucré n’entraîne plus la moindre réaction – et par contre, c’est le signal lui-même qui provoque un shoot de dopamine : l’apprentissage déplace la réponse au plus près du signal qui prédit la récompense.

Parce que le circuit de la dopamine se met à répondre à l’avance, des signaux arbitraires peuvent devenir porteurs de leur propre récompense. Cet effet a été démontré chez l’homme pour les pièces et les billets de banque – ou pour la simple vue d’une seringue chez un toxicomane. Ce signal prédictif est extrêmement utile à l’apprentissage, car il permet au système d’apprendre à se critiquer lui-même et d’anticiper l’erreur ou la réussite d’une action, sans devoir attendre qu’elle soit explicitement sanctionnée de l’extérieur. C’est pourquoi l’architecture acteur-critique, dans lequel un réseau de neurones apprend à

critiquer et à prédire le résultat des actions d'un autre réseau de neurones, est aujourd'hui universellement employée en intelligence artificielle pour résoudre les problèmes les plus complexes. Prédire, détecter son erreur, se corriger sont les fondements mêmes d'un apprentissage efficace.

## *Ne confondons pas erreur et sanction*

« J'ai souvent été frappé du fait que les professeurs de sciences, plus encore que les autres, ne comprennent pas qu'on ne comprenne pas. Peu nombreux sont ceux qui ont creusé la psychologie de l'erreur, de l'ignorance et de l'irréflexion. »

Gaston BACHELARD ,  
*La Formation de l'esprit scientifique* (1938).

Comment tirer le meilleur parti des signaux d'erreur qu'échangent nos aires cérébrales ? Pour qu'un enfant ou un adulte apprenne efficacement, il faut que son environnement (qu'il s'agisse des parents, de l'école, de l'université... ou d'un jeu vidéo) lui fournisse, le plus rapidement et le plus précisément possible, un retour sur erreur. Un simple signal binaire (« juste » ou « faux ») peut déjà l'aider. C'est le principe de l'apprentissage *non supervisé* . Toutefois, pour accélérer l'apprentissage, l'idéal est un retour sur erreur détaillé qui indique précisément ce qu'il aurait fallu faire pour ne pas se tromper. C'est le principe de l'apprentissage *supervisé*. En étant le plus précis possible, en fournissant un retour sur erreur rapide et détaillé, l'enseignant enrichit considérablement l'information dont dispose son élève pour se corriger. En intelligence artificielle, cet apprentissage est le plus efficace, car lui seul permet à la machine de remonter rapidement à la source de l'échec et de s'amender dans le bon sens. Il est important de comprendre que ce retour sur erreur n'a rien à voir avec une punition. On ne « punit » pas un réseau de neurones, on l'informe simplement des endroits où il s'est trompé, en étant le plus précis possible sur la nature et le signe de ses erreurs. L'informatique rejoint ici les sciences de l'éducation. En

effet, les méta-analyses réalisées par le spécialiste australien de l'éducation John Hattie le montrent clairement : la qualité du retour sur erreur que reçoivent les élèves est l'un des facteurs déterminants de la réussite scolaire<sup>278</sup> . Fixer clairement le but de l'apprentissage et permettre aux élèves de s'en approcher progressivement, sans dramatiser leurs inévitables erreurs, sont les clés de la réussite.

Le bon enseignant le sait bien, lui qui tente de diagnostiquer, le plus finement possible, l'endroit exact où ses élèves sont à la peine et de les aider à trouver le meilleur remède. Bienveillant, il sait bien que l'erreur est humaine, que personne n'apprend sans jamais se tromper. Avec l'expérience, il se forge un catalogue d'erreurs car tous les élèves, retraçant les mêmes pas, retombent dans les mêmes ornières. Et il trouve, avec indulgence, les mots pour rassurer, restaurer la confiance en soi et amender les représentations mentales erronées. Il est là pour dire le vrai, pas pour juger.

Bien sûr, me diront les plus rationnels d'entre vous, dire à l'élève qu'il s'est trompé, n'est-ce pas la même chose que de lui dire ce qu'il aurait dû faire ? D'un point de vue logique, si une question n'a que deux réponses possibles, A ou B, et que l'élève a choisi A, lui dire « la bonne réponse était B » revient au même que de lui dire « tu t'es trompé ». Mais c'est oublier que les enfants ne sont pas de parfaits logiciens dépourvus de toute affectivité. Pour eux, la déduction supplémentaire « si je me suis trompé, c'est que la bonne réponse était B » ne va pas de soi. Par contre, ils reçoivent cinq sur cinq le principal message, qui a toujours le parfum d'une sanction : je me suis planté. L'expérience a été faite : là où les adultes, avec un détachement cartésien, parviennent à extraire autant d'information des signaux positifs et négatifs, les adolescents échouent : ils apprennent mieux de leurs réussites que de leurs erreurs<sup>279</sup> . Épargnons-leur donc cette peine, en leur donnant un feed-back le plus neutre et le plus informatif possible. Corriger n'est pas punir.

## *La note, piètre retour sur erreur*

Il me faut à présent dire un mot d'une institution scolaire bourrée de défauts et pourtant tellement ancrée dans notre tradition que nous imaginons mal l'école sans elle : la note. Selon la théorie de l'apprentissage, la note est un signal de récompense (ou de punition !). Cependant, l'un de ses travers est d'être totalement dépourvue de précision. Simple cumul de nos fautes, elle résume, sans les distinguer, différentes sources d'erreurs. Elle n'est donc pas suffisamment informative, car elle ne permet pas, à elle seule, de savoir *pourquoi* on s'est trompé ni *comment* se corriger. Parfois même, elle ne décolle jamais du zéro absolu, auquel cas elle ne contient plus la moindre information !

La note sèche, lorsqu'elle n'est pas accompagnée d'appréciations détaillées et constructives, constitue donc un bien piètre retour sur erreur. Non seulement elle n'est pas précise, mais elle est souvent différée de plusieurs semaines, ce qui fait que l'élève a tout oublié de ce qui l'a induit en erreur. La note peut également être profondément injuste, lorsqu'elle sanctionne un devoir sur table qui revient chaque semaine et dont le niveau ne cesse d'augmenter de semaine en semaine.

Prenons l'analogie du jeu vidéo. Lorsque vous découvrez un nouveau jeu, forcément, vous ne savez pas quelle est la manière efficace de progresser. Vous n'avez pas envie qu'on vous rappelle en permanence à quel point vous êtes mauvais ! C'est pourquoi les concepteurs de jeu vidéo introduisent, au départ, des niveaux extrêmement faciles, où vous vous faites plaisir en gagnant presque à coup sûr. Très progressivement, la difficulté augmente et, avec elle, le risque d'échec et de frustration – mais les programmeurs savent la mitiger en mélangeant le facile et le difficile, et en vous donnant toute latitude de rejouer le même niveau autant que nécessaire. Progressivement, votre score s'améliore... et un jour d'allégresse, vous franchissez ce maudit obstacle qui vous bloquait. Maintenant, comparez cela avec le bulletin de notes d'un mauvais élève : il démarre l'année avec une mauvaise note, et au lieu de le remotiver en lui laissant repasser le même test jusqu'à la réussite, on lui impose, chaque semaine, un



exercice nouveau, toujours au-delà de ses capacités. Jour après jour, son « score » reste voisin de zéro. Sur le marché du jeu vidéo, un design aussi désastreux serait un échec cuisant.

Trop souvent, à l'école, la note sert de sanction. Nous ne pouvons ignorer ses effets massifs sur les systèmes émotionnels du cerveau : découragement, sentiment d'impuissance et de stigmatisation chez les enfants. Écoutons la voix d'un expert – Daniel Pennac, aujourd'hui écrivain renommé, qui reçut en 2007 le prix Renaudot pour son livre *Chagrin d'école*, mais qui fut longtemps un cancre : « Mes bulletins scolaires me le confirmaient tous les mois : si j'étais un crétin, c'était absolument de ma faute. D'où une détestation de moi, un complexe d'infériorité et surtout de culpabilité... Je me considérais comme un moins que rien. Car qui n'est bon à rien – ce que me répétaient les profs les uns après les autres – n'est rien... Je ne me voyais d'ailleurs aucun avenir, je n'avais aucune représentation possible de moi adulte. Pas parce que je ne désirais rien, mais parce que je me croyais inapte à tout<sup>280</sup>. »

Même si la résilience de l'élève Pennac est remarquable, un tel état d'esprit est évidemment désastreux pour la plupart des enfants. Les effets du stress ont été particulièrement étudiés dans le domaine des mathématiques, championnes mondiales de l'anxiété scolaire. En cours de maths, certains enfants souffrent d'un authentique mal-être proche de la dépression, tant ils savent que, quoi qu'ils fassent, ils seront sanctionnés d'un échec. L'anxiété mathématique est un syndrome reconnu, mesuré, quantifié : les enfants qui en souffrent montrent des activations des circuits de la douleur et de la peur, notamment l'amygdale, un noyau de neurones situé en deçà du cortex et impliqué dans les émotions négatives<sup>281</sup>. Ils ne sont pas forcément moins intelligents que les autres, mais le tsunami émotionnel qu'ils éprouvent abolit leurs capacités de calcul, de mémoire à court terme et surtout d'apprentissage.

De nombreuses recherches, tant chez l'homme que chez l'animal, le confirment : le stress et l'anxiété nuisent à la faculté d'apprendre. Dans l'hippocampe de la souris, le conditionnement par la peur bloque littéralement la plasticité neuronale – le circuit se retrouve dans un état qui ressemble à la fin de la période critique, où les synapses sont devenues immobiles, figées dans leurs

filets. À l'inverse, le fait d'être plongé dans un environnement enrichi, stimulant, libère la plasticité et rend les neurones plus mobiles, comme un retour en enfance<sup>282</sup>.

Distribuer une mauvaise note en la présentant comme une sanction, c'est donc prendre le risque d'inhiber les progrès de l'enfant, car le stress et le découragement l'empêcheront d'apprendre. À plus long terme, c'est aussi modifier sa personnalité et son image de soi. La psychologue américaine Carol Dweck a longuement étudié les effets négatifs de cette disposition mentale qui consiste à attribuer ses échecs (ou ses réussites) à un aspect fixe, immuable de sa propre personnalité – ce qu'elle appelle une « mentalité fixiste » (*fixed mindset*) : « je suis nul en maths », « les langues étrangères ne sont pas mon fort » et ainsi de suite. Elle l'oppose à l'idée, fondamentalement juste, que tous les enfants peuvent progresser – ce qu'elle appelle une « mentalité progressiste » (*growth mindset*).

Ses recherches suggèrent qu'à performance égale, l'état d'esprit joue un rôle important dans l'apprentissage<sup>283</sup>. Avoir l'idée de progrès chevillée au corps est en soi source de progrès. À l'inverse, les enfants et les enseignants qui adhèrent à l'idée que les compétences sont fixées une fois pour toutes, qu'on est « doué » ou « pas doué », obtiennent de moins bons résultats. En effet, le fixisme démotive : il n'encourage ni l'attention ni l'engagement actif, et il interprète l'erreur comme l'indice d'une nullité intrinsèque... alors que se tromper est naturel, cela prouve simplement qu'on a essayé ! Que serait devenu Grothendieck s'il avait conclu, à 11 ans, qu'il était nul en maths ? La recherche montre d'ailleurs que l'état d'esprit fixiste ne handicape pas seulement les mauvais élèves, mais aussi les bons : eux aussi ont besoin de travailler, de conserver leur motivation, et ce n'est pas un cadeau que de leur laisser entendre que tout leur est acquis d'avance parce qu'ils sont « doués ».

Avoir un état d'esprit progressiste, ce n'est donc pas bérer d'admiration devant l'enfant, sous le simple prétexte de nourrir son estime de soi – mais plutôt remarquer ses progrès, encourager sa participation, stimuler son attention... et, pourquoi pas, lui expliquer les fondements mêmes de l'apprentissage : que tous

les enfants doivent déployer des efforts et s'engager au risque de l'erreur, que c'est la seule façon de développer les circuits de son cerveau.

Laissons le dernier mot à Daniel Pennac :

« Vos professeurs ne sont pas là pour faire peur, mais pour vaincre la peur d'apprendre. Une fois cette peur vaincue, les élèves sont insatiables. »

## *Se tester pour mieux apprendre*

Si la note n'est guère efficace, comment mettre en pratique, dans nos salles de classe, les connaissances scientifiques sur le traitement de l'erreur dans le cerveau ? Les règles du jeu sont simples. D'abord, on doit forcer l'apprenant à s'engager, hasarder une réponse, générer activement une hypothèse, même si elle est incertaine ; ensuite, il faut qu'il reçoive une information objective, non punitive, qui lui permette de se corriger.

Il existe une stratégie qui répond à tous ces critères, et tous les enseignants la connaissent : dans le milieu pédagogique, on l'appelle... un test ! Ce qui est moins connu, c'est que des dizaines de publications scientifiques démontrent son efficacité : le fait de tester régulièrement ses connaissances, que les anglophones appellent *retrieval practise*, est l'une des stratégies pédagogiques les plus efficaces<sup>284</sup>. Se tester régulièrement maximise l'apprentissage à long terme. Le simple fait de mettre à l'épreuve sa mémoire la rend plus forte – c'est l'effet direct des principes d'engagement actif et de retour sur erreur décrits plus haut. Passer un test oblige à se frotter au réel et à se rendre compte de ce qu'on ne sait pas.

L'idée que les tests sont des moments clés de l'apprentissage ne va pas de soi. La plupart des enseignants et des élèves considèrent les contrôles et les examens comme de simples moyens de notation, une évaluation des

connaissances acquises ailleurs, durant le cours ou pendant les révisions. Or la recherche montre que le test joue un rôle au moins aussi important que le cours lui-même.

Dans une série d'expériences demeurées célèbres, le psychologue américain Henry Roediger et ses collaborateurs proposent à leurs élèves de mémoriser des mots en un temps fixe, mais selon plusieurs stratégies. À un groupe, on propose de passer tout son temps à étudier, en huit courtes séances. À un deuxième, on présente six séances d'apprentissage interrompues par deux tests. Enfin, au troisième groupe, on demande d'alterner quatre brèves séances d'étude et quatre tests. Notez bien que, dans la mesure où la durée totale est constante, ces tests supplémentaires réduisent le temps consacré à l'apprentissage. Et pourtant, le résultat est clair : 48 heures plus tard, la mémoire de la liste de mots est d'autant meilleure que les élèves ont eu la possibilité de se tester plus souvent. Alternier régulièrement l'apprentissage et le test les oblige à rester actifs et à recevoir un feedback explicite : je connais tel mot, mais je n'arrive jamais à me souvenir de tel autre. Cette connaissance de soi, cette « métamémoire » est utile, car elle leur permet, lors de la deuxième séance d'apprentissage, de mieux se concentrer sur les mots difficiles<sup>285</sup>. L'effet est manifeste : plus on se teste, mieux on retient le cours.

Autre exemple : imaginez que vous deviez apprendre des mots dans une langue étrangère, par exemple le fait qu'en langue inuit « traîneau » se dit « *qamutiik* ». Vous pouvez présenter les deux mots ensemble, côte à côte, afin de les associer mentalement. Ou bien, vous pouvez présenter d'abord l'un des mots puis, après cinq secondes, le second (figure 34). Notez bien que, une fois de plus, la seconde condition, paradoxalement, réduit la quantité d'information disponible : pendant les premières cinq secondes, vous ne voyez que le mot « *qamutiik* », sans qu'on vous rappelle ce qu'il signifie. Cependant, c'est cette seconde condition qui marche le mieux<sup>286</sup>. Pourquoi ? Parce qu'elle vous force à réfléchir, à essayer de retrouver le sens du mot avant de vous donner du feedback. Engagement actif et retour sur erreur maximisent l'apprentissage.

Le paradoxe, c'est que ni les étudiants ni leurs professeurs n'ont conscience de ces effets. Si on leur demande leur avis, tous considèrent que c'est l'étude qui

importe, pas le test. C'est pourquoi ils prédisent exactement l'inverse de ce qui est observé expérimentalement : selon eux, plus on étudie, mieux on réussit. Spontanément, les étudiants consacrent d'ailleurs tout leur temps à lire et à relire les cours, en stabilotant chaque ligne des couleurs de l'arc-en-ciel... Des stratégies bien moins efficaces que de se mettre à l'épreuve en passant un bref test.

Pourquoi avons-nous l'illusion que potasser le cours suffit à apprendre ? Parce que nous ne faisons pas la différence entre les différents compartiments de notre mémoire. Immédiatement après la lecture du cours, l'information est présente à l'esprit dans la mémoire de travail consciente, sous une forme active. Nous avons l'impression de la connaître... mais cette mémoire immédiate n'a rien à voir avec la mémoire à long terme dont nous aurons besoin pour récupérer l'information quelques jours plus tard. Au bout de cinq minutes, la mémoire de travail se dissipe déjà et, au bout de quelques jours, l'effet devient massif : il vaut beaucoup mieux avoir étudié, puis s'être testé, que d'avoir passé tout son temps à étudier ([figure 34](#)).

Rien de plus facile que de mettre, seul, ces idées en pratique. Il suffit de se préparer des cartes de révision : d'un côté, on écrit une question, et de l'autre la réponse. Pour se mettre à l'épreuve, on sort les cartes l'une après l'autre, et pour chacune on tente de se souvenir de la réponse (prédiction), puis on vérifie en retournant la carte (correction d'erreur). En cas d'échec, on replace la carte vers le dessus de la pile, ce qui la force à revenir à très court terme. En cas de réussite, on la relègue à l'arrière-plan : pas besoin de la réviser dans l'immédiat, mais elle resurgira plus tard, à un moment où l'oubli aura commencé à faire son effet. C'est le système des *flashcards* en vogue dans les facs américaines et qui est à la base de nombreux logiciels d'apprentissage, tel le célèbre Duolingo pour les langues étrangères.

## *Une règle d'or : espacer les apprentissages*

Pourquoi le fait de se tester à intervalles réguliers a-t-il des effets aussi positifs ? Parce qu'il exploite l'une des stratégies les plus efficaces que les sciences de l'apprentissage ont découvertes : l'espacement des apprentissages. C'est la règle d'or : distribuer les périodes d'entraînement plutôt que de les agglomérer. Au lieu de grouper tout l'apprentissage en une seule fois, on alterne les périodes d'étude et de test, et on révise régulièrement à des intervalles de temps de plus en plus espacés.

Des décennies de recherches en psychologie expérimentale montrent que l'espacement est une stratégie bien plus efficace que l'enseignement en une seule fois<sup>287</sup>. Si l'on dispose d'un temps fixe pour apprendre, mieux vaut répartir les leçons dans le temps que de tout apprendre d'un bloc. La distribution de l'apprentissage sur plusieurs jours a des effets massifs : l'expérience montre que l'on peut multiplier sa mémoire d'un facteur 3 lorsqu'on révise à intervalles réguliers plutôt que de tenter d'apprendre en une seule fois. La règle est simple, et tous les musiciens la connaissent : mieux vaut quinze minutes de travail tous les jours de la semaine que deux heures concentrées sur une seule journée.

Pourquoi la stratégie d'espacement fonctionne-t-elle si bien ? L'imagerie cérébrale<sup>288</sup> montre que le regroupement des problèmes en une seule session diminue l'activité cérébrale, peut-être parce que l'information répétée perd progressivement de sa nouveauté. La répétition semble également créer une illusion de savoir, un excès de confiance dû à la présence de l'information en mémoire de travail – elle paraît disponible, on l'a en tête, on ne voit pas l'intérêt de travailler plus. Au contraire, l'espacement des apprentissages augmente l'activité cérébrale. Il semble créer un effet de « difficulté désirable » en interdisant le simple stockage en mémoire de travail, et en forçant ainsi les circuits sollicités à travailler plus.

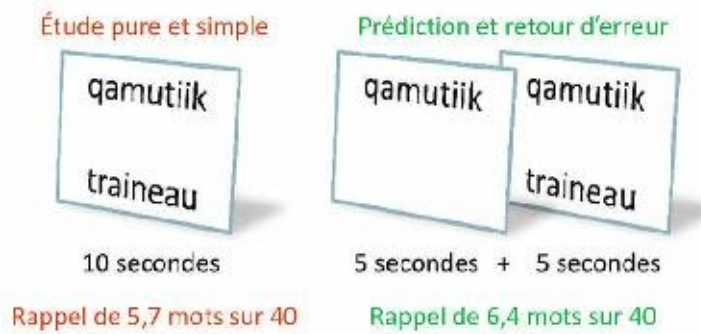
Quel est l'intervalle de temps le plus efficace entre deux répétitions de la même leçon ? Une forte amélioration est observée lorsque l'intervalle est de

vingt-quatre heures – probablement parce que le sommeil, comme nous le verrons dans un instant, joue un rôle central dans la consolidation des apprentissages. Mais le psychologue américain Hal Pashler et ses collègues montrent que l'intervalle optimal dépend de la durée de rétention en mémoire que l'on souhaite obtenir. Si l'on a besoin de se souvenir d'une information pendant quelques jours ou quelques semaines, alors il est idéal de la réviser tous les jours. Si, par contre, on souhaite que les connaissances soient préservées pendant plusieurs mois ou plusieurs années, il faut rallonger l'intervalle de révision en proportion. L'effet est massif : une seule répétition d'une leçon, quelques semaines après la première, multiplie par trois les nombres d'items qu'on parvient à rappeler quelques mois plus tard ! Pour garder l'information en mémoire le plus longtemps possible, le mieux est d'augmenter progressivement l'espacement temporel : on commence avec des leçons tous les jours, puis une révision au bout d'une semaine, d'un mois, d'un an... Cette stratégie garantit une mémoire optimale à chaque instant<sup>289</sup> .

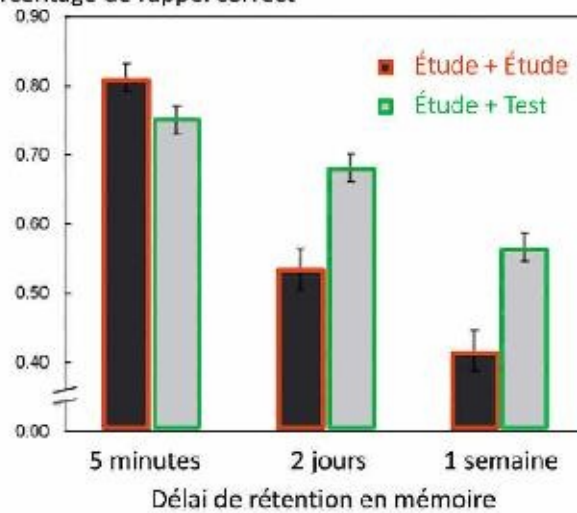
La figure 34 vous montre pourquoi : chaque révision renforce l'apprentissage, elle rafraîchit la force des représentations mentales et combat l'oubli exponentiel qui caractérise notre mémoire. Surtout, l'espacement des apprentissages à travers le temps semble sélectionner, parmi les différents circuits de mémoire disponibles dans notre cerveau, celui qui a la courbe d'oubli la plus lente, c'est-à-dire celui qui projette les informations le plus loin dans le futur.

Car nous nous trompons tous sur le rôle de la mémoire : c'est un système tourné vers le futur et non vers le passé. Le rôle de la mémoire n'est pas de regarder en arrière, mais au contraire d'envoyer une information dans l'avenir, parce que nous estimons qu'elle nous y sera utile. En répétant la même information plusieurs fois, à de longs intervalles, nous aidons notre cerveau à se convaincre que cette information est précieuse, qu'elle le sera pour longtemps et qu'elle mérite donc d'être préservée.

## L'importance de se tester régulièrement



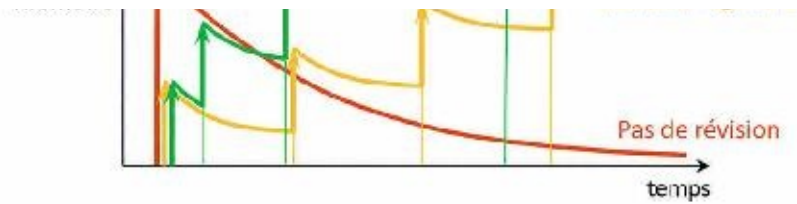
Pourcentage de rappel correct



## L'importance d'espacer les apprentissages







**Figure 34.** Se tester régulièrement est l'une des meilleures stratégies d'apprentissage, parce qu'elle nous force à prendre conscience de nos erreurs. Ainsi, si l'on doit apprendre des paires de mots, mieux vaut commencer par essayer de se souvenir avant de recevoir un retour sur erreur (en haut). L'expérience montre qu'il vaut mieux alterner les périodes d'étude et de test que de passer tout son temps à étudier (au milieu). À long terme, la mémoire est bien meilleure lorsqu'on espace les périodes de révision, surtout si l'on augmente progressivement les intervalles de temps (en bas).

Hal Pashler tire de ces travaux plusieurs enseignements pratiques. Premièrement, l'apprentissage gagne toujours à être réparti en plusieurs fois. Deuxièmement, à l'école, une révision au bout de quelques jours ou quelques semaines ne suffit pas. Si l'on aspire à mémoriser sur le long terme, il faut réviser après un intervalle de quelques mois au minimum. C'est donc toute l'organisation des manuels qu'il faut revoir. La plupart d'entre eux sont organisés en chapitres qui font le point sur un sujet spécifique (ce qui est bien) et qui sont suivis d'exercices qui ne portent que sur cette leçon (ce qui est moins bien). Cette organisation a deux conséquences néfastes : les révisions ne sont pas suffisamment espacées, et le travail est prémâché car l'élève n'apprend pas à déterminer par lui-même quelles connaissances ou quelles stratégies permettent de répondre à un problème donné. L'expérience montre qu'il vaut mieux mélanger les exercices, sans se cantonner à ce qui vient d'être étudié, de façon à remettre régulièrement en jeu la totalité de ses connaissances<sup>290</sup>.

*Quid* des partiels ou des examens de fin d'année ? Ce qu'on sait de l'apprentissage suggère qu'ils ne sont pas idéaux, parce qu'ils n'incitent pas à réviser régulièrement, mais qu'ils constituent tout de même un test utile des connaissances acquises. Même si les examens incitent à la révision de dernière minute, cette stratégie n'est pas forcément inefficace : à condition que l'élève ait déjà fait un effort d'apprentissage des mois plus tôt, réviser la veille de l'examen

rafraîchit les connaissances de façon durable. Toutefois, une révision régulière des connaissances, année après année, entraîne vraisemblablement les plus grands bénéfices pour l'élève. Les examens à court terme, qui ne portent que sur les acquis des semaines précédentes, ne suffisent pas à garantir un souvenir à long terme. Mieux vaut un examen cumulatif, qui porte sur l'ensemble du programme depuis le début de l'année.

Quel intérêt, pensez-vous peut-être, qu'un élève révise tout au long de l'année les mêmes choses ? Pourquoi lui faire refaire plusieurs fois, à différents intervalles de temps, un exercice qu'il a déjà réussi ? S'il réussit parfaitement ses examens, sans erreur, ne va-t-il pas rien apprendre du tout ? Pas du tout. Cela peut paraître paradoxal dans un chapitre consacré aux bénéfices de l'erreur, mais le bénéfice du feedback ne se limite pas aux items où l'élève se trompe. Bien au contraire, recevoir un retour sur erreur améliore la mémoire, même lorsqu'on a choisi la bonne réponse<sup>291</sup>. Pourquoi ? Parce que, comme je l'ai expliqué plus haut, tant que la connaissance n'est pas parfaite, le cerveau continue d'apprendre, même faiblement. Tant qu'il y a de l'incertitude, il y a de la surprise et des signaux d'erreur qui se propagent dans notre cerveau. L'incertitude agit comme une erreur virtuelle que nous aurions pu commettre et dont nous pouvons donc apprendre.

Ce qu'on appelle le surapprentissage est donc bénéfique : tant que nos connaissances ne sont pas absolument certaines, les réviser et les remettre à l'épreuve continuent d'améliorer nos performances, surtout sur le long terme. De plus, dans notre cerveau, la répétition présente d'autres bénéfices : elle automatise nos opérations mentales jusqu'à les rendre inconscientes. Cette consolidation constitue le quatrième pilier de l'apprentissage qu'il nous reste à examiner.

# CHAPITRE 10

## La consolidation

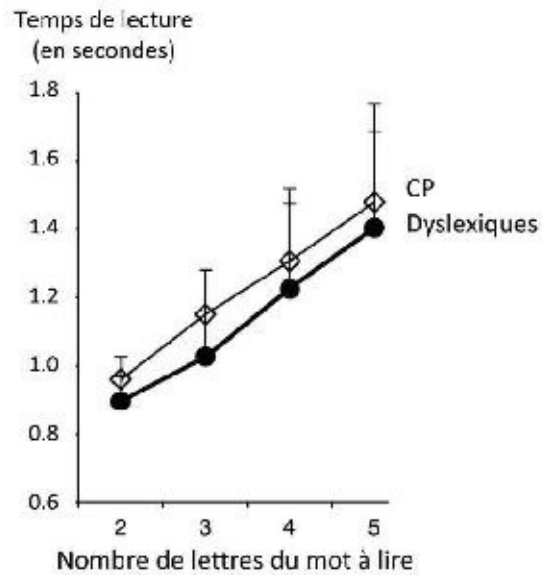
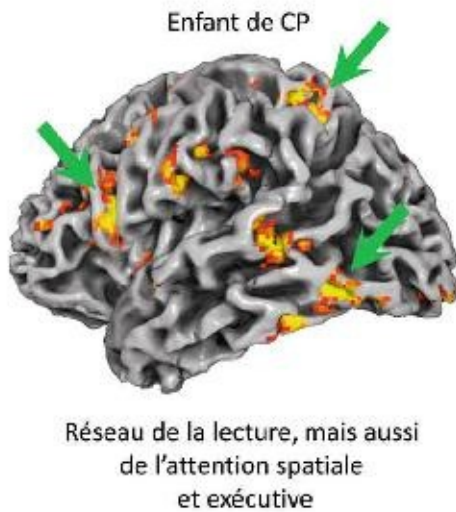
---

Prenez un bon élève en fin de CP. Sa première année d'école s'est bien passée. Soutenu par les trois premiers piliers de l'apprentissage, il a vite appris à lire. Il s'est engagé activement dans la lecture, avec curiosité et enthousiasme. Il a appris à prêter attention à chaque mot, à chaque lettre, de la gauche vers la droite. Et, au fil des mois, il est parvenu à corriger ses erreurs pour décrypter fidèlement les correspondances entre les lettres et les sons, et reconnaître les mots irréguliers. Pourtant, ce n'est pas encore un lecteur fluide. Il lit lentement et avec effort. Que lui manque-t-il ? Il lui reste encore à déployer le quatrième pilier de tout apprentissage : la consolidation. Celle-ci doit encore rendre automatique et inconsciente l'activité de lecture qui, pour l'instant, mobilise toute son attention.

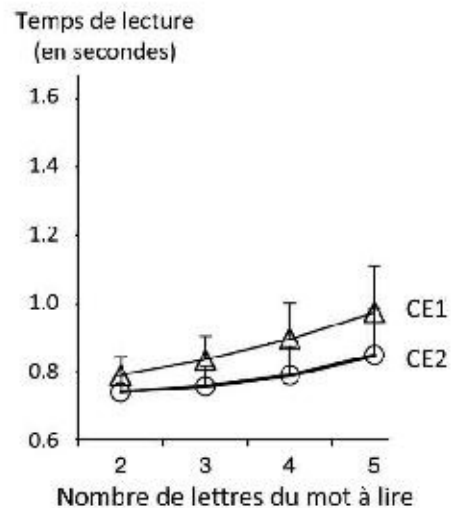
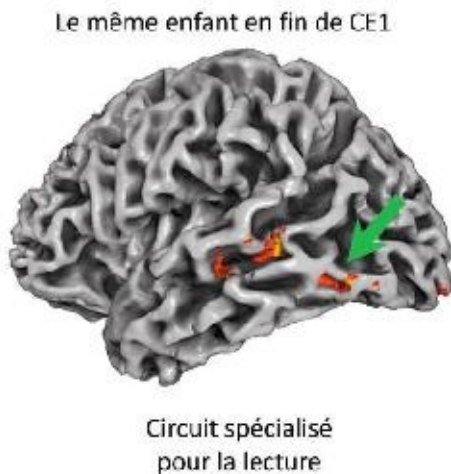
L'analyse de ses temps de réponse est révélatrice. Elle montre que plus les mots sont longs, plus il lui faut de temps pour les déchiffrer (figure 35). La courbe est linéaire, ce qui est caractéristique d'une opération sérielle, pas à pas : chaque lettre supplémentaire ajoute un cinquième de seconde au total. C'est tout à fait normal : à son âge, lire, c'est encore déchiffrer les lettres et les syllabes une par une<sup>292</sup>. Mais ce n'est pas définitif : avec la pratique, dans les deux années

qui suivent, sa lecture va s'accélérer et devenir plus fluide. Après deux ou trois années de pratique intensive, l'effet de la longueur des mots aura disparu. Devenu un lecteur efficace, il mettra le même temps pour lire un mot de 3 lettres ou de 8 lettres. Sa reconnaissance visuelle sera passée d'un traitement séquentiel à une opération parallèle : l'identification simultanée de toutes les lettres du mot.

### Lecture avec effort



### Lecture automatisée



**Figure 35.** La consolidation est le quatrième pilier de l'apprentissage. Au départ, tout apprentissage exige des efforts intensifs, et les régions pariétales et frontales associées à l'attention spatiale et exécutive s'activent. Ainsi, chez un lecteur débutant, le déchiffrement d'un mot demande de gros efforts. La lecture est séquentielle : plus le mot compte de lettres, plus sa lecture est lente (en haut). Avec la pratique, l'apprentissage s'automatise : il devient rapide, parallèle et inconscient (en bas). Un circuit spécialisé se met en place, libérant les ressources du cortex pour d'autres tâches.

La consolidation, c'est cela : passer d'un traitement lent, conscient, avec effort, à un fonctionnement rapide, inconscient, automatique. Notre cerveau ne s'arrête jamais d'apprendre. Même lorsqu'une compétence est maîtrisée, il continue de la surapprendre. Il dispose de mécanismes de routinisation qui « compilent » les opérations que nous utilisons régulièrement sous la forme de routines plus efficaces. Il les transfère dans d'autres régions du cerveau où elles pourront se dérouler inconsciemment, en toute autonomie, sans perturber les autres opérations en cours.

## *Libérer les ressources cérébrales*

Lorsqu'on scanne le cerveau d'un lecteur débutant, que voit-on ? Le circuit normal de la lecture, qui comprend à la fois les aires visuelles de reconnaissance des lettres et les régions temporales de traitement des phonèmes, des syllabes et des mots, est déjà activé – mais il s'accompagne également d'une activation massive de régions pariétales et préfrontales qui reflètent l'effort, l'attention spatiale et les stratégies conscientes<sup>293</sup>. Cette activité intense et gourmande en énergie va progressivement disparaître au fur et à mesure que l'apprentissage se consolide (figure 35). Chez un lecteur expert, les aires pariétales ne contribuent plus guère à la lecture – on ne les voit s'activer que si l'on perturbe la lecture, par exemple en espaçant les l e t t r e s, ou en les faisant tourner<sup>294</sup>.

Routiniser la lecture, c'est mettre en place un circuit restreint et spécialisé pour le traitement efficace des chaînes de lettres que nous rencontrons régulièrement. Avec l'apprentissage, nous développons un circuit extraordinairement efficace pour reconnaître les caractères et leurs combinaisons les plus fréquentes<sup>295</sup>. Notre cerveau compile des statistiques et détermine quelles lettres sont les plus fréquentes, où elles apparaissent le plus souvent et dans quelles associations elles surviennent. Même le cortex visuel primaire s'adapte à la forme et à la position des lettres familières<sup>296</sup>. Après quelques années de surapprentissage, ce circuit passe en mode de routine et parvient à fonctionner sans la moindre intervention consciente<sup>297</sup>. À ce stade, l'activation du cortex pariétal et préfrontal a disparu : nous lisons sans effort.

Ce qui est vrai pour la lecture vaut aussi pour tous les autres domaines de l'apprentissage. Lorsque nous pratiquons un instrument de musique, que nous apprenons à conduire une voiture ou à taper à la machine, nos gestes sont initialement sous le contrôle du cortex préfrontal : nous les produisons lentement, consciemment, un par un. Au bout de quelques séances, tout effort a disparu, et nous pouvons parler ou penser à autre chose : l'activité motrice s'est transférée dans le cortex moteur et surtout dans les noyaux gris centraux, un groupe de circuits sous-corticaux qui enregistrent nos comportements automatiques et routiniers (y compris les prières et les jurons !). Même chose dans le domaine de l'arithmétique : pour un enfant débutant, chaque calcul est une montagne qui demande de gros efforts d'attention et mobilise les circuits du cortex préfrontal. À ce stade, le calcul est séquentiel : pour résoudre  $6 + 3$ , l'enfant va typiquement compter les étapes une par une « 7, 8... 9 ! ». Au fil de la consolidation l'activité préfrontale s'évanouit, au profit des circuits spécialisés du cortex pariétal et temporal ventral<sup>298</sup>.

Pourquoi la routinisation est-elle si importante ? Parce qu'elle libère les ressources du cortex. Souvenez-vous que les réseaux de contrôle exécutif du cortex pariétal et préfrontal imposent un goulot d'étranglement cognitif : ils ne peuvent pas faire deux choses à la fois. Pendant qu'ils se concentrent sur l'exécution d'une tâche donnée, toutes les autres décisions conscientes sont ralenties ou abolies. Ainsi, tant qu'un apprentissage n'est pas automatisé, il

absorbe les précieuses ressources de l'attention exécutive et empêche l'enfant de se concentrer sur toute autre chose. Consolider un apprentissage, c'est rendre les ressources du cerveau disponibles pour d'autres objectifs.

Prenons un exemple concret. Imaginez que vous deviez, simultanément, vous concentrer sur la lecture d'un texte et résoudre un problème de maths, comme un lecteur débutant :

« Hin hotomobilist kit Nanth pour Pari à katorzeure. La distensse ai de troa sans quilomaitre. Ile harive à dice-setteure. Kaile été sa vitaisse moi hyène ? »

La difficulté est claire : il est pratiquement impossible de faire les deux choses en même temps. La difficulté de lecture abolit toute capacité de réflexion arithmétique. Pour progresser, il est indispensable que les outils mentaux qui nous sont les plus utiles, tels que la lecture ou l'arithmétique, deviennent comme une seconde nature, qu'ils opèrent inconsciemment et sans effort. Nous ne pourrions pas construire les plus hauts niveaux de la pyramide éducative sans en avoir d'abord consolidé les fondations.

## *Le rôle clé du sommeil*

Nous avons déjà vu que l'apprentissage bénéficiait beaucoup d'un espacement régulier : plutôt que d'avaler toute une leçon en une seule journée, mieux vaut distribuer les apprentissages tous les jours. La raison en est simple : toutes les nuits, notre cerveau consolide ce qu'il a appris pendant la journée. C'est sans doute l'une des plus importantes découvertes des neurosciences des trente dernières années : le sommeil n'est pas une simple période d'inactivité ou de nettoyage des déchets que le cerveau a accumulés pendant la journée. Bien au contraire : tandis que nous dormons, notre cerveau se répète les événements

importants qu'il a enregistrés pendant la veille et, progressivement, il les transfère dans un compartiment plus efficace de notre mémoire.

La découverte remonte à 1924<sup>299</sup>. Deux psychologues américains, John Jenkins et Karl Dallenbach, reprennent les études classiques sur la mémoire. Ils réexaminent les travaux du pionnier de la mémoire, l'Allemand Hermann Ebbinghaus. Dès la fin du XIX<sup>e</sup> siècle, celui-ci avait démontré que l'oubli suivait une courbe exponentielle : plus on laisse passer de temps et moins on se souvient de ce qu'on a appris. Magnifique loi mathématique de la mémoire – sauf qu'il y a une anomalie dans la courbe d'oubli d'Ebbinghaus : elle ne montre aucune perte de mémoire entre 8 et 14 heures après l'apprentissage. Jenkins et Dallenbach ont une illumination : dans l'expérience d'Ebbinghaus, le délai de 8 heures correspond à des tests passés au cours de la même journée, tandis que celui de 14 heures correspond à des tests espacés d'une nuit de repos. Pour en avoir le cœur net, ils conçoivent une nouvelle expérience qui sépare les deux variables : le délai écoulé et le fait d'avoir dormi ou pas. Pour ce faire, ils enseignent à leurs étudiants des syllabes aléatoires, soit vers minuit, juste avant de dormir, soit le matin. Le résultat est clair : l'apprentissage du matin s'effondre avec le temps selon la loi exponentielle d'Ebbinghaus ; par contre, l'apprentissage de minuit reste stable dans le temps (à condition d'avoir bénéficié d'au moins deux heures de sommeil). En d'autres termes, dormir semble empêcher d'oublier.

Plusieurs autres interprétations de ces résultats viennent cependant à l'esprit. Peut-être le problème vient-il de la veille : pendant qu'il est éveillé, le cerveau accumule peut-être des toxiques qui sont éliminés pendant le sommeil ; ou bien, peut-être la mémoire souffre-t-elle de l'interférence causée par les événements qui surviennent dans l'intervalle entre l'apprentissage et le test, et qui n'ont pas lieu durant le sommeil. Mais toutes ces idées sont définitivement rejetées en 1994 lorsque des chercheurs israéliens démontrent que le sommeil *améliore* l'apprentissage<sup>300</sup>. Sans aucun entraînement supplémentaire, les performances cognitives et motrices s'améliorent après une période de sommeil. L'expérience est simple. Pendant la veille, un volontaire apprend à détecter une barre à un endroit précis de la rétine. Ses performances s'améliorent lentement, mais au



bout de quelques heures d'entraînement elles ne bougent plus : la limite semble être atteinte. On laisse alors la personne dormir, et surprise, lorsqu'on la reteste au réveil, sa réussite est bien meilleure, et elle le reste tout au long de la journée et des suivantes. Et c'est bien le sommeil qui cause le surcroît d'apprentissage, car si l'on réveille le sujet pendant la nuit, à chaque fois qu'il entre en sommeil paradoxal, il ne montre aucune amélioration à son réveil.

De nombreux travaux viendront ensuite confirmer ces découvertes<sup>301</sup>. La quantité d'apprentissage varie directement en fonction de la durée de sommeil, et surtout de sa profondeur. En plaçant des électrodes sur le scalp, on peut évaluer la qualité du sommeil, et cette mesure prédit à quel point les performances sont meilleures au réveil. Et la relation semble fonctionner dans le sens inverse : le besoin de sommeil semble dépendre des événements qui se sont produits pendant la veille. Chez l'animal, un gène impliqué dans la plasticité cérébrale, zif-268, voit son expression augmenter dans l'hippocampe et le cortex au cours du sommeil paradoxal, spécifiquement chez les rats qui ont été préalablement exposés à un environnement enrichi : le surcroît de stimulation entraîne une augmentation de la plasticité nocturne<sup>302</sup>.

Les rôles respectifs des différents stades de sommeil ne sont pas encore parfaitement établis, mais il semblerait que le sommeil profond permette la consolidation et la généralisation des connaissances (mémoire sémantique ou déclarative), tandis que le sommeil paradoxal, pendant lequel l'activité cérébrale est rapide et proche de l'éveil, renforce les apprentissages perceptifs et moteurs (mémoire procédurale).

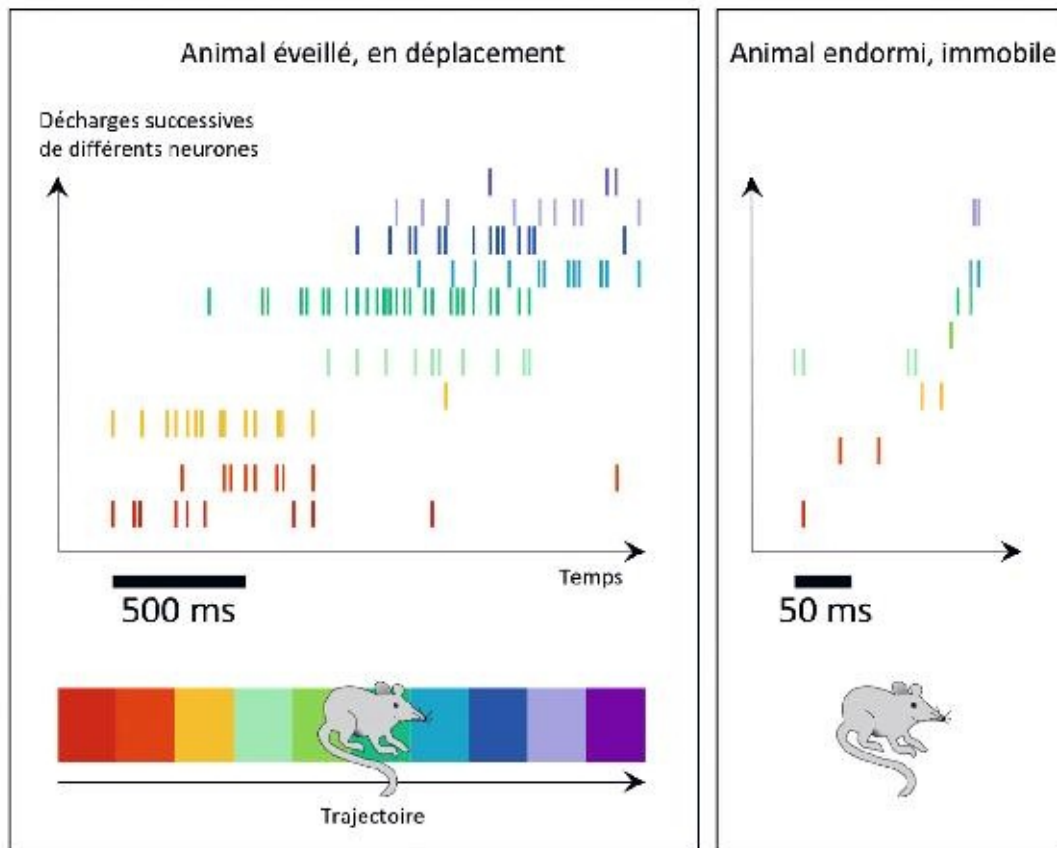
## *Le cerveau endormi revit les épisodes de la veille*

Restait encore à comprendre le mécanisme neuronal par lequel un cerveau endormi parvenait à apprendre, mieux encore qu'à l'état de veille.

En 1994, les neurophysiologistes Matt Wilson et Bruce McNaughton font une découverte remarquable : en l'absence de toute stimulation extérieure, les neurones de l'hippocampe s'activent spontanément pendant le sommeil<sup>303</sup>. Et cette activité n'est pas aléatoire : elle retrace les pas que l'animal a faits dans la journée ! Comme nous l'avons vu au [chapitre 3](#), en effet, l'hippocampe contient des cellules de lieux, c'est-à-dire des neurones qui déchargent lorsque l'animal se balade en un certain point de l'espace. On y trouve toute une série de neurones qui, chacun, préfère un endroit différent, de sorte que si on en enregistre suffisamment, ils finissent par paver tout l'espace dans lequel se promène l'animal. Lorsque qu'il se déplace, disons, dans un couloir, certains neurones déchargent à l'entrée du couloir, d'autres au milieu, d'autres plutôt à son extrémité. Ainsi, la promenade du rat entraîne l'entrée en activité successive de toute une série de cellules de lieux : le mouvement dans l'espace réel devient une séquence temporelle dans l'espace neuronal.

Et c'est là qu'intervient la découverte de l'équipe de McNaughton : lorsque le rat s'immobilise et s'endort, les mêmes neurones de l'hippocampe se remettent à décharger dans le même ordre. Seule différence : la vitesse. Elle peut être normale, mais aussi accélérée d'un facteur 20. En dormant, le rat rêve d'une balade à vitesse accélérée !

La relation entre l'activité des neurones de l'hippocampe et la position de l'animal est si fidèle que l'on parvient aujourd'hui à inverser le processus et à décoder, sur la base des décharges neuronales, le contenu du rêve<sup>304</sup>. On utilise d'abord les données enregistrées au cours de la veille, lorsque l'animal se promène dans le monde réel. Elles permettent d'entraîner un décodeur : un logiciel informatique découvre la relation systématique entre les décharges neuronales et la position de l'animal. Ensuite, on applique ce décodeur aux données enregistrées durant le sommeil – et on constate que, tandis qu'il dort, le cerveau de l'animal trace, en rêve, des trajectoires virtuelles dans l'espace.



**Figure 36.** Le sommeil joue un rôle important dans la consolidation des apprentissages. Lorsqu'un rat s'endort, les neurones de l'hippocampe rejouent, souvent à vitesse accélérée, les séquences d'activité qu'il a vécues lorsqu'il était éveillé. Cette activité, qui s'étend jusqu'au cortex, se répète des centaines de fois pendant la nuit. Au réveil, les apprentissages des jours précédents sont mieux consolidés et automatisés. Notre cerveau peut même découvrir, pendant le sommeil, des régularités qui nous avaient échappé pendant la veille.

Le cerveau du rat rejoue ainsi les patrons d'activité qu'il a éprouvés la veille. Chaque nuit voit renaître les souvenirs de la journée. Ce phénomène n'est pas confiné à l'hippocampe, mais s'étend au cortex. Celui-ci joue un rôle déterminant dans la plasticité synaptique et la consolidation des apprentissages. Grâce à la réactivation nocturne, même un événement unique, une fois enregistré dans la mémoire épisodique, pourra être rejoué des centaines de fois (figure 36). Ce transfert pourrait même être la fonction principale du sommeil<sup>305</sup>. L'hippocampe engrangerait les souvenirs de la journée dans une mémoire

rapide, tandis qu'au cours de la nuit, la réactivation de ces signaux permettrait d'entraîner d'autres réseaux neuronaux supplémentaires, principalement situés dans le cortex – une mémoire plus lente, mais capable d'extraire un maximum d'informations de chaque épisode. Et effectivement, dans le cortex d'un rat qui apprend à réaliser une tâche nouvelle, plus un neurone se réactive pendant la nuit, plus il augmente sa participation à la tâche au cours de la journée qui suit<sup>306</sup>. Réactiver aide à automatiser.

Le même phénomène existe-t-il chez l'homme ? Oui. L'imagerie cérébrale du cerveau humain montre que, durant le sommeil, les circuits sollicités au cours de la journée précédente se réactivent<sup>307</sup>. Après avoir beaucoup joué à Tetris, un jeu vidéo où des pièces de puzzle ne cessent de tomber du haut de l'écran, au cours de la nuit qui suit vous allez littéralement halluciner, dans votre rêve, une cascade de formes géométriques, et vos yeux feront même des mouvements correspondants, du haut vers le bas. Mieux : une expérience récente a fait dormir des volontaires dans l'IRM, et les a réveillés dès que leur électroencéphalogramme suggérait qu'ils étaient en train de rêver. L'IRM montrait que de nombreuses régions de leur cerveau s'étaient spontanément activées juste avant qu'on ne les sorte du sommeil, et que leur activité prédisait le contenu de leur rêve. Ainsi, lorsque le dormeur disait, par exemple, avoir rencontré des personnes en rêve, on pouvait voir que l'aire corticale des visages s'était activée pendant qu'il dormait. D'autres expériences montrent que l'étendue de cette réactivation prédit non seulement le contenu du rêve, mais aussi la quantité de consolidation de la mémoire après le réveil. On commence même à enregistrer des neurones dans le cerveau humain, et on constate que, comme chez le rat, leurs décharges retracent les séquences observées dans la journée qui précède.

Le lien entre sommeil et apprentissage est solide. De nombreuses expériences montrent que les variations spontanées de la profondeur du sommeil corrélaient avec les performances après l'éveil. Lorsqu'on apprend à se servir d'un joystick, par exemple, au cours de la nuit qui suit, la fréquence et l'intensité des ondes lentes de sommeil augmentent en regard des régions pariétales du cerveau qui sont impliquées dans cet apprentissage – et plus cette augmentation est forte,

plus les performances de la personne progressent<sup>308</sup>. De même, après un apprentissage moteur, l'imagerie cérébrale montre que l'activité augmente dans le cortex moteur, l'hippocampe et le cervelet, tandis qu'elle diminue dans certaines aires frontales, pariétales et temporales<sup>309</sup>. Tous ces résultats suggèrent que le sommeil facilite l'automatisation. Après avoir dormi, l'activité cérébrale se déplace : une partie des connaissances acquises dans la journée est transférée vers des circuits plus automatiques et plus spécialisés.

Pour vérifier que ce lien est causal, il est aujourd'hui possible d'augmenter artificiellement la profondeur du sommeil. L'idée est de créer un effet de résonance dans le cerveau. Au cours du sommeil, l'activité cérébrale oscille spontanément à une fréquence lente, de l'ordre de 40 à 50 cycles par minute. En injectant une petite impulsion à la bonne fréquence, on peut faire entrer ces rythmes en résonance et augmenter leur intensité – un peu comme lorsqu'on pousse une balançoire jusqu'à la faire osciller avec une amplitude énorme. L'équipe de Jan Born l'a fait de deux manières différentes : soit en injectant de minuscules courants à travers la boîte crânienne, soit, de façon bien plus simple, en jouant dans un haut-parleur un bruit diffus synchronisé avec les ondes cérébrales du sommeil. Qu'il soit bercé par ce bruit de vagues ou biaisé par la stimulation électromagnétique, le cerveau du dormeur se laisse entraîner par ce rythme irrésistible, et il produit significativement plus d'ondes lentes caractéristiques du sommeil profond. Et dans les deux cas, cet entraînement s'accompagne d'une meilleure consolidation des apprentissages<sup>310</sup>.

Une start-up française commence à exploiter cet effet : elle vend un bandeau censé faciliter l'endormissement et augmenter la profondeur du sommeil, parce qu'il envoie de petits bruits qui stimulent les rythmes lents du cerveau nocturne. D'autres chercheurs envisagent d'augmenter l'apprentissage en forçant le cerveau à réactiver, pendant la nuit, certains souvenirs plus que d'autres. Imaginez que la salle de classe soit parfumée d'une entêtante odeur de rose. Au cours de votre sommeil profond, on pulvérise cette odeur dans votre chambre. Le résultat est net : le lendemain matin, l'apprentissage est bien mieux consolidé que si vous aviez dormi sans être exposé à cette odeur<sup>311</sup>. L'odeur de rose sert d'indice inconscient qui incite votre cerveau à réactiver cet épisode particulier de

vosre journée, ce qui augmente sa consolidation en mémoire. On peut obtenir le même effet avec des indices auditifs. Imaginez qu'on vous demande de mémoriser la position d'une cinquantaine d'images, chacune associée à un son (le chat fait miaou, la vache fait meuh...). C'est une tâche difficile... mais la nuit porte conseil. Pendant le sommeil, on rejoue certains de ces sons. Les entendre, même inconsciemment, au plus profond de votre sommeil, biaise la réactivation nocturne de vos neurones et accroît votre mémoire visuelle de la carte correspondante<sup>312</sup>.

À l'avenir, manipulerons-nous tous notre sommeil afin de mieux apprendre ? Beaucoup d'étudiants le font déjà spontanément : ils révisent une leçon importante juste avant de s'endormir. Attention toutefois à ne pas confondre ces découvertes, dont l'efficacité est démontrée, avec l'idée fausse que l'on peut acquérir des compétences nouvelles tout en dormant. Certains charlatans vendent des enregistrements audio censés vous faire apprendre une langue étrangère pendant la nuit – l'expérience montre qu'ils n'ont aucun effet<sup>313</sup>. Le cerveau endormi n'apprend pratiquement rien, il ne peut que rejouer ce qu'il a déjà éprouvé. Pour apprendre une compétence aussi complexe qu'une nouvelle langue, la seule chose qui fonctionne, c'est de s'entraîner pendant la journée, puis de dormir pour laisser le sommeil réactiver et consolider ce que nous avons appris.

## *Peut-on faire des découvertes en dormant ?*

Le cerveau endormi se contente-t-il d'enregistrer ce que nous avons appris ? De nombreux scientifiques pensent que non : ils disent faire des découvertes en dormant. Le plus célèbre d'entre eux est le chimiste allemand Friedrich August Kekulé von Stradonitz, qui a découvert la structure du benzène – une molécule inhabituelle, puisque ses six atomes de carbone forment une boucle fermée,

comme un anneau ou comme... un serpent qui se mord la queue. Car c'est ainsi que Kekulé décrit son rêve : « Une fois encore les atomes tourbillonnaient devant moi [...]. Dans mon esprit, qui était empli d'innombrables images semblables, je voyais de grandes formes étranges et de longues chaînes. Les formes se tordaient comme des serpents. Soudain, quelque chose se passa. Un serpent se mordit la queue et engendra une forme semblable à un anneau qui tournait devant mes yeux. Je me sentis comme frappé par une illumination et je me réveillai. » Et Kekulé de conclure : « Apprenons à rêver, mesdames et messieurs, et seulement alors, nous pourrons apprendre la vérité ! »

Le sommeil a-t-il vraiment le pouvoir de nous conduire sur le chemin de la vérité et d'augmenter notre créativité ? Les historiens des sciences sont partagés sur l'authenticité de l'épisode du serpent Ouroboros de Kekulé, mais l'idée d'incubation nocturne est répandue chez les scientifiques et les artistes. Le designer Philippe Starck, créateur de centaines d'objets aux formes inédites, le disait avec humour dans une interview récente : « Quand je vais me coucher, je dis à ma femme : bon, je vais au boulot<sup>314</sup> ! »

J'ai moi-même fait maintes fois l'expérience de découvrir, à mon réveil, la solution d'un problème difficile de la veille. Cependant, une collection d'anecdotes, si longue soit-elle, n'a aucune valeur probante. Il faut expérimenter – et c'est exactement ce qu'a fait l'équipe de Jan Born, spécialiste allemand du sommeil<sup>315</sup>. Pendant la journée, il a enseigné à des volontaires un algorithme complexe. Il s'agissait d'appliquer toute une série de calculs à un chiffre de départ. Or, à leur insu, le problème possédait un raccourci caché : une astuce permettait de réduire considérablement le temps de calcul. Avant de dormir, bien peu de sujets en prenaient conscience. Cependant, une bonne nuit de sommeil doublait le nombre de volontaires qui découvraient le raccourci, alors qu'aucun *eurêka* ne survenait chez ceux qu'on empêchait de dormir. Les résultats ne dépendaient pas de l'heure de la journée à laquelle ils étaient testés. Ainsi le temps écoulé n'était pas le facteur déterminant, mais bien le sommeil lui-même.

La consolidation nocturne ne se réduit donc pas au simple renforcement des connaissances existantes. Les découvertes de la journée ne sont pas seulement stockées, mais également recodées sous une forme plus abstraite et plus

générale. La réactivation neuronale nocturne joue sans doute ici un rôle crucial. Chaque nuit, nos idées flottantes de la journée se rejouent des centaines de fois, à un rythme accéléré, ce qui multiplie les chances que notre cortex y découvre des règles qui fassent sens. De plus, l'accélération des décharges neuronales comprime l'information. Grâce à elle, des neurones qui, à l'état de veille, n'étaient activés qu'à de longs intervalles se retrouvent adjacents dans la séquence de la nuit. Ce mécanisme semble idéal pour regrouper, synthétiser, « convertir des informations brutes en connaissances utiles et exploitables » – la définition même de l'intelligence selon le spécialiste d'intelligence artificielle Demis Hassabis.

Dans le futur, les machines intelligentes dormiront-elles ? La question semble farfelue, et pourtant je pense que oui – leur algorithme d'apprentissage aura probablement besoin d'une phase analogue à ce que nous appelons le sommeil. En effet, les informaticiens ont d'ores et déjà conçu plusieurs algorithmes d'apprentissage par ordinateur qui exploitent l'alternance veille/sommeil<sup>316</sup>. Ces algorithmes jouent un rôle essentiel dans la vision nouvelle de l'apprentissage que je défends dans ce livre, et selon laquelle apprendre consiste à sélectionner un modèle génératif du monde extérieur. Souvenez-vous : notre cerveau contient d'immenses modèles internes, capables de resynthétiser des images mentales plus vraies que nature, des dialogues réalistes et des déductions qui font sens. À l'état de veille, nous ajustons ces modèles : les données sensorielles que nous recevons du monde extérieur nous servent à sélectionner ceux qui collent le mieux au monde qui nous entoure. C'est un fonctionnement du bas vers le haut : les messages d'erreur, issus des aires sensorielles de bas niveau, montent dans la hiérarchie du cortex, et y ajustent nos pondérations statistiques, afin que notre modèle descendant soit plus juste.

L'idée nouvelle est qu'au cours du sommeil, notre cerveau fonctionne plutôt dans le sens inverse, du haut vers le bas : nous utilisons nos modèles génératifs pour synthétiser des images nouvelles, jamais vues auparavant, et une partie de notre cerveau s'entraîne à reconnaître ces images créées de toutes pièces. Pendant le sommeil, ce sont donc nos connexions montantes qui s'ajustent : elles



détectent quelles régularités du monde extérieur pointent vers tel ou tel modèle interne. C'est ainsi que nous devenons de plus en plus efficaces. Après une bonne nuit de sommeil, le moindre indice nous suffit pour retrouver le meilleur modèle mental de la réalité, même s'il est excessivement abstrait.

Selon cette idée, nos rêves ne seraient rien d'autre que des images d'entraînement, des reconstitutions de la réalité que nous utilisons pour démultiplier les expériences (forcément limitées) que nous avons faites pendant la journée. Le sommeil résoudrait un problème auquel sont confrontés tous les algorithmes d'apprentissage par ordinateur : la petite taille des données disponibles. En effet, les réseaux de neurones artificiels ont besoin d'énormes jeux de données pour apprendre – mais la vie est trop courte, et le cerveau doit se contenter du nombre limité d'informations qu'il parvient à recueillir durant la veille. Grâce au sommeil, le cerveau simule peut-être, en accéléré, le temps d'une nuit, des myriades d'expériences qu'une vie entière ne suffirait pas à faire.

Au cours de ces « expériences de pensée », il se peut que nous fassions des découvertes nouvelles. Il n'y a là rien de magique : nous faisons tourner, dans notre tête, notre machine à simuler le monde, et nous y découvrons des conséquences inattendues – un peu comme un joueur d'échecs, une fois les règles maîtrisées, passe des années à découvrir certaines de leurs conséquences.

On doit à l'imagerie mentale certaines des plus grandes découvertes scientifiques de l'humanité – lorsque Einstein rêve qu'il chevauche un photon ou que Newton imagine que la lune tombe comme une pomme. Sait-on que la célèbre expérience de Galilée, dans laquelle il laisse tomber des objets de la tour de Pise afin de prouver que la chute des corps ne dépend pas de leur masse, n'a peut-être jamais eu lieu ? C'est avant tout une expérience de pensée : Galilée imagine qu'on laisse tomber deux sphères, l'une légère et l'autre lourde, du haut de la tour de Pise. Il raisonne par l'absurde. Supposons, dit-il, que la plus lourde tombe plus vite. Imaginons que je l'attache à l'autre par un fil de masse négligeable. L'ensemble des deux sphères, puisqu'il forme un objet plus lourd, devrait tomber encore plus vite. Mais c'est absurde, car la sphère légère, qui tombe moins vite, devrait ralentir la plus lourde. Ce faisceau de contradictions

n'a qu'une seule issue : tous les objets tombent à la même vitesse, quelle que soit leur masse.

Voilà le type d'argument dont notre simulateur mental est capable, dans la veille comme dans le sommeil. Le fait même que nous sachions dérouler de tels faisceaux imaginaires met en valeur l'extraordinaire panoplie d'algorithmes de notre cerveau. Nous apprenons bien sûr pendant la veille, mais l'activité nocturne de notre cerveau démultiplie notre potentiel. Il n'est d'ailleurs pas impossible que ce soit l'un des secrets de l'espèce humaine : il semblerait que, parmi tous les primates, notre sommeil soit l'un des plus profonds et des plus efficaces<sup>317</sup>.

## *Sommeil, enfance et école*

Qu'en est-il chez l'enfant ? Chacun sait que les bébés passent la plupart de leur temps à dormir, et que le sommeil se raccourcit avec l'âge. C'est logique : la petite enfance est une période privilégiée d'apprentissages, et les algorithmes de notre cerveau doivent sans doute tourner beaucoup plus. De fait, l'expérience montre qu'à durée égale, le sommeil de l'enfant est deux à trois fois plus efficace que celui de l'adulte. Après un apprentissage intensif, un enfant de 10 ans plonge beaucoup plus vite en sommeil profond qu'un adulte. Ses ondes lentes de sommeil sont plus intenses, et le résultat est net : lorsqu'il étudie une séquence, sombre dans le sommeil et se réveille le lendemain frais et dispos, on constate qu'il a découvert nettement plus de régularités qu'un adulte<sup>318</sup>.

La consolidation nocturne fonctionne dès les premiers mois de la vie. Chez le bébé de moins de 1 an, l'apprentissage des mots dépend étroitement du sommeil. Un bébé qui fait la sieste, ne fût-ce qu'une heure et demie, retient bien mieux les mots qu'on lui a enseignés avant de s'endormir<sup>319</sup>. Surtout, il les généralise mieux : la première fois qu'il entend le mot « cheval », il ne lui associe qu'une ou deux images spécifiques, mais après avoir dormi, son cerveau parvient à associer ce mot à de nouveaux exemplaires jamais vus auparavant. Tel

un Kekulé en couche-culotte, il fait des découvertes pendant son sommeil et se réveille avec une bien meilleure théorie du mot « cheval ».

Et à l'école ? Les expériences sont très claires : chez l'enfant de maternelle, même une brève sieste dans l'après-midi améliore l'apprentissage de la matinée<sup>320</sup>. Pour un bénéfice maximal, le sommeil doit survenir dans les heures qui suivent cet apprentissage. Ce bénéfice, toutefois, n'existe que chez les enfants qui font régulièrement la sieste. Il ne semble donc pas utile de forcer tous les enfants à dormir, mais seulement de laisser dormir ceux qui en éprouvent le besoin. Notre cerveau régule naturellement son besoin de sommeil en fonction de la quantité de stimulations nouvelles de la journée.

Hélas, le sommeil des enfants, comme celui des adultes, est aujourd'hui menacé de toutes parts. Quelles en sont les conséquences ? Le manque chronique de sommeil va-t-il aller jusqu'à causer des troubles spécifiques des apprentissages ? Ce n'est encore qu'une hypothèse, mais les arguments s'accumulent : une fraction non négligeable des enfants hyperactifs et atteints de troubles de l'attention pourrait ne souffrir que d'un manque chronique de sommeil. Chez certains, on détecte des troubles de la respiration qui les empêchent de plonger en sommeil profond – et un simple dégagement des voies respiratoires suffit parfois à faire disparaître non seulement les troubles du sommeil, mais aussi ceux de l'attention. Une expérience récente suggère que la stimulation électrique du cerveau, en augmentant la profondeur des ondes lentes du sommeil, pourrait également corriger le déficit d'apprentissage chez certains enfants hyperactifs<sup>321</sup>.

Entendons-nous bien : ces données demandent encore à être répliquées, et je ne suis aucunement en train de nier l'existence d'authentiques troubles de l'attention (des enfants chez qui un médicament, la Ritaline, peut avoir des effets très positifs). Mais, sur le plan éducatif, il ne fait aucun doute que l'amélioration de la durée et de la qualité du sommeil constitue une intervention efficace chez *tous* les enfants, et notamment ceux en difficulté scolaire ou qui présentent des troubles des apprentissages.

L'idée a d'ailleurs été testée chez les adolescents. Chez eux, la chronobiologie a démontré qu'avec la puberté, le cycle de sommeil se décale :

les adolescents ne ressentent pas le besoin de se coucher tôt, mais, comme chacun a pu le constater, ils éprouvent les plus grandes difficultés à se lever. Ce n'est pas de la mauvaise volonté, mais la simple conséquence d'une réorganisation des réseaux neuronaux et hormonaux qui contrôlent le cycle veille/sommeil. Malheureusement, personne ne semble en avoir informé les directeurs d'école qui continuent d'exiger que les élèves fassent acte de présence en tout début de matinée. Et si on changeait cette convention somme toute arbitraire ? L'expérience a été menée avec succès : différez l'entrée à l'école d'une demi-heure ou d'une heure et tout change : les adolescents dorment mieux, leur absentéisme se réduit, leur attention en classe augmente, et leurs performances scolaires s'améliorent<sup>322</sup> (et je pourrais prolonger la liste : l'obésité régresse, tout comme la dépression, les adolescents ont moins d'accidents, etc.). C'est un magnifique exemple d'adaptation du système éducatif aux contraintes de la biologie cérébrale !

# CONCLUSION

## Réconcilier l'éducation avec les neurosciences

---

« Traiter de la façon d'élever et d'éduquer les enfants semble être la chose la plus importante et la plus difficile de toute la science humaine. »

MONTAIGNE , *Essais* (1580).

« La pédagogie est comme la médecine : un art, mais qui s'appuie – ou devrait s'appuyer – sur des connaissances scientifiques précises. »

Jean PIAGET ,  
*La Pédagogie moderne* (1949).

Nous arrivons au terme de ce voyage. J'espère vous avoir convaincu que, grâce aux avancées de la psychologie cognitive, des neurosciences, de l'intelligence artificielle et des sciences de l'éducation, nous disposons de connaissances détaillées sur la manière dont notre cerveau apprend. Ces connaissances ne vont pas de soi, et la plupart de nos idées sur l'apprentissage sont à revoir :

- Non, le bébé n'est pas une ardoise vierge de tout savoir – il dispose, dès la première année de vie, d'une vaste panoplie de connaissances sur les objets, les nombres, les probabilités, l'espace ou les personnes.
- Non, l'enfant n'est pas une éponge qui absorbe docilement la structure de son environnement – l'expérience de Felipe, aveugle et tétraplégique, ou de Nicholas Saunderson, mathématicien aveugle, nous montre que les entrées sensorielles peuvent être bouleversées ou absentes sans que cela abolisse les capacités d'abstraction.
- Non, le cerveau n'est pas qu'un réseau de neurones malléable qui se laisse façonner par ses entrées – tous les grands faisceaux de connexions sont présents dès la naissance, et la plasticité cérébrale, pour indispensable qu'elle soit, ne fait qu'affiner les derniers millimètres de nos connexions.
- Non, l'apprentissage ne se produit pas de façon passive, par simple exposition à des données ou à un cours magistral. Au contraire, la psychologie cognitive et l'imagerie cérébrale nous montrent que l'enfant est un scientifique en herbe, sans cesse en train de générer des hypothèses nouvelles, et le cerveau un organe toujours en alerte qui apprend en mettant à l'épreuve les modèles qu'il projette sur le monde extérieur.
- Non, l'erreur n'est pas la marque des mauvais élèves – se tromper fait partie intégrante de l'apprentissage, car notre cerveau ne peut ajuster ses modèles que lorsqu'il découvre un décalage entre ce qu'il envisageait et la réalité.
- Non, le sommeil n'est pas qu'une période de repos – c'est une partie intégrante de notre algorithme d'apprentissage, un moment privilégié où notre cerveau fait tourner ses modèles en boucle et amplifie, d'un facteur 10 ou 100, les acquis de la journée.
- Et non, les machines actuelles ne sont pas près de dépasser le cerveau humain – celui-ci reste, pour l'instant, le plus rapide, le plus efficace et le moins gourmand en énergie de tous les dispositifs de traitement de l'information. Véritable machine probabiliste, il parvient à extraire les

moindres bribes d'information de chaque épisode de la journée, et à les transformer pendant la nuit en connaissances abstraites et générales, d'une manière que nous ne savons pas encore reproduire dans nos ordinateurs.

Dans la bataille prométhéenne à laquelle se livrent la puce et le neurone, la machine et le cerveau, ce dernier conserve encore l'avantage. En principe, il n'y a rien dans la mécanique cérébrale qu'une machine ne puisse imiter – d'ailleurs, toutes les idées que j'ai énoncées dans ce livre sont d'ores et déjà entre les mains des informaticiens, qui s'inspirent ouvertement du cerveau<sup>323</sup>. En pratique, cependant, les machines ont encore bien du chemin à parcourir. Pour progresser, il leur faudra disposer d'un langage interne qui permette de recombinaison les concepts avec flexibilité ; d'algorithmes capables de raisonner avec des distributions de probabilités ; d'une fonction de curiosité ; de systèmes efficaces de gestion de l'attention et de la mémoire ; et peut-être, comme nous l'avons vu, d'une période de sommeil qui décuple l'apprentissage et la découverte. Des algorithmes de ce type commencent à apparaître mais, pour l'instant, ils restent à des années-lumière des performances d'un bébé de quelques mois. Le cerveau conserve donc la dragée haute.

## *Treize maximes pour l'épanouissement des enfants*

Plus nous l'étudions, plus le cerveau humain nous impressionne. Mais nous savons aussi que ses performances sont fragiles, car elles dépendent fortement de l'environnement dans lequel il se développe. Trop d'enfants ne réalisent pas pleinement leur potentiel d'apprentissage, parce que leur famille ou leur école ne leur fournissent pas les conditions idéales.

En France, les résultats des comparaisons internationales sont alarmants : depuis quinze ou vingt ans, les performances de nos enfants plongent<sup>324</sup>. L'enquête PISA, qui a lieu tous les quatre ans, montre qu'en lecture et en

compréhension de phrases, les élèves français occupent le milieu ou le bas du peloton des pays européens. En mathématiques, les résultats ont fortement baissé entre 2003 et 2015. Cette année-là, coup de tonnerre : on apprend que la France occupe la toute dernière place en Europe dans l'enquête TIMMS, qui évalue les acquis des élèves de 15 ans en maths et en sciences – c'est, de tous les pays, la plus forte baisse sur vingt ans. En 2016, même son de cloche de l'enquête PIRLS, qui évalue la lecture en CM1 : la France a plongé en quinze ans et occupe à présent le bas du classement européen, particulièrement lorsqu'il s'agit d'interpréter un texte un tant soit peu complexe. Pire, le pays dont la devise vante l'Égalité et la Fraternité présente les plus mauvais résultats dans le domaine de l'intégration scolaire : c'est en France que le niveau socio-économique des familles a le plus d'impact sur les résultats scolaires.

Face à ces mauvais résultats, les enseignants sont trop souvent pointés du doigt. En réalité, personne ne connaît les causes de cette chute récente : société, parents, école, sommeil, inattention, jeux vidéo, fratries réduites ? Quoi qu'il en soit, je suis persuadé que les avancées des sciences de l'apprentissage peuvent nous aider à inverser cette sombre tendance. Nous en savons long sur les conditions qui maximisent l'apprentissage et la mémoire. Parents, enseignants, chercheurs, nous devons tous apprendre à les mettre en œuvre dans la vie de tous les jours, à la maison comme en classe.

Tout au long de ce livre, j'ai présenté une série de résultats scientifiques qui convergent vers des idées simples et faciles à appliquer. Révisons-les ensemble.

- **Ne sous-estimons pas les enfants.** Dès la naissance, chaque enfant possède de riches noyaux de compétence. La connaissance des objets, le sens des nombres, le goût des langues, la compréhension des personnes... : autant de modules cérébraux qui sont déjà présents chez les tout-petits et que l'éducation recyclera plus tard pendant les cours de physique, de mathématiques, de langues ou de philosophie. Appuyons-nous sur les intuitions précoces des enfants : chaque mot, chaque symbole, même abstrait, doit venir se connecter à des connaissances préalables. C'est cette connexion qui leur donnera du sens.



- **Profitions des périodes sensibles.** Dans les toutes premières années de la vie, ce sont des milliards de synapses qui se font et se défont chaque jour. Ce bouillonnement rend le cerveau particulièrement réceptif à certains apprentissages, au premier chef celui des langues. Sachons profiter de cette période sensible pour l'exposer, dès que possible, à une seconde langue – en gardant en tête que la plasticité se prolonge au moins jusqu'à l'adolescence. Durant toute cette période, l'immersion dans une langue étrangère peut transformer le cerveau en quelques mois.
- **Enrichissons l'environnement.** Sur le plan de l'apprentissage, le cerveau de l'enfant demeure le plus puissant des superordinateurs. Respectons-le en lui fournissant, dès le plus jeune âge, des données à la hauteur : jeux de mots ou de construction, histoires, défis, casse-tête... Parlons-lui avec sérieux, répondons à ses questions, même les plus ardues, sans hésiter à employer un vocabulaire élaboré. Expliquons-lui le monde qui l'entoure. Donner à l'enfant un environnement enrichi, c'est maximiser la croissance de son cerveau et préserver, le plus longtemps possible, sa plasticité juvénile.
- **Ne croyons pas que les enfants sont tous différents.** L'idée que chacun d'entre nous possède son style d'apprentissage est un mythe. L'imagerie cérébrale montre que nous possédons tous des circuits et des règles d'apprentissage très semblables. Les aires cérébrales responsables de la lecture, du calcul mental ou des mathématiques sont, à quelques millimètres près, les mêmes chez tous – y compris les enfants aveugles. Tous les enfants rencontrent des difficultés similaires, et les mêmes méthodes pédagogiques permettent de les surmonter. Les différences, bien réelles, portent sur la vitesse d'apprentissage et les goûts de chacun : sachons nous y adapter et stimuler l'enfant avec des problèmes adaptés à son niveau et à ses envies, mais assurons-nous en premier lieu que tous les enfants acquièrent les fondamentaux dont ils auront besoin.
- **Faisons attention à l'attention.** L'attention est la porte d'entrée des apprentissages : pratiquement aucune information ne sera mémorisée si

elle n'a pas été, auparavant, amplifiée par l'attention et la prise de conscience. À l'école, sachons captiver l'attention des élèves et l'orienter vers le niveau approprié. Cela implique d'écarter avec soin toute source de distraction : les manuels trop illustrés, les classes trop décorées, les lettres ou les chiffres déformés ou figurés par des personnages ne font que distraire l'enfant et l'empêcher de se concentrer.

- **Rendons l'enfant actif, curieux, engagé, autonome.** Un élève passif n'apprend guère. Rendons-le plus actif. Sollicitons en permanence son intelligence afin que son esprit pétille de curiosité et génère en permanence des hypothèses nouvelles. Mais n'espérons pas qu'il découvre tout par lui-même : guidons-le par une pédagogie structurée.
- **Faisons de chaque jour d'école un plaisir.** Les circuits de la récompense sont des modulateurs essentiels de la plasticité cérébrale. Utilisons-les en faisant de chaque heure de cours un moment de plaisir, et en récompensant chaque effort. Aucun enfant n'est insensible aux récompenses matérielles – mais son cerveau social répond tout autant aux sourires et aux encouragements. Le sentiment d'être apprécié, mais aussi la conscience de progresser apportent leur propre récompense. À l'inverse, bannissons l'anxiété et le stress qui bloquent les apprentissages – surtout en mathématiques.
- **Encourageons les efforts.** L'école du plaisir n'est pas synonyme de « sans effort ». Au contraire, les plus intéressants des apprentissages – lire, faire des mathématiques, jouer d'un instrument... – exigent des mois ou des années d'apprentissage. Cessons de faire croire aux enfants que tout sera facile, ce qui leur laisse à entendre que, s'ils n'y arrivent pas, c'est qu'ils ne sont pas doués. Au contraire, expliquons-leur que tous les élèves doivent faire des efforts et que, lorsqu'ils le font, tous progressent. Adoptons une mentalité progressiste, pas une mentalité fixiste.
- **Aidons les élèves à approfondir leur pensée.** Le cerveau retient mieux les informations qu'il a traitées en profondeur. Creusons sans relâche pour atteindre le fond du sujet. Et souvenons-nous des paroles de Henry

Roediger : « Rendre les conditions d'apprentissage plus difficiles, ce qui oblige les étudiants à un surcroît d'engagement et d'effort cognitif, conduit souvent à une meilleure rétention. »

- **Fixons des objectifs clairs à l'apprentissage.** Les élèves apprennent mieux lorsqu'on leur énonce clairement le but de l'apprentissage, et qu'ils peuvent constater que tout ce qui leur est proposé converge vers ce but. Expliquons-leur clairement ce que l'on attend d'eux, et restons concentrés sur ce but.
- **Acceptons et corrigeons les erreurs.** Pour mettre à jour nos modèles mentaux, nos aires cérébrales doivent échanger des messages d'erreur. L'erreur est donc la condition même de l'apprentissage. Ne la sanctionnons pas, mais corrigeons-la rapidement, en donnant à l'enfant un retour sur erreur détaillé et décomplexé. D'après la synthèse de l'Education Endowment Fund, la qualité du retour que l'enseignant fournit à ses élèves est le levier le plus efficace de progrès scolaire <sup>325</sup> .
- **Révisions, encore et encore.** Apprendre ne suffit pas – encore faut-il consolider afin que le contenu de l'apprentissage devienne automatique, inconscient, réflexe. Seule l'automatisation libère le cortex préfrontal, qui devient ainsi disponible pour d'autres activités. La stratégie la plus efficace consiste à distribuer les apprentissages : un peu tous les jours. Par la suite, espacer progressivement les révisions permettra à l'information de s'imprimer définitivement en mémoire.
- **Laissons les enfants dormir.** Le sommeil est un ingrédient essentiel de notre algorithme d'apprentissage. Chaque période de sommeil, y compris la sieste, apporte un bénéfice supplémentaire. Assurons-nous donc que nos enfants dorment longuement et profondément. Pour tirer le maximum de bénéfices de cette période nocturne de travail inconscient, réviser une leçon ou relire un problème juste avant de s'endormir peut s'avérer une stratégie utile. Et comme la chronobiologie indique que les cycles de sommeil se décalent chez l'adolescent, ne les réveillons pas trop tôt !

Ce n'est qu'en apprenant à mieux nous connaître que nous parviendrons à tirer le meilleur parti des puissants algorithmes dont notre cerveau est équipé. Tous les enfants gagneraient sans doute à connaître les quatre piliers de l'apprentissage que sont l'attention, l'engagement actif, le retour sur erreur et la consolidation. Quatre slogans les résument efficacement : « Concentrez-vous totalement » ; « Participez en classe » ; « Faites des exercices » ; « Profitez de chaque jour et de chaque nuit<sup>326</sup> ». Ce sont des messages très simples que nous devrions tous connaître et mettre en œuvre dès le plus jeune âge.

## *Une alliance pour l'école de demain*

Saurons-nous mettre nos écoles en harmonie avec les acquis des neurosciences et des sciences cognitives ? Je suis persuadé qu'en unissant nos forces, nous pouvons y parvenir. L'école de demain devra rassembler les enseignants, les parents et les scientifiques autour d'une cause commune : ressusciter chez tous les enfants la curiosité et le bonheur d'apprendre.

Experts de la classe, les enseignants exercent leur métier dans des conditions souvent difficiles. Nous leur confions le plus précieux de nos vies et l'avenir de notre pays : ils méritent notre respect et un investissement massif. Or, trop souvent, ils ne reçoivent qu'une formation bien trop courte et presque totalement dépourvue d'introduction aux sciences de l'apprentissage. J'espère que ce livre contribuera à ce que les universités et les écoles supérieures du professorat et de l'éducation leur offrent un programme de formation à la hauteur de leur engagement au service de nos enfants.

Ils auront également compris que l'objectif de ce livre n'est pas de restreindre leur liberté pédagogique, mais de leur permettre de l'exercer au mieux, en toute conscience des conséquences de leurs choix. Pour être pleinement libre, il faut maîtriser tout l'éventail des stratégies disponibles et choisir parmi elles en pleine connaissance de leur impact. Les principes que j'ai énoncés tout au long de ce livre sont compatibles avec de multiples approches

pédagogiques, et tout reste à inventer pour les traduire en actes dans les classes. J'attends beaucoup de l'inventivité des enseignants, car elle me paraît essentielle à l'enthousiasme des enfants.

L'école du futur passera également par un meilleur accueil des parents, car ils sont les premiers acteurs de l'épanouissement de l'enfant. Leur action précède et prolonge celle de l'école. C'est en famille que l'enfant poursuit, par des travaux et des jeux, les apprentissages commencés en classe, et tire ainsi pleinement profit de chaque alternance de veille et de sommeil, d'apprentissage et de consolidation. L'école devrait consacrer plus de temps à former les parents, car c'est l'une des interventions les plus efficaces : ils peuvent être de précieux relais de l'enseignant, et de fins observateurs des difficultés de leurs enfants.

Les scientifiques, enfin, doivent s'engager auprès des enseignants. En France, la recherche en éducation est encore insuffisante. Les organismes de recherche devraient inciter les chercheurs à mener de grands programmes de recherche fondamentale dans tous les domaines des sciences de l'apprentissage, depuis les neurosciences et l'imagerie cérébrale jusqu'à la neuropsychologie des enfants « dys », la psychologie cognitive et la sociologie de l'éducation. Le passage à l'échelle, du laboratoire à la salle de classe, n'est pas aussi facile qu'il y paraît. Nous avons grand besoin d'expérimentations en grandeur réelle dans les écoles. Les outils des sciences cognitives peuvent aider à concevoir et à évaluer des projets éducatifs innovants. Tout comme la médecine s'appuie sur la biologie, l'Éducation nationale doit faire éclore un nouvel écosystème de recherche, plus systématique, plus rigoureux, qui associe enseignants et chercheurs dans la quête incessante d'une éducation plus efficace, fondée sur des données factuelles et bien éprouvées.

Enseignants, parents, scientifiques, ensemble pour faire progresser les sciences de l'éducation et leur mise en œuvre à l'école : la recette me paraît claire – mais l'urgence est réelle, car comme le souligne l'humoriste Francis Blanche :

« Face au monde qui bouge, il vaut mieux penser le changement que changer le pansement ! »

# Notes

---

## INTRODUCTION

- 1 . Voir les films *Miracle en Alabama* (1962) et *Marie Heurtin* (2014) ainsi que les livres suivants : Arnould, 1900 ; Keller, 1903.
- 2 . Apprentissage chez le nématode *C. elegans* : Bessa, Maciel & Rodrigues, 2013 ; Kano *et al.* , 2008 ; Rankin, 2004.
- 3 . Site de l'Education Endowment Fund (EEF) : <https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/> .

## PREMIÈRE PARTIE

### Qu'est-ce qu'apprendre ?

## CHAPITRE 1

### COMMENT UN RÉSEAU DE NEURONES APPREND

- 4 . Réseau de neurones LeNet : LeCun, Bottou, Bengio & Haffner, 1998.
- 5 . Visualisation de la hiérarchie d'unités du réseau GoogLeNet : Olah, Mordvintsev & Schubert, 2017 ; voir <https://doi.org/10.23915/distill.00007> .
- 6 . Séparation progressive des chiffres par un réseau de neurones profond : Guerguiev, Lillicrap & Richards, 2017.
- 7 . Apprentissage par récompense : Mnih *et al.* , 2015 ; Sutton & Barto, 1998.

- 8 . Réseau de neurones artificiel qui apprend à jouer aux jeux vidéo : Mnih *et al.* , 2015.
- 9 . Réseau de neurones artificiel qui apprend à jouer au go : Banino *et al.* , 2018 ; Silver *et al.* , 2016.
- 10 . Réseau de neurones qui lutte contre un adversaire : Goodfellow *et al.* , 2014.
- 11 . Réseaux de neurones à convolution : LeCun, Bengio & Hinton, 2015 ; LeCun *et al.* , 1998.
- 12 . L'évolution darwinienne considérée comme un algorithme : Dennett, 1996.

## CHAPITRE 2

### POURQUOI NOTRE CERVEAU APPREND MIEUX QUE LES MACHINES

- 13 . Les réseaux de neurones conventionnels correspondent aux opérations inconscientes du cerveau : Dehaene, Lau & Kouider, 2017.
- 14 . Les réseaux de neurones ont tendance à apprendre des régularités superficielles : Jo & Bengio, 2017.
- 15 . Génération d'images qui confondent les humains et les réseaux artificiels : Elsayed *et al.* , 2018.
- 16 . Réseau de neurones artificiel qui reconnaît les CAPTCHA : George *et al.* , 2017.
- 17 . Critique de la vitesse d'apprentissage des réseaux de neurones artificiels : Lake, Ullman, Tenenbaum & Gershman, 2017.
- 18 . Absence de systématisme dans les réseaux de neurones artificiels : Fodor & Pylyshyn, 1988 ; Fodor & McLaughlin, 1990.
- 19 . Existence d'un langage de la pensée : Amalric, Wang *et al.* , 2017 ; Fodor, 1975.
- 20 . Apprendre à compter, c'est sélectionner un programme mental : Piantadosi, Tenenbaum & Goldman, 2012 ; voir aussi Piantadosi, Tenenbaum & Goldman, 2016.
- 21 . Les représentations syntaxiques récursives, singularité de l'espèce humaine : Dehaene, Meyniel, Wacogne, Wang & Pallier, 2015 ; Everaert, Huybregts, Chomsky, Berwick & Bolhuis, 2015 ; Hauser, Chomsky & Fitch, 2002 ; Hauser & Watumull, 2017.
- 22 . Singularité de l'espèce humaine lors de l'écoute d'une séquence de sons : Wang, Uhrig, Jarraya & Dehaene, 2015.
- 23 . Apprentissage de règles géométriques, lent chez le singe, ultrarapide chez l'enfant : Jiang *et al.* , 2018.
- 24 . Le cerveau humain se comporte comme une machine de Turing : Sackur & Dehaene, 2009 ; Zylberberg, Dehaene, Roelfsema & Sigman, 2011.
- 25 . Apprentissage ultrarapide du sens des mots : Tenenbaum, Kemp, Griffiths & Goodman, 2011 ; Xu & Tenenbaum, 2007.
- 26 . L'apprentissage des mots dépend de l'attention partagée : Baldwin *et al.* , 1996.
- 27 . Connaissance des principaux articles et mots grammaticaux dès 12 mois : Cyr & Shi, 2013 ; Shi & Lepage, 2008.



- 28 . Règle d'exclusivité dans l'apprentissage des mots : Carey & Bartlett, 1978 ; Clark, 1988 ; Markman & Wachtel, 1988 ; Markman, Wasow & Hansen, 2003. Ce comportement a été élevé au rang de principe linguistique par Ellen Markman (*mutual exclusivity assumption* , 1989) et Eve Clark (*principle of contrast* ).
- 29 . Inhibition du principe d'exclusivité chez les bilingues : Byers-Heinlein & Werker, 2009.
- 30 . Rico, un chien qui apprend quelques centaines de mots : Kaminski, Call & Fischer, 2004.
- 31 . Modélisation d'un « scientifique artificiel » : Kemp & Tenenbaum, 2008.
- 32 . Découverte du principe de causalité : Goodman, Ullman & Tenenbaum, 2011 ; Tenenbaum *et al.* , 2011.
- 33 . Le cerveau vu comme un modèle génératif du monde extérieur : Lake, Salakhutdinov & Tenenbaum, 2015 ; Lake, Ullman, Tenenbaum & Gershman, 2016.
- 34 . La probabilité, logique de la science : Jaynes, 2003.
- 35 . Modèle bayésien des échanges d'information dans le cortex : Friston, 2005.

## DEUXIÈME PARTIE

### Comment notre cerveau apprend

#### CHAPITRE 3

##### LE SAVOIR INVISIBLE : LES ÉTONNANTES INTUITIONS DES BÉBÉS

- 36 . Concept d'objet chez le bébé : Baillargeon & DeVos, 1991 ; Kellman & Spelke, 1983.
- 37 . Apprentissage rapide de la chute des objets : Baillargeon, Needham & Devos, 1992 ; Hespos & Baillargeon, 2008.
- 38 . Concept de nombre chez le bébé : Izard, Dehaene-Lambertz & Dehaene, 2008 ; Izard, Sann, Spelke & Streri, 2009 ; Starkey & Cooper Jr, 1980 ; Starkey, Spelke & Gelman, 1990. On trouvera une revue détaillée de ces résultats dans la seconde édition de mon livre *La Bosse des maths. Quinze ans après* (Dehaene, 2010).
- 39 . Connaissance multimodale des nombres chez le bébé : Izard *et al.* , 2009.
- 40 . Addition et soustraction des petits nombres chez le bébé : Koechlin, Dehaene & Mehler, 1997 ; Wynn, 1992.
- 41 . Addition et soustraction des grands nombres chez le bébé : McCrink & Wynn, 2004.
- 42 . La précision du sens des nombres s'affine avec l'âge et l'éducation : Halberda & Feigenson, 2008 ; Piazza *et al.* , 2010 ; Piazza, Pica, Izard, Spelke & Dehaene, 2013.

- 43 . Sens des nombres chez le poussin : Rugani, Fontanari, Simoni, Regolin & Vallortigara, 2009 ; Rugani, Vallortigara, Priftis & Regolin, 2015.
- 44 . Neurones qui codent pour le nombre en l'absence d'entraînement : Ditz & Nieder, 2015 ; Viswanathan & Nieder, 2013.
- 45 . Existence d'un noyau de connaissances chez le bébé : Spelke, 2003.
- 46 . Raisonnement bayésien chez le bébé : Xu & Garcia, 2008.
- 47 . L'enfant agit comme un scientifique en herbe : Gopnik, 2000 ; Gopnik *et al.* , 2004.
- 48 . Le bébé comprend les probabilités, les urnes et les tirages aléatoires : Denison & Xu, 2010 ; Gweon, Tenenbaum & Schulz, 2010 ; Kushnir, Xu & Wellman, 2010.
- 49 . Le bébé comprend si c'est un humain ou une machine qui effectue un tirage : Ma & Xu, 2013.
- 50 . Compréhension des intentions chez le bébé : Gergely, Bekkering & Kiraly, 2002 ; Gergely & Csibra, 2003. Voir aussi Warneken & Tomasello, 2006.
- 51 . Les bébés de 10 mois infèrent les préférences des autres : Liu, Ullman, Tenenbaum & Spelke, 2017.
- 52 . Évaluation des personnes gentilles ou méchantes chez le bébé : Buon *et al.* , 2014.
- 53 . Distinction entre accident et intention chez le bébé : Behne, Carpenter, Call & Tomasello, 2005.
- 54 . Traitement des visages chez le fœtus *in utero* : Reid *et al.* , 2017.
- 55 . Reconnaissance des visages chez le bébé et développement de l'aire des visages : Adibpour, Dubois & Dehaene-Lambertz, 2018 ; Deen *et al.* , 2017 ; Livingstone *et al.* , 2017.
- 56 . Apprentissage des visages dans la première année de vie : Morton & Johnson, 1991.
- 57 . Un bébé de 4 jours préfère écouter sa langue maternelle : Mehler *et al.* , 1988.
- 58 . « L'enfant a tressailli... » : Évangile de Luc 1, 44.
- 59 . Voir mon livre *Le Code de la conscience* (Dehaene, 2014).
- 60 . Latéralisation du langage et de la voix chez le prématuré : Mahmoudzadeh *et al.* , 2013.
- 61 . Segmentation des mots chez le bébé : Hay, Pelucchi, Graf Estes & Saffran, 2011 ; Saffran, Aslin & Newport, 1996.
- 62 . Détection des violations grammaticales comme « je biberon » chez le jeune enfant : Bernal, Dehaene-Lambertz, Millotte & Christophe, 2010.
- 63 . Échec de l'apprentissage du langage chez l'animal, voir par exemple : Penn, Holyoak & Povinelli, 2008 ; Terrace, Petitto, Sanders & Bever, 1979 ; Yang, 2013.
- 64 . Émergence rapide du langage dans les communautés de personnes sourdes : Senghas, Kita & Özyürek, 2004.

## CHAPITRE 4

### NAISSANCE D'UN CERVEAU

- 65 . Imagerie cérébrale du langage chez le bébé : Dehaene-Lambertz *et al.* , 2006 ; Dehaene-Lambertz, Dehaene & Hertz-Pannier, 2002.
- 66 . Vision empiriste de l'apprentissage chez le bébé : voir par exemple Elman *et al.* , 1996 ; Quartz & Sejnowski, 1997.
- 67 . Évolution des aires corticales : Krubitzer, 2007.
- 68 . Hiérarchie de réponses au langage dans le cortex humain : Lerner, Honey, Silbert & Hasson, 2011 ; Pallier, Devauchelle & Dehaene, 2011.
- 69 . Présence des grands faisceaux de connexions dès la naissance : Dehaene-Lambertz & Spelke, 2015 ; Dubois *et al.* , 2015.
- 70 . Hypothèse du cortex initialement désorganisé et instruit par l'environnement : Quartz & Sejnowski, 1997.
- 71 . Le système nerveux périphérique est déjà remarquablement organisé dès deux mois de grossesse : Belle *et al.* , 2017.
- 72 . Séparation du cortex en aires de Brodmann : Amunts *et al.* , 2010 ; Amunts & Zilles, 2015 ; Brodmann, 1909.
- 73 . Expression précoce de gènes spécifiques à certaines aires corticales : Kwan *et al.* , 2012 ; Sun *et al.* , 2005.
- 74 . Origine précoce des asymétries cérébrales : Dubois *et al.* , 2009 ; Leroy *et al.* , 2015.
- 75 . Asymétrie cérébrale chez les droitiers et les gauchers : Sun *et al.* , 2012.
- 76 . Modèle d'auto-organisation des plissements du cortex : Lefèvre & Mangin, 2010.
- 77 . Cellules de grille chez le rat : Banino *et al.* , 2018 ; Brun *et al.* , 2008 ; Fyhn, Molden, Witter, Moser & Moser, 2004 ; Hafting, Fyhn, Molden, Moser & Moser, 2005.
- 78 . Hypothèse d'auto-organisation des cellules de grille : Kropff & Treves, 2008 ; Shipston-Sharman, Solanka & Nolan, 2016 ; Widloski & Fiete, 2014 ; Yoon *et al.* , 2013.
- 79 . Émergence rapide des cellules de lieu, de grille et de direction de la tête : Langston *et al.* , 2010 ; Wills, Cacucci, Burgess & O'Keefe, 2010.
- 80 . Cellules de grille dans le cerveau humain : Doeller, Barry & Burgess, 2010 ; Nau, Navarro Schröder, Bellmund & Doeller, 2018.
- 81 . Navigation dans l'espace chez un enfant aveugle : Landau, Gleitman & Spelke, 1981.
- 82 . Développement rapide de régions spécialisées pour les visages et les maisons : Deen *et al.* , 2017 ; Livingstone *et al.* , 2017.
- 83 . Réponse au nombre dans le cortex pariétal : Nieder & Dehaene, 2009.
- 84 . Modèle d'auto-organisation des neurones des nombres : Hannagan, Nieder, Viswanathan & Dehaene, 2017.
- 85 . Auto-organisation par apprentissage à partir d'un modèle interne du monde physique : Lake *et al.* , 2016.

## CHAPITRE 5

## LA PART DE L'ACQUIS

- 86 . Hypothèse synaptique de la plasticité cérébrale : Holtmaat & Caroni, 2016 ; Takeuchi, Duzskiewicz & Morris, 2014.
- 87 . Activation du circuit de la récompense par la musique : Salimpoor *et al.* , 2013.
- 88 . Potentiation à long terme des synapses : Bliss & Lomo, 1973 ; Lømo, 2018.
- 89 . Aplysie, hippocampe et plasticité synaptique : Pittenger & Kandel, 2003.
- 90 . Hippocampe et mémoire des lieux : Whitlock, Heynen, Shuler & Bear, 2006.
- 91 . Souvenir de la peur d'un son chez la souris : Kim & Cho, 2017.
- 92 . Rôle causal des modifications synaptiques : Takeuchi *et al.* , 2014.
- 93 . Nature de l'engramme, la base neuronale d'un souvenir : Josselyn, Köhler & Frankland, 2015 ; Poo *et al.* , 2016.
- 94 . Mémoire de travail et activité neuronale soutenue : Courtney, Ungerleider, Keil & Haxby, 1997 ; Ester, Sprague & Serences, 2015 ; Goldman-Rakic, 1995 ; Kerkoerle, Self & Roelfsema, 2017 ; Vogel & Machizawa, 2004.
- 95 . Mémoire de travail et modifications synaptiques rapides : Mongillo, Barak & Tsodyks, 2008.
- 96 . Rôle de l'hippocampe dans l'apprentissage rapide d'informations nouvelles : Genzel *et al.* , 2017 ; Lisman *et al.* , 2017 ; Schapiro, Turk-Browne, Norman & Botvinick, 2016 ; Shohamy & Turk-Browne, 2013.
- 97 . Déplacement d'un engramme de l'hippocampe vers le cortex : Kitamura *et al.* , 2017.
- 98 . Création d'un faux souvenir chez la souris : Ramirez *et al.* , 2013.
- 99 . Transformation d'un mauvais souvenir en bon : Ramirez *et al.* , 2015.
- 100 . Effacement d'un souvenir traumatique : Kim & Cho, 2017.
- 101 . Apprentissage d'outils ou de symboles chez le singe macaque : Iriki, 2005 ; Obayashi *et al.* , 2001 ; Srihasam, Mandeville, Morocz, Sullivan & Livingstone, 2012.
- 102 . Modifications synaptiques à distance : Fitzsimonds, Song & Poo, 1997.
- 103 . Modifications anatomiques dues à l'apprentissage de la musique : Gaser & Schlaug, 2003 ; Oechslin, Gschwind & James, 2018 ; Schlaug, Jancke, Huang, Staiger & Steinmetz, 1995.
- 104 . Modifications anatomiques dues à l'apprentissage de la lecture : Carreiras *et al.* , 2009 ; Thiebaut de Schotten, Cohen, Amemiya, Braga & Dehaene, 2014.
- 105 . Modifications anatomiques après avoir appris à jongler : Draganski *et al.* , 2004 ; Gerber *et al.* , 2014.
- 106 . Modifications cérébrales chez les chauffeurs de taxi londoniens : Maguire *et al.* , 2000, 2003.
- 107 . Mécanisme non synaptique d'apprentissage : Johansson, Jirenhed, Rasmussen, Zucca & Hesslow, 2014 ; Rasmussen, Jirenhed & Hesslow, 2008.
- 108 . Effets de l'exercice physique et de la nutrition sur le cerveau : Prado & Dewey, 2014 ; Voss, Vivar, Kramer & Van Praag, 2013.

- 109 . Déficits cognitifs chez les enfants avec carence en vitamine B1 (thiamine) : Fattal, Friedmann & Fattal-Valevski, 2011.
- 110 . Plasticité cérébrale chez un enfant sans hémisphère droit : Muckli, Naumer & Singer, 2009.
- 111 . Transformer le cortex auditif en cortex visuel : Sur, Garraghty & Roe, 1988 ; Sur & Rubenstein, 2005.
- 112 . Hypothèse du cortex désorganisé et instruit par l'environnement : Quartz & Sejnowski, 1997.
- 113 . Auto-organisation visuelle par le biais d'ondes qui se propagent sur la rétine : Goodman & Shatz, 1993 ; Shatz, 1996.
- 114 . Ajustement progressif de l'activité spontanée du cortex : Berkes, Orbán, Lengyel & Fiser, 2011 ; Orbán, Berkes, Fiser & Lengyel, 2016.
- 115 . Revue de la notion de période sensible pour l'apprentissage : Werker & Hensch, 2014.
- 116 . Développement des neurones corticaux dans le cerveau humain : Conel, 1939 ; Courchesne *et al.*, 2007.
- 117 . Surproduction et élimination de synapses au cours du développement : Rakic, Bourgeois, Eckenhoff, Zecevic & Goldman-Rakic, 1986.
- 118 . Vagues d'élimination synaptique dans le cerveau humain : Huttenlocher & Dabholkar, 1997.
- 119 . Myélinisation progressive du cortex : Dubois *et al.*, 2007, 2015 ; Flechsig, 1876.
- 120 . Accélération des réponses visuelles en quelques mois chez le bébé : Adibpour *et al.*, 2018 ; Dehaene-Lambertz & Spelke, 2015.
- 121 . Lenteur du traitement conscient chez le bébé : Kouider *et al.*, 2013.
- 122 . Période sensible pour le développement de la vision binoculaire : Epelbaum, Milleret, Buisseret & Duffer, 1993 ; Fawcett, Wang & Birch, 2005 ; Hensch, 2005.
- 123 . Perte de la discrimination des phonèmes non natifs vers 1 an : Dehaene-Lambertz & Spelke, 2015 ; Maye, Werker & Gerken, 2002 ; Pena, Werker & Dehaene-Lambertz, 2012 ; Werker & Tees, 1984.
- 124 . Récupération de la discrimination r/l chez les Japonais : McCandliss, Fiez, Protopapas, Conway & McClelland, 2002.
- 125 . L'anatomie du cortex auditif prédit la capacité d'apprendre des phonèmes étrangers : Golestani, Molko, Dehaene, Le Bihan & Pallier, 2007.
- 126 . Période sensible pour l'apprentissage d'une seconde langue : Flege, Munro & MacKay, 1995 ; Hartshorne, Tenenbaum & Pinker, 2018 ; Johnson & Newport, 1989 ; Weber-Fox & Neville, 1996.
- 127 . Déclin brutal de la vitesse d'apprentissage de la syntaxe d'une seconde langue vers 17 ans : Hartshorne *et al.*, 2018.
- 128 . Période sensible pour l'apprentissage du langage après la pose d'un implant cochléaire : Friedmann & Rusou, 2015.
- 129 . Mécanismes biologiques de l'ouverture et de la fermeture des périodes sensibles : Caroni, Donato & Muller, 2012 ; Friedmann & Rusou, 2015 ; Werker & Hensch, 2014.
- 130 . Restauration de la plasticité cérébrale : Krause *et al.*, 2017.

131 . Réorganisation des aires du langage chez les enfants adoptés : Pallier *et al.* , 2003. Des résultats similaires ont été trouvés dans le domaine de la reconnaissance des visages : adoptés avant 9 ans, les enfants coréens réapprennent le prototype du visage occidental et perdent l'avantage que l'on observe habituellement pour les visages de son propre pays d'origine (Sangrigoli, Pallier, Argenti, Ventureyra & Schonen, 2005).

132 . Représentation dormante de la première langue chez les enfants adoptés : Pierce, Klein, Chen, Delcenserie & Genesee, 2014.

133 . Connexions dormantes chez la chouette : Knudsen & Knudsen, 1990 ; Knudsen, Zheng & DeBello, 2000.

134 . Effet d'âge d'acquisition des mots : Ellis, Ralph & Matthew, 2000 ; Gerhand & Barry, 1999 ; Morrison & Ellis, 1995.

135 . Projet d'intervention auprès des orphelins de Bucarest : Almas *et al.* , 2012 ; Berens & Nelson, 2015 ; Nelson *et al.* , 2007 ; Sheridan, Fox, Zeanah, McLaughlin & Nelson, 2012 ; Windsor, Moraru, Nelson, Fox & Zeanah, 2013.

136 . Éthique du projet de Bucarest : Millum & Emanuel, 2007.

## CHAPITRE 6

### RECYCLEZ VOTRE CERVEAU

137 . Vladimir Nabokov, *Feu pâle* , traduction R. Girard et M.-E. Coindreau, Paris, Gallimard, 1955.

138 . Difficultés des analphabètes en reconnaissance d'images : Kolinsky *et al.* , 2011 ; Kolinsky, Morais, Content & Cary, 1987 ; Szwed, Ventura, Querido, Cohen & Dehaene, 2012.

139 . Difficultés des analphabètes à distinguer des images en miroir : Kolinsky *et al.* , 2011, 1987 ; Pegado, Nakamura *et al.* , 2014.

140 . Difficultés des analphabètes à prêter attention à une partie d'un visage : Ventura *et al.* , 2013.

141 . Difficultés des analphabètes à reconnaître et mémoriser des mots : Castro-Caldas, Petersson, Reis, Stone-Elander & Ingvar, 1998 ; Morais, 2017 ; Morais, Bertelson, Cary & Alegria, 1986 ; Morais & Kolinsky, 2005.

142 . Impact de l'éducation sur l'arithmétique élémentaire : Dehaene, Izard, Pica & Spelke, 2006 ; Dehaene, Izard, Spelke & Pica, 2008 ; Piazza *et al.* , 2013 ; Pica, Lemer, Izard & Dehaene, 2004.

143 . Apprentissage de la ligne numérique : Dehaene, 2003 ; Dehaene *et al.* , 2008 ; Siegler & Opfer, 2003.

144 . Théorie du recyclage neuronal : Dehaene, 2005, 2014 ; Dehaene & Cohen, 2007.

145 . Évolution par duplication de circuits cérébraux : Chakraborty & Jarvis, 2015 ; Fukuchi-Shimogori & Grove, 2001.

146 . Apprentissage confiné à un sous-espace d'activité neuronale : Galgali & Mante, 2018 ; Golub *et al.* , 2018 ; Sadtler *et al.* , 2014.

147 . Codage unidimensionnel (une ligne) dans le cortex pariétal : Chafee, 2013 ; Fitzgerald *et al.* , 2013.

- 148 . Rôle du cortex pariétal dans la comparaison du statut social : Chiao, 2010.
- 149 . Codage bidimensionnel (un plan) dans le cortex entorhinal : Yoon *et al.* , 2013.
- 150 . Codage d'un espace bidimensionnel arbitraire par des cellules de grille : Constantinescu, O'Reilly & Behrens, 2016.
- 151 . Codage des arbres syntaxiques dans la région de Broca : Musso *et al.* , 2003 ; Nelson *et al.* , 2017 ; Pallier *et al.* , 2011.
- 152 . *La Bosse des maths* : Dehaene, 2010.
- 153 . Neurones qui codent pour le nombre en l'absence d'entraînement : Viswanathan & Nieder, 2013.
- 154 . Effet de l'entraînement sur les neurones du cortex préfrontal : Viswanathan & Nieder, 2015.
- 155 . Apprentissage des chiffres arabes chez le singe : Diester & Nieder, 2007.
- 156 . Relation entre addition, soustraction et mouvements de l'attention visuelle : Knops, Thirion, Hubbard, Michel & Dehaene, 2009 ; Knops, Viarouge & Dehaene, 2009.
- 157 . IRM fonctionnelle des mathématiciens professionnels : Amalric & Dehaene, 2016, 2017.
- 158 . Imagerie des nombres chez le bébé : Izard *et al.* , 2008.
- 159 . IRM fonctionnelle des nombres et de *Sesame Street* chez l'enfant de 4 ans : Cantlon, Brannon, Carter & Pelphrey, 2006 ; Cantlon & Li, 2013.
- 160 . Mathématiciens aveugles : Amalric, Denghien & Dehaene, 2017.
- 161 . Recyclage du cortex occipital pour l'arithmétique et les mathématiques chez les aveugles : Amalric, Denghien *et al.* , 2017 ; Kanjlia, Lane, Feigenson & Bedny, 2016.
- 162 . Traitement du langage dans le cortex occipital des aveugles : Amedi, Raz, Pianka, Malach & Zohary, 2003 ; Bedny, Pascual-Leone, Dodell-Feder, Fedorenko & Saxe, 2011 ; Lane, Kanjlia, Omaki & Bedny, 2015 ; Sabbah *et al.* , 2016.
- 163 . Débat sur la flexibilité du cortex chez les aveugles : Bedny, 2017 ; Hannagan, Amedi, Cohen, Dehaene-Lambertz & Dehaene, 2015.
- 164 . Cartes rétinotopiques chez les aveugles : Bock *et al.* , 2015.
- 165 . Recyclage du cortex occipital chez les aveugles : Abboud, Maidenbaum, Dehaene & Amedi, 2015 ; Amedi *et al.* , 2003 ; Bedny *et al.* , 2011 ; Mahon, Anzellotti, Schwarzbach, Zampini & Caramazza, 2009 ; Reich, Szwed, Cohen & Amedi, 2011 ; Striem-Amit & Amedi, 2014 ; Strnad, Peelen, Bedny & Caramazza, 2013.
- 166 . La connectivité prédit la fonction dans le cortex visuel : Bouhali *et al.* , 2014 ; Hannagan *et al.* , 2015 ; Saygin *et al.* , 2012, 2013, 2016.
- 167 . Effet de distance quand nous comparons deux nombres : Dehaene, 2007b ; Dehaene, Dupoux & Mehler, 1990 ; Moyer & Landauer, 1967.
- 168 . Effet de distance quand nous décidons que deux nombres sont différents : Dehaene & Akhavein, 1995 ; Diester & Nieder, 2010.
- 169 . Effet de distance en addition et en soustraction : Groen & Parkman, 1972 ; Pinheiro-Chagas, Dotan, Piazza & Dehaene, 2017.
- 170 . Représentation mentale des prix : Dehaene & Marques, 2002 ; Marques & Dehaene, 2004.

- 171 . Représentation mentale de la parité : Dehaene, Bossini & Giraux, 1993. Nombres négatifs : Blair, Rosenberg-Lee, Tsang, Schwartz & Menon, 2012 ; Fischer, 2003 ; Gullick & Wolford, 2013. Fractions : Jacob & Nieder, 2009 ; Siegler, Thompson & Schneider, 2011.
- 172 . Langage de la pensée en mathématiques : Amalric, Wang *et al.* , 2017 ; Piantadosi, Tenenbaum & Goodman, 2012, 2016.
- 173 . Voir mon livre *Les Neurones de la lecture* (Dehaene, 2007a).
- 174 . Mécanismes cérébraux de la reconnaissance invariante des mots écrits : Dehaene *et al.* , 2001, 2004.
- 175 . Connexions entre l'aire de la forme visuelle des mots et les aires du langage : Bouhali *et al.* , 2014 ; Saygin *et al.* , 2016.
- 176 . Imagerie du cerveau de personnes analphabètes : Dehaene *et al.* , 2010 ; Dehaene, Cohen, Morais & Kolinsky, 2015 ; Pegado, Comerlato *et al.* , 2014.
- 177 . Spécialisation du cortex visuel précoce pour la lecture : Chang *et al.* , 2015 ; Dehaene *et al.* , 2010 ; Szwed, Qiao, Jobert, Dehaene & Cohen, 2014.
- 178 . Modifications anatomiques dues à l'apprentissage de la lecture : Carreiras *et al.* , 2009 ; Thiebaut de Schotten, Cohen, Amemiya, Braga & Dehaene, 2014.
- 179 . Visages chassés de l'hémisphère gauche : Dehaene *et al.* , 2010 ; Pegado, Comerlato *et al.* , 2014.
- 180 . Développement de la lecture et de la représentation des visages : Dehaene-Lambertz, Monzalvo & Dehaene, 2018 ; Dundas, Plaut & Behrmann, 2013 ; Li *et al.* , 2013 ; Monzalvo, Fluss, Billard, Dehaene & Dehaene-Lambertz, 2012.
- 181 . Activité insuffisante évoquée par les mots et les visages chez les enfants dyslexiques : Monzalvo *et al.* , 2012.
- 182 . Marqueur universel des difficultés de lecture : Rueckl *et al.* , 2015.
- 183 . Compétition entre mots et visages : knock-out ou blocus ? Dehaene-Lambertz *et al.* , 2018.
- 184 . (Ré)apprentissage de la lecture à l'âge adulte : Braga *et al.* , 2017 ; Cohen, Dehaene, McCormick, Durant & Zanker, 2016.
- 185 . Déplacement de la représentation corticale des mots chez les musiciens : Mongelli *et al.* , 2017.
- 186 . Réponse réduite aux visages chez les mathématiciens : Amalric & Dehaene, 2016.
- 187 . Nombreux effets à long terme d'une éducation précoce de qualité : voir le programme Abecedarian (Campbell *et al.* , 2012, 2014 ; Martin, Ramey & Ramey, 1990), le programme Perry de maternelle (Heckman, Moon, Pinto, Savelyev & Yavitz, 2010 ; Schweinhart, 1993) et le programme de la Jamaïque (Gertler *et al.* , 2014 ; Grantham-McGregor, Powell, Walker & Himes, 1991 ; Walker, Chang, Powell & Grantham-McGregor, 2005).
- 188 . Langage adressé à l'enfant et croissance du vocabulaire : Shneidman, Arroyo, Levine & Goldin-Meadow, 2013 ; Shneidman & Goldin-Meadow, 2012.
- 189 . Meilleure réponse au langage parlé chez les enfants à qui on lit des histoires : Hutton *et al.* , 2015, 2017. Voir aussi Romeo *et al.* , 2018.
- 190 . Avantages du bilinguisme précoce : Bialystok, Craik, Green & Gollan, 2009 ; Costa & Sebastián-Gallés, 2014 ; Li, Legault & Litcofsky, 2014.



191 . Bénéfices d'un environnement enrichi : Donato, Rompani & Caroni, 2013 ; Knudsen *et al.* , 2000 ; Van Praag, Kempermann & Gage, 2000 ; Voss *et al.* , 2013 ; Zhu *et al.* , 2014.

## TROISIÈME PARTIE

### Les quatre piliers de l'apprentissage

## CHAPITRE 7

### L'ATTENTION

192 . Attention chez la souris : Wang & Krauzlis, 2018.

193 . Attention dans un réseau de neurones artificiels : Bahdanau, Cho & Bengio, 2014 ; Cho, Courville & Bengio, 2015.

194 . Attention dans un réseau de neurones qui apprend à légender une image : Xu *et al.* , 2015.

195 . L'inattention réduit considérablement l'apprentissage : Ahissar & Hochstein, 1993.

196 . Un peu d'apprentissage reste possible sans attention ni conscience : Seitz, Lefebvre, Watanabe & Jolicœur, 2005 ; Watanabe, Nanez & Sasaki, 2001.

197 . Embrassement préfrontal et prise de conscience : Dehaene & Changeux, 2011 ; Van Vugt *et al.* , 2018.

198 . Acétylcholine, dopamine, plasticité cérébrale et modification des cartes corticales : Bao, Chan & Merzenich, 2001 ; Froemke, Merzenich & Schreiner, 2007 ; Kilgard & Merzenich, 1998.

199 . Équilibre entre inhibition et excitation, et réouverture de la plasticité cérébrale : Werker & Hensch, 2014.

200 . Jeux vidéo et mobilisation des circuits d'alerte et de récompense : Koepp *et al.* , 1998.

201 . Facultés cognitives entraînées par les jeux vidéo : Bavelier *et al.* , 2011 ; Cardoso-Leite & Bavelier, 2014 ; Green & Bavelier, 2003.

202 . Jeux vidéo pour l'apprentissage : voir les logiciels en accès libre La Course aux nombres ([www.thenumberrace.com](http://www.thenumberrace.com)) et L'Attrape-Nombre ([www.thenumbercatcher.com](http://www.thenumbercatcher.com)). Pour l'apprentissage de la lecture, voir le GraphoGame (<http://grapholearn.fr/>) et Élan pour la lecture (<http://elanpourlalecture.fr/>), prochainement renommé Kalulu.

203 . Orientation de l'attention spatiale : Posner, 1994.

204 . Effet amplificateur de l'attention : Çukur, Nishimoto, Huth & Gallant, 2013 ; Desimone & Duncan, 1995 ; Kastner & Ungerleider, 2000.

205 . Cécité causée par l'inattention : Mack & Rock, 1998 ; Simons & Chabris, 1999.

- 206 . L'inattention réduit massivement la vitesse d'apprentissage : Leong, Radulescu, Daniel, DeWoskin & Niv, 2017.
- 207 . Expérience de lecture locale ou globale : Yoncheva, Blau, Maurer & McCandliss, 2010.
- 208 . Organisation du contrôle exécutif dans le cortex préfrontal : D'Esposito & Grossman, 1996 ; Koechlin, Ody & Kouneiher, 2003 ; Rouault & Koechlin, 2018.
- 209 . Expansion du cortex préfrontal dans l'espèce humaine : Elston, 2003 ; Sakai *et al.* , 2011 ; Schoenemann, Sheehan & Glotzer, 2005 ; Smaers, Gómez-Robles, Parks & Sherwood, 2017.
- 210 . Organisation hiérarchique du cortex préfrontal et contrôle métacognitif : Fleming, Weil, Nagy, Dolan & Rees, 2010 ; Koechlin *et al.* , 2003 ; Rouault & Koechlin, 2018.
- 211 . Espace de travail neuronal global : Dehaene & Changeux, 2011 ; Dehaene, Changeux, Naccache, Sackur & Sergent, 2006 ; Dehaene, Kerszberg & Changeux, 1998 ; Dehaene & Naccache, 2001.
- 212 . Goulot d'étranglement central : Chun & Marois, 2002 ; Marti, King & Dehaene, 2015 ; Marti, Sigman & Dehaene, 2012 ; Sigman & Dehaene, 2008.
- 213 . Non-conscience du délai d'exécution d'une double tâche : Corallo, Sackur, Dehaene & Sigman, 2008 ; Marti *et al.* , 2012.
- 214 . Polémique sur la capacité de partager l'attention et d'effectuer deux tâches en parallèle : Tombu & Jolicoeur, 2004.
- 215 . Une classe trop décorée distrait l'enfant : Fisher, Godwin & Seltman, 2014.
- 216 . Erreur « A, non B » et développement du cortex préfrontal : Diamond & Doar, 1989 ; Diamond & Goldman-Rakic, 1989.
- 217 . Développement du contrôle exécutif dans la perception du nombre : Borst, Poirel, Pineau, Cassotti & Houdé, 2013 ; Piazza, DeFeo, Panzeri & Dehaene, 2018 ; Poirel *et al.* , 2012.
- 218 . Entraînement des capacités numériques et cortex préfrontal : Viswanathan & Nieder, 2015.
- 219 . Rôle du contrôle exécutif au cours du développement cognitif et émotionnel : Houdé *et al.* , 2000 ; Isingrini, Perrotin & Souchay, 2008 ; Posner & Rothbart, 1998 ; Sheese, Rothbart, Posner, White & Fraundorf, 2008 ; Siegler, 1989.
- 220 . Entraînement du contrôle exécutif et de la mémoire de travail : Diamond & Lee, 2011 ; Habibi, Damasio, Ilari, Elliott Sachs & Damasio, 2018 ; Jaeggi, Buschkuhl, Jonides & Shah, 2011 ; Klingberg, 2010 ; Moreno *et al.* , 2011 ; Olesen, Westerberg & Klingberg, 2004 ; Rueda, Rothbart, McCandliss, Saccomanno & Posner, 2005.
- 221 . Étude randomisée de l'efficacité de la pédagogie Montessori : Lillard & Else-Quest, 2006 ; Marshall, 2017.
- 222 . Effets de la pratique musicale sur le cerveau : Bermudez, Lerch, Evans & Zatorre, 2009 ; James *et al.* , 2014 ; Moreno *et al.* , 2011.
- 223 . Relation entre contrôle exécutif, cortex préfrontal et intelligence : Duncan, 2003, 2010, 2013.
- 224 . Effets de l'entraînement sur l'intelligence fluide : Au *et al.* , 2015.
- 225 . Effets de l'éducation et de l'adoption sur le QI : Duyme, Dumaret & Tomkiewicz, 1999 ; Ritchie & Tucker-Drob, 2017.
- 226 . Effets de l'entraînement cognitif sur la concentration, la lecture et l'arithmétique : Bergman-Nutley & Klingberg, 2014 ; Blair & Raver, 2014 ; Klingberg, 2010 ; Spencer-Smith & Klingberg,

2015.

227 . Corrélacion entre mémoire de travail et réussite ultérieure en mathématiques : Dumontheil & Klingberg, 2011 ; Gathercole, Pickering, Knight & Stegmann, 2004 ; Geary, 2011.

228 . Entraînement conjoint de la ligne numérique et de la mémoire de travail : Nemmi *et al.* , 2016.

229 . L'apprentissage des mots dépend de l'attention partagée : Baldwin *et al.* , 1996.

230 . Apprentissage du chinois avec une nounou, mais pas en vidéo : Kuhl, Tsao & Liu, 2003.

231 . Attention partagée et posture pédagogique : Csibra & Gergely, 2009 ; Egyed, Király & Gergely, 2013.

232 . Pointage vers un objet et mémoire de son identité : Yoon, Johnson & Csibra, 2008.

233 . « Enseignement » chez les suricates : Thornton & McAuliffe, 2006.

234 . Copie intelligente des intentions *versus* copie servile de l'action par un bébé de 14 mois : Gergely *et al.* , 2002.

235 . Conformisme social dans la perception : voir par exemple Bond & Smith, 1996.

## CHAPITRE 8

### L'ENGAGEMENT ACTIF

236 . Expérience classique des chatons passifs ou actifs : Held & Hein, 1963.

237 . Apprentissage des statistiques des syllabes : Hay *et al.* , 2011 ; Saffran *et al.* , 1996 ; travaux en cours de l'équipe de G. Dehaene-Lambertz sur l'apprentissage chez le nourrisson pendant le sommeil.

238 . Effet de la profondeur du traitement des mots sur la mémoire explicite : Craik & Tulving, 1975 ; Jacoby & Dallas, 1981.

239 . Mémoire des phrases : Auble & Franks, 1978 ; Auble, Franks & Soraci, 1979.

240 . « Rendre les conditions d'apprentissage plus difficiles... » : Zaromb, Karpicke & Roediger, 2010.

241 . Imagerie cérébrale de l'effet de profondeur de traitement des mots : Kapur *et al.* , 1994.

242 . L'activité des boucles entre le cortex préfrontal et l'hippocampe prédit la mémoire : Brewer, Zhao, Desmond, Glover & Gabrieli, 1998 ; Paller, McCarthy & Wood, 1988 ; Sederberg *et al.* , 2006 ; Sederberg, Kahana, Howard, Donner & Madsen, 2003 ; Wagner *et al.* , 1998.

243 . Mémoire de mots conscients ou inconscients : Dehaene *et al.* , 2001.

244 . Apprentissage actif de la physique : Kontra, Goldin-Meadow & Beilock, 2012 ; Kontra, Lyons, Fischer & Beilock, 2015.

245 . Comparaison du cours magistral et des pédagogies actives : Freeman *et al.* , 2014.

246 . Réfutation des pédagogies de la découverte : Hattie, Brodeur & St-Cyr, 2017 ; Kirschner, Sweller & Clark, 2006 ; Kirschner & Van Merriënboer, 2013 ; Mayer, 2004 ; Mottint, 2018.

247 . Comment additionner tous les nombres de 1 à 100 ? On apparie 1 avec 100, 2 avec 99, 3 avec 98..., de sorte que leur somme fasse toujours 101 ; et on voit que la somme totale comprend 50 fois

101, soit 5 050.

248 . Étude randomisée de l'efficacité de la pédagogie Montessori : Lillard & Else-Quest, 2006 ; Marshall, 2017.

249 . Mythes éducatifs : Kirschner & Van Merriënboer, 2013.

250 . Mythe des styles d'apprentissage : Pashler, McDaniel, Rohrer & Bjork, 2008.

251 . Les enfants curieux réussissent mieux : Shah, Weeks, Richards & Kaciroti, 2018.

252 . Neurones dopaminergiques sensibles à l'information : Bromberg-Martin & Hikosaka, 2009.

253 . Curiosité pour un lieu rempli d'objets nouveaux chez le rat : Bevins, 2001.

254 . Imagerie cérébrale de la curiosité : Gruber, Gelman & Ranganath, 2014. Voir aussi Kang *et al.* , 2009.

255 . Le rire comme une émotion épistémique propre à l'espèce humaine : Hurley, Dennett & Adams, 2011.

256 . Rire et apprentissage : Esseily, Rat-Fischer, Somogyi, O'Regan & Fagard, 2016.

257 . Revue des théories psychologiques de la curiosité : Loewenstein, 1994.

258 . La curiosité suit une courbe en U inversé : Kang *et al.* , 2009 ; Kidd, Piantadosi & Aslin, 2012, 2014 ; Loewenstein, 1994.

259 . Implémentation de la curiosité dans un robot : Gottlieb, Oudeyer, Lopes & Baranes, 2013 ; Kaplan & Oudeyer, 2007.

260 . Courbe en U inversé et effet Boucle d'Or chez le bébé de 8 mois : Kidd *et al.* , 2012, 2014.

261 . Métacognition chez le très jeune enfant : Dehaene *et al.* , 2017 ; Goupil, Romand-Monnier & Kouider, 2016 ; Lyons & Ghetti, 2011.

262 . Stéréotypes de genre ou de race en mathématiques : Spencer, Steele & Quinn, 1999 ; Steele & Aronson, 1995.

263 . Stress, anxiété, impuissance acquise et blocage de l'apprentissage : Caroni *et al.* , 2012 ; Donato *et al.* , 2013 ; Kim & Diamond, 2002 ; Noble, Norman & Farah, 2005.

264 . L'enseignement explicite peut tuer la curiosité : Bonawitz *et al.* , 2011.

## CHAPITRE 9

### LE RETOUR SUR ERREUR

265 . *Récoltes et semailles* : Grothendieck, 1986.

266 . La méta-analyse de John Hattie estime que le retour sur erreur a une taille d'effet de 0,73 écart-type, ce qui en fait l'un des facteurs les plus puissants dans le domaine de l'apprentissage (Hattie, 2008).

267 . Règle d'apprentissage par erreur de prédiction : Rescorla & Wagner, 1972.

268 . Pour une critique détaillée de la vision associationniste de l'apprentissage, voir : Balsam & Gallistel, 2009 ; Gallistel, 1990.

- 269 . Blocage de l'apprentissage associatif : Beckers, Miller, De Houwer & Urushihara, 2006 ; Fanselow, 1998 ; Waelti, Dickinson & Schultz, 2001.
- 270 . La surprise du bébé guide ses apprentissages et son exploration : Stahl & Feigenson, 2015.
- 271 . Échange de messages d'erreur dans le cerveau : Friston, 2005 ; Naatanen, Paavilainen, Rinne & Alho, 2007 ; Schultz, Dayan & Montague, 1997.
- 272 . La surprise reflète la violation d'une prédiction : Strauss *et al.* , 2015 ; Todorovic & de Lange, 2012.
- 273 . Hiérarchie de messages d'erreur locaux et globaux : Bekinschtein *et al.* , 2009 ; Strauss *et al.* , 2015 ; Uhrig, Dehaene & Jarraya, 2014 ; Wang *et al.* , 2015.
- 274 . Surprise liée à une image inattendue : Meyer & Olson, 2011.
- 275 . Surprise liée à une violation sémantique : Curran, Tucker, Kutas & Posner, 1993 ; Kutas & Federmeier, 2011 ; Kutas & Hillyard, 1980.
- 276 . Surprise liée à une violation de la grammaire : Friederici, 2002 ; Hahne & Friederici, 1999 ; mais voir aussi Steinhauer & Drury, 2012, pour une discussion critique.
- 277 . Prédications et erreurs de prédiction dans le circuit de la dopamine : Pessiglione, Seymour, Flandin, Dolan & Frith, 2006 ; Schultz *et al.* , 1997 ; Waelti *et al.* , 2001.
- 278 . L'importance de la qualité du retour sur erreur à l'école : Hattie, 2008.
- 279 . Apprentissage par essais et erreurs chez l'adulte et l'adolescent : Palminteri, Kilford, Coricelli & Blakemore, 2016.
- 280 . Daniel Pennac, *Le Monde* , 11 février 2017, [http://www.lemonde.fr/culture/article/2017/02/12/-daniel-pennac-j-ai-ete-d-abord-et-avant-tout-professeur\\_5078432\\_3246.html#WGr7tHhf7v-qGZm0g.99](http://www.lemonde.fr/culture/article/2017/02/12/-daniel-pennac-j-ai-ete-d-abord-et-avant-tout-professeur_5078432_3246.html#WGr7tHhf7v-qGZm0g.99) .
- 281 . Syndrome d'anxiété mathématique : Ashcraft, 2002 ; Lyons & Beilock, 2012 ; Maloney & Beilock, 2012 ; Young, Wu & Menon, 2012.
- 282 . Effet du conditionnement à la peur sur la plasticité synaptique : Caroni *et al.* , 2012 ; Donato *et al.* , 2013.
- 283 . État d'esprit fixiste ou progressiste : Claro, Paunesku & Dweck, 2016 ; Dweck, 2006 ; Rattan, Savani, Chugh & Dweck, 2015. L'importance de cet effet a toutefois été récemment contestée : Sisk, Burgoyne, Sun, Butler & Macnamara, 2018.
- 284 . Effet massif des tests sur l'apprentissage (*retrieval practise* ) : Carrier & Pashler, 1992 ; Karpicke & Roediger, 2008 ; Roediger & Karpicke, 2006 ; Szpunar, Khan & Schacter, 2013 ; Zaromb & Roediger, 2010. Pour une excellente revue détaillée des pratiques pédagogiques et de leur efficacité relative, voir : Dunlosky, Rawson, Marsh, Nathan & Willingham, 2013.
- 285 . Évaluer sa mémoire améliore l'efficacité des révisions : Robey, Dougherty & Buttaccio, 2017.
- 286 . Se tester améliore l'apprentissage d'une langue étrangère : Carrier & Pashler, 1992.
- 287 . Effet de l'espacement sur l'apprentissage : Cepeda *et al.* , 2009 ; Cepeda, Pashler, Vul, Wixted & Rohrer, 2006 ; Rohrer & Taylor, 2006 ; Schmidt & Bjork, 1992.
- 288 . Imagerie cérébrale de l'effet d'espacement sur l'apprentissage : Bradley *et al.* , 2015 ; Callan & Schweighofer, 2010.
- 289 . Importance de l'espacement croissant des révisions : Kang, Lindsey, Mozer & Pashler, 2014.

290 . Le mélange des exercices améliore les performances en mathématiques : Rohrer & Taylor, 2006, 2007.

291 . Même en cas de réponse correcte, le retour sur erreur améliore la mémoire : Butler, Karpicke & Roediger, 2008.

## CHAPITRE 10

### LA CONSOLIDATION

292 . Passage d'une lecture sérielle à une lecture parallèle au cours de l'apprentissage : Zoccolotti *et al.* , 2005.

293 . Imagerie cérébrale longitudinale du développement de la lecture : Dehaene-Lambertz *et al.* , 2018.

294 . Contribution du cortex pariétal chez les lecteurs experts seulement lorsque les mots sont dégradés : Cohen, Dehaene, Vinckier, Jobert & Montavont, 2008 ; Vinckier *et al.* , 2006.

295 . Reconnaissance visuelle des combinaisons fréquentes de lettres : Binder, Medler, Westbury, Liebenthal & Buchanan, 2006 ; Dehaene, Cohen, Sigman & Vinckier, 2005 ; Grainger & Whitney, 2004 ; Vinckier *et al.* , 2007.

296 . Adaptation du cortex visuel à la perception des lettres : Chang *et al.* , 2015 ; Dehaene *et al.* , 2010 ; Sigman *et al.* , 2005 ; Szwed *et al.* , 2011, 2014.

297 . Lecture inconsciente : Dehaene *et al.* , 2001, 2004.

298 . Automatisation de l'arithmétique : Ansari & Dhital, 2006 ; Rivera, Reiss, Eckert & Menon, 2005. L'hippocampe semble également contribuer fortement à la mémorisation des faits arithmétiques : Qin *et al.* , 2014.

299 . Le sommeil ralentit l'oubli : Jenkins & Dallenbach, 1924.

300 . Le sommeil paradoxal améliore l'apprentissage : Karni, Tanne, Rubenstein, Askenasy & Sagi, 1994.

301 . Sommeil et consolidation des apprentissages de la veille : Huber, Ghilardi, Massimini & Tononi, 2004 ; Stickgold, 2005 ; Walker, Brakefield, Hobson & Stickgold, 2003 ; Walker & Stickgold, 2004.

302 . Surexpression du gène zif-268 durant le sommeil après une journée dans un environnement enrichi : Ribeiro, Goyal, Mello & Pavlides, 1999.

303 . Le cerveau rejoue pendant la nuit les activités neuronales vécues dans la journée : Ji & Wilson, 2007 ; Louie & Wilson, 2001 ; Skaggs & McNaughton, 1996 ; Wilson & McNaughton, 1994.

304 . Déchiffrement du code neural au cours du sommeil : Chen & Wilson, 2017 ; Horikawa, Tamaki, Miyawaki & Kamitani, 2013.

305 . Théories de la fonction mnésique du sommeil : Diekelmann & Born, 2010.

306 . La réactivation au cours du sommeil permet la consolidation : Ramanathan, Gulati & Ganguly, 2015. Voir aussi Norimoto *et al.* , 2018, pour un effet direct du sommeil sur la plasticité synaptique.

- 307 . Réactivation corticale et hippocampique au cours du sommeil chez l'homme : Horikawa *et al.* , 2013 ; Jiang *et al.* , 2017 ; Peigneux *et al.* , 2004.
- 308 . Corrélation entre ondes lentes et amélioration des performances : Huber *et al.* , 2004.
- 309 . Imagerie cérébrale des effets du sommeil sur l'apprentissage moteur : Walker, Stickgold, Alsop, Gaab & Schlaug, 2005.
- 310 . Entraînement des ondes lentes du sommeil et amélioration de l'apprentissage : Marshall, Helgadottir, Mölle & Born, 2006 ; Ngo, Martinetz, Born & Mölle, 2013.
- 311 . Les odeurs peuvent biaiser la consolidation durant le sommeil : Rasch, Büchel, Gais & Born, 2007.
- 312 . Les sons présentés durant le sommeil peuvent biaiser la réactivation et faciliter la mémoire : Antony, Gobel, O'Hare, Reber & Paller, 2012 ; Bendor & Wilson, 2012 ; Rudoy, Voss, Westerberg & Paller, 2009.
- 313 . Une association nouvelle peut être apprise pendant le sommeil : Arzi *et al.* , 2012.
- 314 . Philippe Starck, *Boomerang* , France Inter, vendredi 8 juin 2018.
- 315 . Découverte de régularités pendant le sommeil : Wagner, Gais, Haider, Verleger & Born, 2004.
- 316 . Algorithmes d'apprentissage avec une alternance veille/sommeil : Hinton, Dayan, Frey & Neal, 1995 ; Hinton, Osindero & Teh, 2006.
- 317 . Hypothèse d'un sommeil plus profond et plus efficace dans l'espèce humaine : Samson & Nunn, 2015.
- 318 . Le sommeil de l'enfant est plus efficace que celui de l'adulte : Wilhelm *et al.* , 2013.
- 319 . Les bébés retiennent et généralisent mieux les mots après avoir dormi : Friedrich, Wilhelm, Born & Friederici, 2015 ; Seehagen, Konrad, Herbert & Schneider, 2015.
- 320 . Effet positif de la sieste en maternelle : Kurdziel, Duclos & Spencer, 2013.
- 321 . Sommeil et amélioration des troubles de l'attention : Avior *et al.* , 2004 ; Cortese *et al.* , 2013 ; Hiscock *et al.* , 2015 ; Prehn-Kristensen *et al.* , 2014.
- 322 . Amélioration des performances scolaires des adolescents lorsqu'on les laisse dormir : American Academy of Pediatrics, 2014.

## CONCLUSION

### RÉCONCILIER L'ÉDUCATION AVEC LES NEUROSCIENCES

- 323 . Intelligence artificielle inspirée par les neurosciences et les sciences cognitives : Hassabis, Kumaran, Summerfield & Botvinick, 2017 ; Lake *et al.* , 2016.
- 324 . Voir les enquêtes PISA (*Program for International Student Assessment* , <http://www.oecd.org/pisa-fr/> ), TIMMS (*Trends in International Mathematics and Science Study* ) et PIRLS (*Progress in International Reading Literacy Study* , <https://timssandpirls.bc.edu/> ).
- 325 . Site de l'Education Endowment Fund (EEF) : <https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/> .

326 . Ces idées sur la formation des élèves aux piliers de l'apprentissage sont notamment diffusées en France par l'association Énergie Jeunes.



# Bibliographie

---

- Abboud, S., Maidenbaum, S., Dehaene, S., Amedi, A. (2015), « A number-form area in the blind », *Nature Communications* , 6, 6026.
- Adibpour, P., Dubois, J., Dehaene-Lambertz, G. (2018), « Right but not left hemispheric discrimination of faces in infancy », *Nature Human Behaviour* , 2 (1), 67-79.
- Ahissar, M., Hochstein, S. (1993), « Attentional control of early perceptual learning », *Proc. Natl Acad. Sci. USA* , 90 (12), 5718-5722.
- Almas, A. N., Degnan, K. A., Radulescu, A., Nelson, C. A., Zeanah, C. H., Fox, N. A. (2012), « Effects of early intervention and the moderating effects of brain activity on institutionalized children's social skills at age 8 », *Proc. Natl Acad. Sci. USA* , 109 (suppl. 2), 17228-17231.
- Amalric, M., Dehaene, S. (2016), « Origins of the brain networks for advanced mathematics in expert mathematicians », *Proceedings of the National Academy of Sciences* , 201603205.
- Amalric, M., Dehaene, S. (2017), « Cortical circuits for mathematical knowledge : Evidence for a major subdivision within the brain's semantic networks », *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* , 373 (1740).

- Amalric, M., Denghien, I., Dehaene, S. (2017), « On the role of visual experience in mathematical development : Evidence from blind mathematicians », *Developmental Cognitive Neuroscience* .
- Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., Dehaene, S. (2017), « The language of geometry : Fast comprehension of geometrical primitives and rules in human adults and preschoolers », *PLoS Computational Biology* , 13 (1), e1005273.
- Amedi, A., Raz, N., Pianka, P., Malach, R., Zohary, E. (2003), « Early “visual” cortex activation correlates with superior verbal memory performance in the blind », *Nature Neuroscience* , 6 (7), 758-766.
- American Academy of Pediatrics (2014), « School start times for adolescents », *Pediatrics* , 134 (3), 642-649.
- Amunts, K., Lenzen, M., Friederici, A. D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., Zilles, K. (2010), « Broca’s region : Novel organizational principles and multiple receptor mapping », *PLoS Biology* , 8 (9).
- Amunts, K., Zilles, K. (2015), « Architectonic mapping of the human brain beyond Brodmann », *Neuron* , 88 (6), 1086-1107.
- Ansari, D., Dhital, B. (2006), « Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing : An event-related functional magnetic resonance imaging study », *J. Cogn. Neurosci.* , 18 (11), 1820-1828.
- Antony, J. W., Gobel, E. W., O’Hare, J. K., Reber, P. J., Paller, K. A. (2012), « Cued memory reactivation during sleep influences skill learning », *Nature Neuroscience* , 15 (8), 1114-1116.
- Arnould, L. (1900), *Une âme en prison. Histoire de l’éducation d’une aveugle-sourde-muette de naissance* , Paris, Oudin.
- Arzi, A., Shedlesky, L., Ben-Shaul, M., Nasser, K., Oksenberg, A., Hairston, I. S., Sobel, N. (2012), « Humans can learn new information during sleep », *Nature Neuroscience* , 15 (10), 1460-1465.

- Ashcraft, M. H. (2002), « Math anxiety : Personal, educational, and cognitive consequences », *Current Directions in Psychological Science* , 11, 181-185.
- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuhl, M., Jaeggi, S. M. (2015), « Improving fluid intelligence with training on working memory : A meta-analysis », *Psychonomic Bulletin and Review* , 22 (2), 366-377.
- Auble, P. M., Franks, J. J. (1978), « The effects of effort toward comprehension on recall », *Memory and Cognition* , 6 (1), 20-25.
- Auble, P. M., Franks, J. J., Soraci, S. A., (1979), « Effort toward comprehension : Elaboration or “aha” ? », *Memory and Cognition* , 7 (6), 426-434.
- Avior, G., Fishman, G., Leor, A., Sivan, Y., Kaysar, N., Derowe, A. (2004), « The effect of tonsillectomy and adenoidectomy on inattention and impulsivity as measured by the Test of Variables of Attention (TOVA) in children with obstructive sleep apnea syndrome », *Otolaryngology* , 131 (4), 367-371.
- Bahdanau, D., Cho, K., Bengio, Y. (2014), « Neural machine translation by jointly learning to align and translate », <https://arxiv.org/abs/1409.0473> .
- Baillargeon, R., DeVos, J. (1991), « Object permanence in young infants : Further evidence », *Child Development* , 62 (6), 1227-1246.
- Baillargeon, R., Needham, A., Devos, J. (1992), « The development of young infants' intuitions about support », *Early Development and Parenting* , 1 (2), 69-78.
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., Tidball, G. (1996), « Infants' reliance on a social criterion for establishing word-object relations », *Child Development* , 67 (6), 3135-3153.
- Balsam, P. D., Gallistel, C. R. (2009), « Temporal maps and informativeness in associative learning », *Trends in Neurosciences* , 32 (2), 73-78.
- Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., ..., Kumaran, D. (2018), « Vector-based navigation using gridlike representations in artificial agents », *Nature* , 557 (7705), 429-433.

- Bao, S., Chan, V. T., Merzenich, M. M. (2001), « Cortical remodelling induced by activity of ventral tegmental dopamine neurons », *Nature* , 412 (6842), 79-83.
- Bavelier, D., Green, C. S., Han, D. H., Renshaw, P. F., Merzenich, M. M., Gentile, D. A. (2011), « Brains on video games », *Nature Reviews. Neuroscience* , 12 (12), 763-768.
- Beckers, T., Miller, R. R., De Houwer, J., Urushihara, K. (2006), « Reasoning rats : Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference », *Journal of Experimental Psychology : General* , 135 (1), 92.
- Bedny, M. (2017), « Evidence from blindness for a cognitively pluripotent cortex », *Trends Cogn. Sci.* , 21 (9), 637-648.
- Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E., Saxe, R. (2011), « Language processing in the occipital cortex of congenitally blind adults », *Proc. Natl Acad. Sci. USA* , 108 (11), 4429-4434.
- Behne, T., Carpenter, M., Call, J., Tomasello, M. (2005), « Unwilling versus unable : infants' understanding of intentional action », *Developmental Psychology* , 41 (2), 328.
- Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., Naccache, L. (2009), « Neural signature of the conscious processing of auditory regularities », *Proc. Natl Acad. Sci. USA* , 106 (5), 1672-1677.
- Belle, M., Godefroy, D., Couly, G., Malone, S. A., Collier, F., Giacobini, P., Chédotal, A. (2017), « Tridimensional visualization and analysis of early human development », *Cell* , 169 (1), 161-173.e12.
- Bendor, D., Wilson, M. A. (2012), « Biasing the content of hippocampal replay during sleep », *Nature Neuroscience* , 15 (10), 1439-1444.
- Berens, A. E., Nelson, C. A. (2015), « The science of early adversity : Is there a role for large institutions in the care of vulnerable children ? », *The Lancet* , 386 (9991), 388-398.

- Bergman-Nutley, S., Klingberg, T. (2014), « Effect of working memory training on working memory, arithmetic and following instructions », *Psychological Research* , 78 (6), 869-877.
- Berkes, P., Orban, G., Lengyel, M., Fiser, J. (2011), « Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment », *Science* , 331 (6013), 83-87.
- Bermudez, P., Lerch, J. P., Evans, A. C., Zatorre, R. J. (2009), « Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry », *Cereb. Cortex*, 19 (7), 1583-1596.
- Bernal, S., Dehaene-Lambertz, G., Millotte, S., Christophe, A. (2010), « Two-year-olds compute syntactic structure on-line », *Dev. Sci.* , 13 (1), 69-76.
- Bessa, C., Maciel, P., Rodrigues, A. J. (2013), « Using *C. elegans* to decipher the cellular and molecular mechanisms underlying neurodevelopmental disorders », *Molecular Neurobiology* , 48 (3), 465-489.
- Bevins, R. A. (2001), « Novelty seeking and reward : Implications for the study of high-risk behaviors », *Current Directions in Psychological Science* , 10 (6), 189-193.
- Bialystok, E., Craik, F. I. M., Green, D. W., Gollan, T. H. (2009), « Bilingual minds », *Psychological Science in the Public Interest* , 10 (3), 89-129.
- Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., Buchanan, L. (2006), « Tuning of the human left fusiform gyrus to sublexical orthographic structure », *NeuroImage* , 33 (2), 739-748.
- Blair, C., Raver, C. C. (2014), « Closing the achievement gap through modification of neurocognitive and neuroendocrine function : Results from a Cluster randomized controlled trial of an innovative approach to the education of children in kindergarten », *PLoS ONE* , 9 (11), e112393.
- Blair, K. P., Rosenberg-Lee, M., Tsang, J. M., Schwartz, D. L., Menon, V. (2012), « Beyond natural numbers : Negative number representation in parietal cortex », *Frontiers in Human Neuroscience* , 6, 7.

- Bliss, T. V., Lomo, T. (1973), « Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path », *The Journal of Physiology* , 232 (2), 331-356.
- Bock, A. S., Binda, P., Benson, N. C., Bridge, H., Watkins, K. E., Fine, I. (2015), « Resting-state retinotopic organization in the absence of retinal input and visual experience », *J. Neurosci.* , 35 (36), 12366-12382.
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., Schulz, L. (2011), « The double-edged sword of pedagogy : Instruction limits spontaneous exploration and discovery », *Cognition* , 120 (3), 322-330.
- Bond, R., Smith, P. B. (1996), « Culture and conformity : A meta-analysis of studies using Asch's (1952b, 1956) line judgment task », *Psychological Bulletin* , 119 (1), 111.
- Borst, G., Poirel, N., Pineau, A., Cassotti, M., Houdé, O. (2013), « Inhibitory control efficiency in a Piaget-like class-inclusion task in school-age children and adults : A developmental negative priming study », *Developmental Psychology* , 49 (7), 1366-1374.
- Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J.-F., Dehaene, S., Cohen, L. (2014), « Anatomical connections of the visual word form area », *J. Neurosci.* , 34 (46), 15402-15414.
- Bradley, M. M., Costa, V. D., Ferrari, V., Codispoti, M., Fitzsimmons, J. R., Lang, P. J. (2015), « Imaging distributed and massed repetitions of natural scenes : Spontaneous retrieval and maintenance », *Human Brain Mapping* , 36 (4), 1381-1392.
- Braga, L. W., Amemiya, E., Tauil, A., Sugueida, D., Lacerda, C., Klein, E., ..., Dehaene, S. (2017), « Tracking adult literacy acquisition with functional MRI : A single-case study », *Mind, Brain and Education* .
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., Gabrieli, J. D. (1998), « Making memories : Brain activity that predicts how well visual experience will be remembered », *Science* , 281 (5380), 1185-1187.

- Brodman, K. (1909), *Localisation in the Cerebral Cortex* , Springer.
- Bromberg-Martin, E. S., Hikosaka, O. (2009), « Midbrain dopamine neurons signal preference for advance information about upcoming rewards », *Neuron* , 63 (1), 119-126.
- Brun, V. H., Leutgeb, S., Wu, H. Q., Schwarcz, R., Witter, M. P., Moser, E. I., Moser, M. B. (2008), « Impaired spatial representation in CA1 after lesion of direct input from entorhinal cortex », *Neuron* , 57 (2), 290-302.
- Buon, M., Jacob, P., Margules, S., Brunet, I., Dutat, M., Cabrol, D., Dupoux, E. (2014), « Friend or foe ? Early social evaluation of human interactions », *PLoS ONE* , 9 (2), e88612.
- Butler, A. C., Karpicke, J. D., Roediger, H. L. (2008), « Correcting a metacognitive error : Feedback increases retention of low-confidence correct responses », *J. Exp. Psychol. Learn. Mem. Cogn.* , 34 (4), 918-928.
- Byers-Heinlein, K., Werker, J. F. (2009), « Monolingual, bilingual, trilingual : Infants' language experience influences the development of a word-learning heuristic », *Dev. Sci.* , 12 (5), 815-823.
- Callan, D. E., Schweighofer, N. (2010), « Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory », *Human Brain Mapping* , 31 (4), 645-659.
- Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., ..., Ramey, C. T. (2012), « Adult outcomes as a function of an early childhood educational program : An Abecedarian Project follow-up », *Developmental Psychology* , 48 (4), 1033-1043.
- Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., Pan, Y. (2014), « Early childhood investments substantially boost adult health », *Science* , 343 (6178), 1478-1485.
- Cantlon, J. F., Brannon, E. M., Carter, E. J., Pelphrey, K. A. (2006), « Functional imaging of numerical processing in adults and 4-y-old children », *PLoS Biology* , 4 (5), e125.

- Cantlon, J. F., Li, R. (2013), « Neural activity during natural viewing of *Sesame Street* statistically predicts test scores in early childhood », *PLoS Biology* , 11 (1), e1001462.
- Cardoso-Leite, P., Bavelier, D. (2014), « Video game play, attention, and learning : How to shape the development of attention and influence learning ? », *Current Opinion in Neurology* , 27 (2), 185-191.
- Carey, S., Bartlett, E. (1978), « Acquiring a single new word », *Proceedings of the Stanford Child Language Conference* , 15, 17-29.
- Caroni, P., Donato, F., Muller, D. (2012), « Structural plasticity upon learning : Regulation and functions », *Nature Reviews. Neuroscience* , 13 (7), 478-490.
- Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., Price, C. J. (2009), « An anatomical signature for literacy », *Nature* , 461 (7266), 983-986.
- Carrier, M., Pashler, H. (1992), « The influence of retrieval on retention », *Memory and Cognition* , 20 (6), 633-642.
- Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., Ingvar, M. (1998), « The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain », *Brain* , 121 (Pt 6), 1053-1063.
- Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., Pashler, H. (2009), « Optimizing distributed practice : Theoretical analysis and practical implications », *Experimental Psychology* , 56 (4), 236-246.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., Rohrer, D. (2006), « Distributed practice in verbal recall tasks : A review and quantitative synthesis », *Psychological Bulletin* , 132 (3), 354-380.
- Chafee, M. V. (2013), « A scalar neural code for categories in parietal cortex : Representing cognitive variables as “more” or “less” », *Neuron* , 77 (1), 7-9.
- Chakraborty, M., Jarvis, E. D. (2015), « Brain evolution by brain pathway duplication », *Philos. Trans. R. Soc. B* , 370 (1684), 20150056.



- Chang, C. H. C., Pallier, C., Wu, D. H., Nakamura, K., Jobert, A., Kuo, W. J., Dehaene, S. (2015), « Adaptation of the human visual system to the statistics of letters and line configurations », *NeuroImage* , 120, 428-440.
- Chen, Z., Wilson, M. A. (2017), « Deciphering neural codes of memory during sleep », *Trends in Neurosciences* , 40 (5), 260-275.
- Chiao, J. Y. (2010), « Neural basis of social status hierarchy across species », *Current Opinion in Neurobiology* , 20 (6), 803-809.
- Cho, K., Courville, A., Bengio, Y. (2015), « Describing multimedia content using attention-based encoder-decoder networks », *IEEE Transactions on Multimedia* , 17 (11), 1875-1886.
- Chun, M. M., Marois, R. (2002), « The dark side of visual attention », *Current Opinion in Neurobiology* , 12 (2), 184-189.
- Clark, E. V. (1988), « On the logic of contrast », *Journal of Child Language* , 15 (2), 317-335.
- Claro, S., Paunesku, D., Dweck, C. S. (2016), « Growth mindset tempers the effects of poverty on academic achievement », *Proc. Natl Acad. Sci USA* , 113 (31), 8664-8668.
- Cohen, L., Dehaene, S., McCormick, S., Durant, S., Zanker, J. M. (2016), « Brain mechanisms of recovery from pure alexia : A single case study with multiple longitudinal scans », *Neuropsychologia* , 91, 36-49.
- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., Montavont, A. (2008), « Reading normal and degraded words : Contribution of the dorsal and ventral visual pathways » *NeuroImage* , 40 (1), 353-366.
- Conel, J. L. (1939), *The Postnatal Development of the Human Cerebral Cortex* , Cambridge (Mass.), Harvard University Press, vol. 1-8.
- Constantinescu, A. O., O'Reilly, J. X., Behrens, T. E. J. (2016), « Organizing conceptual knowledge in humans with a gridlike code », *Science* , 352 (6292), 1464-1468.

- Corallo, G., Sackur, J., Dehaene, S., Sigman, M. (2008), « Limits on introspection : Distorted subjective time during the dual-task bottleneck », *Psychological Science*, 19 (11), 1110-1117.
- Cortese, S., Brown, T. E., Corkum, P., Gruber, R., O'Brien, L. M., Stein, M., ..., Owens, J. (2013), « Assessment and management of sleep problems in youths with attention-deficit/hyperactivity disorder », *Journal of the American Academy of Child and Adolescent Psychiatry*, 52 (8), 784-796.
- Costa, A., Sebastián-Gallés, N. (2014), « How does the bilingual experience sculpt the brain ? », *Nature Reviews. Neuroscience*, 15 (5), 336-345.
- Courchesne, E., Pierce, K., Schumann, C. M., Redcay, E., Buckwalter, J. A., Kennedy, D. P., Morgan, J. (2007), « Mapping early brain development in autism », *Neuron*, 56 (2), 399-413.
- Courtney, S. M., Ungerleider, L. G., Keil, K., Haxby, J. V. (1997), « Transient and sustained activity in a distributed neural system for human working memory », *Nature*, 386 (6625), 608-611.
- Craik, F. I., Tulving, E. (1975), « Depth of processing and the retention of words in episodic memory », *Journal of Experimental Psychology : General*, 104 (3), 268.
- Csibra, G., Gergely, G. (2009), « Natural pedagogy », *Trends Cogn. Sci.*, 13 (4), 148-153.
- Çukur, T., Nishimoto, S., Huth, A. G., Gallant, J. L. (2013), « Attention during natural vision warps semantic representation across the human brain », *Nature Neuroscience*, 16 (6), 763-770.
- Curran, T., Tucker, D. M., Kutas, M., Posner, M. I. (1993), « Topography of the N400 : Brain electrical activity reflecting semantic expectancy », *Electroencephalography and Clinical Neurophysiology*, 88, 188-209.
- Cyr, M., Shi, R. (2013), « Development of abstract grammatical categorization in infants », *Child Development*, 84 (2), 617-629.

- Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., ..., Saxe, R. (2017), « Organization of high-level visual cortex in human infants », *Nature Communications* , 8, 13995.
- Dehaene, S. (2003), « The neural basis of the Weber-Fechner law : A logarithmic mental number line », *Trends Cogn. Sci.* , 7 (4), 145-147.
- Dehaene, S. (2005), « Evolution of human cortical circuits for reading and arithmetic : The “neuronal recycling” hypothesis », in S. Dehaene, J. R. Duhamel, M. Hauser, G. Rizzolatti (éd.), *From Monkey Brain to Human Brain* , Cambridge (Mass.), MIT Press, 133-157.
- Dehaene, S. (2007a), *Les Neurones de la lecture* , Paris, Odile Jacob (trad. angl. : *Reading in the brain*, New York, Penguin Viking, 2009).
- Dehaene, S. (2007b), « Symbols and quantities in parietal cortex : Elements of a mathematical theory of number representation and manipulation », in P. Haggard, Y. Rossetti (éd.), *Attention and Performance XXII. Sensorimotor foundations of higher cognition* , Cambridge (Mass), Harvard University Press, 527-574.
- Dehaene, S. (2010), *La Bosse des maths. Quinze ans après* , Paris, Odile Jacob.
- Dehaene, S. (2014), *Le Code de la conscience* , Paris, Odile Jacob.
- Dehaene, S., Akhavein, R. (1995), « Attention, automaticity and levels of representation in number processing », *J. Exp. Psychol. Learn. Mem. Cogn.* , 21, 314-326.
- Dehaene, S., Bossini, S., Giraux, P. (1993), « The mental representation of parity and numerical magnitude », *Journal of Experimental Psychology : General* , 122, 371-396.
- Dehaene, S., Changeux, J.-P. (2011), « Experimental and theoretical approaches to conscious processing », *Neuron* , 70 (2), 200-227.
- Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., Sergent, C. (2006), « Conscious, preconscious, and subliminal processing : A testable taxonomy », *Trends Cogn. Sci.* , 10 (5), 204-211.

- Dehaene, S., Cohen, L. (2007), « Cultural recycling of cortical maps », *Neuron* , 56 (2), 384-398.
- Dehaene, S., Cohen, L., Morais, J., Kolinsky, R. (2015), « Illiterate to literate : Behavioural and cerebral changes induced by reading acquisition », *Nature Reviews. Neuroscience* , 16 (4), 234-244.
- Dehaene, S., Cohen, L., Sigman, M., Vinckier, F. (2005), « The neural code for written words : A proposal », *Trends Cogn. Sci.* , 9 (7), 335-341.
- Dehaene, S., Dupoux, E., Mehler, J. (1990), « Is numerical comparison digital : Analogical and symbolic effects in two-digit number comparison », *J. Exp. Psychol. Hum. Percep. Perf.* , 16, 626-641.
- Dehaene, S., Izard, V., Pica, P., Spelke, E. (2006), « Core knowledge of geometry in an Amazonian indigene group », *Science* , 311, 381-384.
- Dehaene, S., Izard, V., Spelke, E., Pica, P. (2008), « Log or linear ? Distinct intuitions of the number scale in Western and Amazonian indigene cultures », *Science* , 320 (5880), 1217-1220.
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., Cohen, L. (2004), « Letter binding and invariant recognition of masked words : Behavioral and neuroimaging evidence », *Psychological Science*, 15 (5), 307-313.
- Dehaene, S., Kerszberg, M., Changeux, J.-P. (1998), « A neuronal model of a global workspace in effortful cognitive tasks », *Proc. Natl Acad. Sci. USA* , 95 (24), 14529-14534.
- Dehaene, S., Lau, H., Kouider, S. (2017), « What is consciousness, and could machines have it ? », *Science* , 358 (6362), 486-492.
- Dehaene, S., Marques, J. F. (2002), « Cognitive Euroscience : Scalar variability in price estimation and the cognitive consequences of switching to the Euro », *Quarterly Journal of Experimental Psychology* , 55 (3), 705-731.
- Dehaene, S., Meyniel, F., Wacogne, C., Wang, L., Pallier, C. (2015), « The neural representation of sequences : From transition probabilities to

- algebraic patterns and linguistic trees », *Neuron* , 88 (1), 2-19.
- Dehaene, S., Naccache, L. (2001), « Towards a cognitive neuroscience of consciousness : Basic evidence and a workspace framework », *Cognition* , 79, 1-37.
- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., Rivière, D. (2001), « Cerebral mechanisms of word masking and unconscious repetition priming », *Nature Neuroscience* , 4 (7), 752-758.
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ..., Cohen, L. (2010), « How learning to read changes the cortical networks for vision and language », *Science* , 330 (6009), 1359-1364.
- Dehaene-Lambertz, G., Dehaene, S., Hertz-Pannier, L. (2002), « Functional neuroimaging of speech perception in infants », *Science* , 298 (5600), 2013-2015.
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M., Dehaene, S. (2006), « Functional organization of perisylvian activation during presentation of sentences in preverbal infants », *Proc. Natl Acad. Sci. USA* , 103 (38), 14240-14245.
- Dehaene-Lambertz, G., Monzalvo, K., Dehaene, S. (2018), « The emergence of the visual word form : Longitudinal evolution of category-specific ventral visual areas during reading acquisition », *PLoS Biology* , 16 (3), e2004103.
- Dehaene-Lambertz, G., Spelke, E. S. (2015), « The infancy of the human brain », *Neuron* , 88 (1), 93-109.
- Denison, S., Xu, F. (2010), « Integrating physical constraints in statistical inference by 11-month-old infants », *Cogn. Sci.* , 34 (5), 885-908.
- Dennett, D. C. (1996), *Darwin's Dangerous Idea : Evolution and the Meaning of Life*, New York, Simon & Schuster.
- Desimone, R., Duncan, J. (1995), « Neural mechanisms of selective visual attention », *Annu. Rev. Neurosci.* , 18, 193-222.

- D'Esposito, M., Grossman, M. (1996), « The physiological basis of executive function and working memory », *The Neuroscientist* , 2, 345-352.
- Diamond, A., Doar, B. (1989), « The performance of human infants on a measure of frontal cortex function the delayed response task », *Developmental Psychobiology* , 22.
- Diamond, A., Goldman-Rakic, P. S. (1989), « Comparison of human infants and rhesus monkeys on Piaget's A-not-B task : Evidence for dependence on dorsolateral prefrontal cortex », *Experimental Brain Research* , 74, 24-40.
- Diamond, A., Lee, K. (2011), « Interventions shown to aid executive function development in children 4 to 12 years old », *Science* , 333 (6045), 959-964.
- Diekelmann, S., Born, J. (2010), « The memory function of sleep », *Nature Reviews. Neuroscience* , 11 (2), 114-126.
- Diester, I., Nieder, A. (2007), « Semantic associations between signs and numerical categories in the prefrontal cortex », *PLoS Biology*, 5 (11), e294.
- Diester, I., Nieder, A. (2010), « Numerical values leave a semantic imprint on associated signs in monkeys », *J. Cogn. Neurosci.* , 22 (1), 174-183.
- Ditz, H. M., Nieder, A. (2015), « Neurons selective to the number of visual items in the corvid songbird endbrain », *Proc. Natl Acad. Sci. USA* , 112 (25), 7827-7832.
- Doeller, C. F., Barry, C., Burgess, N. (2010), « Evidence for grid cells in a human memory network », *Nature* , 463 (7281), 657-661.
- Donato, F., Rompani, S. B., Caroni, P. (2013), « Parvalbumin-expressing basket-cell network plasticity induced by experience regulates adult learning », *Nature* , 504 (7479), 272-276.
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., May, A. (2004), « Neuroplasticity : Changes in grey matter induced by training », *Nature* , 427 (6972), 311-312.
- Dubois, J., Dehaene-Lambertz, G., Perrin, M., Mangin, J.-F., Cointepas, Y., Duchesnay, E., ..., Hertz-Pannier, L. (2007), « Asynchrony of the early

- maturation of white matter bundles in healthy infants : Quantitative landmarks revealed noninvasively by diffusion tensor imaging », *Human Brain Mapping* .
- Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J.-F., Le Bihan, D., Dehaene-Lambertz, G. (2009), « Structural asymmetries in the infant language and sensorimotor networks », *Cereb. Cortex*, 19 (2), 414-423.
- Dubois, J., Poupon, C., Thirion, B., Simonnet, H., Kulikova, S., Leroy, F., Hertz-Pannier, L., Dehaene-Lambertz, G. (2015), « Exploring the early organization and maturation of linguistic pathways in the human infant brain », *Cereb. Cortex*, 26 (5), 2283-2298.
- Dumontheil, I., Klingberg, T. (2011), « Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later », *Cereb. Cortex*, 22 (5), 1078-1085.
- Duncan, J. (2003), « Intelligence tests predict brain response to demanding task events », *Nature Neuroscience* , 6 (3), 207-208.
- Duncan, J. (2010), « The multiple-demand (MD) system of the primate brain : Mental programs for intelligent behaviour », *Trends Cogn. Sci.* , 14 (4), 172-179.
- Duncan, J. (2013), « The structure of cognition : Attentional episodes in mind and brain », *Neuron* , 80 (1), 35-50.
- Dundas, E. M., Plaut, D. C., Behrmann, M. (2013), « The joint development of hemispheric lateralization for words and faces », *Journal of Experimental Psychology : General* , 142 (2), 348-358.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., Willingham, D. T. (2013), « Improving students' learning with effective learning techniques : Promising directions from cognitive and educational psychology », *Psychological Science in the Public Interest : A Journal of the American Psychological Society* , 14 (1), 4-58.
- Duyme, M., Dumaret, A. C., Tomkiewicz, S. (1999), « How can we boost IQs of “dull children” ? : A late adoption study », *Proc. Natl Acad. Sci. USA* , 96

(15), 8790-8794.

- Dweck, C. S. (2006), *Mindset : The New Psychology of Success* , Random House Incorporated.
- Egyed, K., Király, I., Gergely, G. (2013), « Communicating shared knowledge in infancy », *Psychological Science* , 24 (7), 1348-1353.
- Ellis, A. W., Ralph, L., Matthew, A. (2000), « Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems : Insights from connectionist networks », *J. Exp. Psychol. Learn. Mem. Cogn.* , 26 (5), 1103.
- Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., Plunkett, K. (1996), *Rethinking Innateness : A Connectionist Perspective on Development* , Cambridge (Mass.), MIT Press.
- Elsayed, G. F., Shankar, S., Cheung, B., Papernot, N., Kurakin, A., Goodfellow, I., Sohl-Dickstein, J. (2018), « Adversarial examples that fool both human and computer vision », <https://arxiv.org/abs/1802.08195> .
- Elston, G. N. (2003), « Cortex, cognition and the cell : New insights into the pyramidal neuron and prefrontal function », *Cereb. Cortex*, 13 (11), 1124-1138.
- Epelbaum, M., Milleret, C., Buisseret, P., Duffer, J. L. (1993), « The sensitive period for strabismic amblyopia in humans », *Ophthalmology* , 100 (3), 323-327.
- Esseily, R., Rat-Fischer, L., Somogyi, E., O'Regan, K. J., Fagard, J. (2016), « Humour production may enhance observational learning of a new tool-use action in 18-month-old infants », *Cognition and Emotion* , 30 (4), 817-825.
- Ester, E. F., Sprague, T. C., Serences, J. T. (2015), « Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory », *Neuron* , 87 (4), 893-905.
- Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., Bolhuis, J. J. (2015), « Structures, not strings : Linguistics as part of the cognitive



- sciences », *Trends Cogn. Sci.* , 19 (12), 729-743.
- Fanselow, M. S. (1998), « Pavlovian conditioning, negative feedback, and blocking : Mechanisms that regulate association formation », *Neuron* , 20 (4), 625-627.
- Fattal, I., Friedmann, N., Fattal-Valevski, A. (2011), « The crucial role of thiamine in the development of syntax and lexical retrieval : A study of infantile thiamine deficiency », *Brain* , 134 (Pt 6), 1720-1739.
- Fawcett, S. L., Wang, Y.-Z., Birch, E. E. (2005), « The critical period for susceptibility of human stereopsis », *Investigative Ophthalmology and Visual Science* , 46 (2), 521-525.
- Fischer, M. H. (2003), « Cognitive representation of negative numbers », *Psychological Science*, 14 (3), 278-282.
- Fisher, A. V., Godwin, K. E., Seltman, H. (2014), « Visual environment, attention allocation, and learning in young children when too much of a good thing may be bad », *Psychological Science* , 25 (7), 1362-1370.
- Fitzgerald, J. K., Freedman, D. J., Fanini, A., Bennur, S., Gold, J. I., Assad, J. A. (2013), « Biased associative representations in parietal cortex », *Neuron* , 77 (1), 180-191.
- Fitzsimonds, R. M., Song, H. J., Poo, M. M. (1997), « Propagation of activity-dependent synaptic depression in simple neural networks », *Nature* , 388 (6641), 439-448.
- Flechsig, P. (1876), *Die Leitungsbahnen im Gehirn und Rückenmark des Menschen auf Grund Entwicklungsgeschichtlicher Untersuchungen* , Leipzig, Engelmann.
- Flege, J. E., Munro, M. J., MacKay, I. R. (1995), « Factors affecting strength of perceived foreign accent in a second language », *The Journal of the Acoustical Society of America* , 97 (5), 3125-3134.
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., Rees, G. (2010), « Relating introspective accuracy to individual differences in brain structure », *Science* ,

329 (5998), 1541-1543.

- Fodor, J. A. (1975), *The Language of Thought*, New York, Thomas Y. Crowell.
- Fodor, J. A., Pylyshyn, Z. W. (1988), « Connectionism and cognitive architecture : A critical analysis », *Cognition*, 28, 3-71.
- Fodor, J., McLaughlin, B. P. (1990), « Connectionism and the problem of systematicity : Why Smolensky's solution doesn't work », *Cognition*, 35 (2), 183-204.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., Wenderoth, M. P. (2014), « Active learning increases student performance in science, engineering, and mathematics », *Proc. Natl Acad. Sci. USA*, 111 (23), 8410-8415.
- Friederici, A. D. (2002), « Towards a neural basis of auditory sentence processing », *Trends Cogn. Sci.*, 6 (2), 78-84.
- Friedmann, N., Rusou, D. (2015), « Critical period for first language : The crucial role of language input during the first year of life », *Current Opinion in Neurobiology*, 35, 27-34.
- Friedrich, M., Wilhelm, I., Born, J., Friederici, A. D. (2015), « Generalization of word meanings during infant sleep », *Nature Communications*, 6, 6004.
- Friston, K. (2005), « A theory of cortical responses », *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360 (1456), 815-836.
- Froemke, R. C., Merzenich, M. M., Schreiner, C. E. (2007), « A synaptic memory trace for cortical receptive field plasticity », *Nature*, 450 (7168), 425-429.
- Fukuchi-Shimogori, T., Grove, E. A. (2001), « Neocortex patterning by the secreted signaling molecule FGF8 », *Science*, 294 (5544), 1071-1074.
- Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., Moser, M.-B. (2004), « Spatial representation in the entorhinal cortex », *Science*, 305 (5688), 1258-1264.
- Galgali, A. R., Mante, V. (2018), « Set in one's thoughts », *Nature Neuroscience*, 21 (4), 459-460.

- Gallistel, C. R. (1990), *The Organization of Learning* , Cambridge (Mass.), MIT Press.
- Gaser, C., Schlaug, G. (2003), « Brain structures differ between musicians and non-musicians », *J. Neurosci.* , 23 (27), 9240-9245.
- Gathercole, S. E., Pickering, S. J., Knight, C., Stegmann, Z. (2004), « Working memory skills and educational attainment : Evidence from national curriculum assessments at 7 and 14 years of age », *Applied Cognitive Psychology* , 18 (1), 1-16.
- Geary, D. C. (2011), « Cognitive predictors of achievement growth in mathematics : A five year longitudinal study », *Developmental Psychology* , 47 (6), 1539-1552.
- Genzel, L., Rossato, J. I., Jacobse, J., Grieves, R. M., Spooner, P. A., Battaglia, F. P., Fernandez, G., Morris, R. G. M. (2017), « The yin and yang of memory consolidation : Hippocampal and neocortical », *PLoS Biology* , 15 (1).
- George, D., Lehrach, W., Kansky, K., Lázaro-Gredilla, M., Laan, C., Marthi, B., ..., Phoenix, D. S. (2017), « A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs », *Science* , 358 (6368).
- Gerber, P., Schlaffke, L., Heba, S., Greenlee, M. W., Schultz, T., Schmidt-Wilcke, T. (2014), « Juggling revisited – A voxel-based morphometry study with expert jugglers », *NeuroImage* , 95, 320-325.
- Gergely, G., Bekkering, H., Kiraly, I. (2002), « Rational imitation in preverbal infants », *Nature* , 415 (6873), 755.
- Gergely, G., Csibra, G. (2003), « Teleological reasoning in infancy : The naïve theory of rational action », *Trends Cogn. Sci.* , 7 (7), 287-292.
- Gerhand, S., Barry, C. (1999), « Age of acquisition, word frequency, and the role of phonology in the lexical decision task », *Memory and Cognition* , 27 (4), 592-602.

- Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., ..., Grantham-McGregor, S. (2014), « Labor market returns to an early childhood stimulation intervention in Jamaica », *Science* , 344 (6187), 998-1001.
- Goldman-Rakic, P. S. (1995), « Cellular basis of working memory », *Neuron* , 14 (3), 477-485.
- Golestani, N., Molko, N., Dehaene, S., Le Bihan, D., Pallier, C. (2007), « Brain structure predicts the learning of foreign speech sounds », *Cereb. Cortex*, 17 (3), 575-582.
- Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., ..., Yu, B. M. (2018), « Learning by neural reassociation », *Nature Neuroscience* , 21 (4), 607-616.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ..., Bengio, Y. (2014), « Generative adversarial networks », <https://arxiv.org/abs/1406.2661> .
- Goodman, C. S., Shatz, C. J. (1993), « Developmental mechanisms that generate precise patterns of neuronal connectivity », *Neuron* , 10, 77-98.
- Goodman, N. D., Ullman, T. D., Tenenbaum, J. B. (2011), « Learning a theory of causality » *Psychological Review* , 118 (1), 110-119.
- Gopnik, A. (2000), *The Scientist in the Crib : What Early Learning Tells Us About the Mind* , New York, William Morrow.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., Danks, D. (2004), « A theory of causal learning in children : Causal maps and Bayes nets », *Psychological Review* , 111 (1), 3-32.
- Gottlieb, J., Oudeyer, P.-Y., Lopes, M., Baranes, A. (2013), « Information-seeking, curiosity, and attention : Computational and neural mechanisms », *Trends Cogn. Sci.* , 17 (11), 585-593.
- Goupil, L., Romand-Monnier, M., Kouider, S. (2016), « Infants ask for help when they know they don't know », *Proc. Natl Acad. Sci. USA* , 113 (13),

3492-3496.

- Grainger, J., Whitney, C. (2004), « Does the human mind read words as a whole ? » *Trends Cogn. Sci.* , 8 (2), 58-59.
- Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., Himes, J. H. (1991), « Nutritional supplementation, psychosocial stimulation, and mental development of stunted children : The Jamaican Study », *The Lancet* , 338 (8758), 1-5.
- Green, C. S., Bavelier, D. (2003), « Action video game modifies visual selective attention », *Nature* , 423 (6939), 534-537.
- Groen, G. J., Parkman, J. M. (1972), « A chronometric analysis of simple addition », *Psychological Review* , 79, 329-343.
- Grothendieck, A. (1986), *Récoltes et semailles* , [https://www.quarante-deux.org/archives/klein/prefaces/Romans\\_1965-1969/Recoltes\\_et\\_semailles.pdf](https://www.quarante-deux.org/archives/klein/prefaces/Romans_1965-1969/Recoltes_et_semailles.pdf) .
- Gruber, M. J., Gelman, B. D., Ranganath, C. (2014), « States of curiosity modulate hippocampus : Dependent learning via the dopaminergic circuit », *Neuron* , 84 (2), 486-496.
- Guerguiev, J., Lillicrap, T. P., Richards, B. A. (2017), « Towards deep learning with segregated dendrites », *ELife* , 6, e22901.
- Gullick, M. M., Wolford, G. (2013), « Understanding less than nothing : Children's neural response to negative numbers shifts across age and accuracy », *Frontiers in Psychology* , 4, 584.
- Gweon, H., Tenenbaum, J. B., Schulz, L. E. (2010), « Infants consider both the sample and the sampling process in inductive generalization », *Proc. Natl Acad. Sci. USA* , 107 (20), 9066-9071.
- Habibi, A., Damasio, A., Ilari, B., Elliott Sachs, M., Damasio, H. (2018), « Music training and child development : A review of recent findings from a longitudinal study », *Annals of the New York Academy of Sciences* .

- Hafting, T., Fyhn, M., Molden, S., Moser, M. B., Moser, E. I. (2005), « Microstructure of a spatial map in the entorhinal cortex », *Nature* , 436 (7052), 801-806.
- Hahne, A., Friederici, A. D. (1999), « Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes », *J. Cogn. Neurosci.*, 11 (2), 194-205.
- Halberda, J., Feigenson, L. (2008), « Developmental change in the acuity of the “number sense” : The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults », *Dev. Psychol.* , 44 (5), 1457-1465.
- Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., Dehaene, S. (2015), « Origins of the specialization for letters and numbers in ventral occipitotemporal cortex », *Trends Cogn. Sci.* , 19 (7), 374-382.
- Hannagan, T., Nieder, A., Viswanathan, P., Dehaene, S. (2017), « A random-matrix theory of the number sense », *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* , 373 (1740).
- Hartshorne, J. K., Tenenbaum, J. B., Pinker, S. (2018), « A critical period for second language acquisition : Evidence from 2/3 million English speakers » *Cognition* , 177, 263-277.
- Hassabis, D., Kumaran, D., Summerfield, C., Botvinick, M. (2017), « Neuroscience-inspired artificial intelligence », *Neuron* , 95 (2), 245-258.
- Hattie, J. (2008), *Visible Learning* , Londres/New York, Routledge.
- Hattie, J., Brodeur, M., St-Cyr, C. (2017), *L'Apprentissage visible pour les enseignants. Connaître son impact pour maximiser le rendement des élèves* , Québec, Presses de l'Université du Québec.
- Hauser, M. D., Chomsky, N., Fitch, W. T. (2002), « The faculty of language : What is it, who has it, and how did it evolve ? », *Science* , 298 (5598), 1569-1579.
- Hauser, M. D., Watumull, J. (2017), « The Universal Generative Faculty : The source of our expressive power in language, mathematics, morality, and

- music », *Journal of Neurolinguistics* , 43 (B), 78-94.
- Hay, J. F., Pelucchi, B., Graf Estes, K., Saffran, J. R. (2011), « Linking sounds to meanings : Infant statistical learning in a natural language », *Cognitive Psychology* , 63 (2), 93-106.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., Yavitz, A. (2010), « The rate of return to the HighScope Perry Preschool Program », *Journal of Public Economics* , 94 (1), 114-128.
- Held, R., Hein, A. (1963), « Movement-produced stimulation in the development of visually guided behavior », *Journal of Comparative and Physiological Psychology* , 56, 872-876.
- Hensch, T. K. (2005), « Critical period plasticity in local cortical circuits », *Nature Reviews. Neuroscience* , 6 (11), 877.
- Hespos, S. J., Baillargeon, R. (2008), « Young infants' actions reveal their developing knowledge of support variables : Converging evidence for violation-of-expectation findings », *Cognition* , 107 (1), 304-316.
- Hinton, G. E., Dayan, P., Frey, B. J., Neal, R. M. (1995), « The “wake-sleep” algorithm for unsupervised neural networks », *Science* , 268 (5214), 1158-1161.
- Hinton, G. E., Osindero, S., Teh, Y. W. (2006), « A fast learning algorithm for deep belief nets », *Neural Computation* , 18 (7), 1527-1554.
- Hiscock, H., Sciberras, E., Mensah, F., Gerner, B., Efron, D., Khano, S., Oberklaid, F. (2015), « Impact of a behavioural sleep intervention on symptoms and sleep in children with attention deficit hyperactivity disorder, and parental mental health : Randomised controlled trial », *BMJ (Clinical Research Ed.)* , 350, h68.
- Holtmaat, A., Caroni, P. (2016), « Functional and structural underpinnings of neuronal assembly formation in learning », *Nature Neuroscience* , 19 (12), 1553-1562.

- Horikawa, T., Tamaki, M., Miyawaki, Y., Kamitani, Y. (2013), « Neural decoding of visual imagery during sleep », *Science* , 340 (6132), 639-642.
- Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., Tzourio-Mazoyer, N. (2000), « Shifting from the perceptual brain to the logical brain : The neural impact of cognitive inhibition training », *J. Cogn. Neurosci.*, 12 (5), 721-728.
- Huber, R., Ghilardi, M. F., Massimini, M., Tononi, G. (2004), « Local sleep and learning », *Nature* , 430 (6995), 78-81.
- Hurley, M. M., Dennett, D. C., Adams, R. B. (2011), *Inside Jokes : Using Humor to Reverse-Engineer the Mind* , Cambridge (Mass.), MIT Press.
- Huttenlocher, P. R., Dabholkar, A. S. (1997), « Regional differences in synaptogenesis in human cerebral cortex », *J. Comp. Neurol.* , 387 (2), 167-178.
- Hutton, J. S., Horowitz-Kraus, T., Mendelsohn, A. L., DeWitt, T., Holland, S. K., C-MIND Authorship Consortium (2015), « Home reading environment and brain activation in preschool children listening to stories », *Pediatrics* , 136 (3), 466-478.
- Hutton, J. S., Phelan, K., Horowitz-Kraus, T., Dudley, J., Altaye, M., DeWitt, T., Holland, S. K. (2017), « Shared reading quality and brain activation during story listening in preschool-age children », *Journal of Pediatrics* , 191, 204-211.e1.
- Iriki, A. (2005), « A prototype of Homo-Faber : A silent precursor of human intelligence in the tool-using monkey brain », in S. Dehaene, J. R. Duhamel, M. Hauser, G. Rizzolatti (éd.), *From Monkey Brain to Human Brain* , Cambridge (Mass.), MIT Press, 133-157.
- Isingrini, M., Perrotin, A., Souchay, C. (2008), « Aging, metamemory regulation and executive functioning », *Prog. Brain Res.* , 169, 377-392.
- Izard, V., Dehaene-Lambertz, G., Dehaene, S. (2008), « Distinct cerebral pathways for object identity and number in human infants », *PLoS Biology*, 6 (2), 275-285.



- Izard, V., Sann, C., Spelke, E. S., Streri, A. (2009), « Newborn infants perceive abstract numbers », *Proc. Natl Acad. Sci. USA* , 106 (25), 10382-10385.
- Jacob, S. N., Nieder, A. (2009), « Notation-independent representation of fractions in the human parietal cortex », *J. Neurosci.* , 29 (14), 4652-4657.
- Jacoby, L. L., Dallas, M. (1981), « On the relationship between autobiographical memory and perceptual learning », *Journal of Experimental Psychology : General* , 110 (3), 306-340.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., Shah, P. (2011), « Short-and long-term benefits of cognitive training », *Proc. Natl Acad. Sci. USA* , 108 (25), 10081-10086.
- James, C. E., Oechslin, M. S., Van De Ville, D., Hauert, C.-A., Descloux, C., Lazeyras, F. (2014), « Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks », *Brain Structure and Function* , 219 (1), 353-366.
- Jaynes, E. T. (2003), *Probability Theory : The Logic of Science* , Cambridge, Cambridge University Press.
- Jenkins, J. G., Dallenbach, K. M. (1924), « Obliviscence during sleep and waking », *The American Journal of Psychology* , 35 (4), 605-612.
- Ji, D., Wilson, M. A. (2007), « Coordinated memory replay in the visual cortex and hippocampus during sleep », *Nature Neuroscience* , 10 (1), 100-107.
- Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., Wang, L. (2018), « Production of supra-regular spatial sequences by macaque monkeys », *Current Biology* , 28 (12), 1851-1859.
- Jiang, X., Shamie, I., K. Doyle, W., Friedman, D., Dugan, P., Devinsky, O., ..., Halgren, E. (2017), « Replay of large-scale spatio-temporal patterns from waking during subsequent NREM sleep in human cortex », *Scientific Reports* , 7 (1), 17380.
- Jo, J., Bengio, Y. (2017), « Measuring the tendency of CNNs to Learn surface statistical regularities », <https://arxiv.org/abs/1711.11561> .

- Johansson, F., Jirenhed, D. A., Rasmussen, A., Zucca, R., Hesslow, G. (2014), « Memory trace and timing mechanism localized to cerebellar Purkinje cells », *Proc. Natl Acad. Sci. USA* , 111 (41), 14930-14934.
- Johnson, J. S., Newport, E. L. (1989), « Critical period effects in second language learning : The influence of maturational state on the acquisition of English as a second language », *Cognitive Psychology* , 21 (1), 60-99.
- Josselyn, S. A., Köhler, S., Frankland, P. W. (2015), « Finding the engram », *Nature Reviews. Neuroscience* , 16 (9), 521-534.
- Kaminski, J., Call, J., Fischer, J. (2004), « Word learning in a domestic dog : Evidence for “fast mapping” », *Science* , 304 (5677), 1682-1683.
- Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., Camerer, C. F. (2009), « The wick in the candle of learning : Epistemic curiosity activates reward circuitry and enhances memory », *Psychological Science*, 20 (8), 963-973.
- Kang, S. H. K., Lindsey, R. V., Mozer, M. C., Pashler, H. (2014), « Retrieval practice over the long term : should spacing be expanding or equal-interval ? », *Psychonomic Bulletin and Review* , 21 (6), 1544-1550.
- Kanjlia, S., Lane, C., Feigenson, L., Bedny, M. (2016), « Absence of visual experience modifies the neural basis of numerical thinking », *Proc. Natl Acad. Sci. USA* , 113 (40), 11172-11177.
- Kano, T., Brockie, P. J., Sassa, T., Fujimoto, H., Kawahara, Y., Iino, Y., ..., Maricq, A. V. (2008), « Memory in *Caenorhabditis elegans* is mediated by NMDA-type ionotropic glutamate receptors », *Current Biology* , 18 (13), 1010-1015.
- Kaplan, F., Oudeyer, P.-Y. (2007), « In search of the neural circuits of intrinsic motivation », *Frontiers in Neuroscience* , 1 (1), 225.
- Kapur, S., Craik, F. I., Tulving, E., Wilson, A. A., Houle, S., Brown, G. M. (1994), « Neuroanatomical correlates of encoding in episodic memory : Levels of processing effect », *Proc. Natl Acad. Sci. USA* , 91 (6), 2008-2011.

- Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J., Sagi, D. (1994), « Dependence on REM sleep of overnight improvement of a perceptual skill », *Science* , 265 (5172), 679-682.
- Karpicke, J. D., Roediger, H. L. (2008), « The critical importance of retrieval for learning », *Science* , 319 (5865), 966-968.
- Kastner, S., Ungerleider, L. G. (2000), « Mechanisms of visual attention in the human cortex », *Annu. Rev. Neurosci.* , 23, 315-341.
- Keller, H. (1903), *The Story of My Life* , New York, Simon & Schuster.
- Kellman, P. J., Spelke, E. S. (1983), « Perception of partly occluded objects in infancy », *Cognitive Psychology* , 15, 483-524.
- Kemp, C., Tenenbaum, J. B. (2008), « The discovery of structural form », *Proc. Natl Acad. Sci. USA* , 105 (31), 10687-10692.
- Kerkoerle, T. van, Self, M. W., Roelfsema, P. R. (2017), « Layer-specificity in the effects of attention and working memory on activity in primary visual cortex », *Nature Communications* , 8, 13804.
- Kidd, C., Piantadosi, S. T., Aslin, R. N. (2012), « The Goldilocks effect : Human infants allocate attention to visual sequences that are neither too simple nor too complex », *PLoS ONE* , 7 (5), e36399.
- Kidd, C., Piantadosi, S. T., Aslin, R. N. (2014), « The Goldilocks effect in infant auditory attention », *Child Development* , 85 (5), 1795-1804.
- Kilgard, M. P., Merzenich, M. M. (1998), « Cortical map reorganization enabled by nucleus basalis activity », *Science* , 279 (5357), 1714-1718.
- Kim, J. J., Diamond, D. M. (2002), « The stressed hippocampus, synaptic plasticity and lost memories », *Nature Reviews. Neuroscience* , 3 (6), 453-462.
- Kim, W. B., Cho, J. H. (2017), « Encoding of discriminative fear memory by input-specific LTP in the amygdala », *Neuron* , 95 (5), 1129-1146.e5.
- Kirschner, P. A., Sweller, J., Clark, R. E. (2006), « Why minimal guidance during instruction does not work : An analysis of the failure of

- constructivist, discovery, problem-based, experiential, and inquiry-based teaching », *Educational Psychologist* , 41 (2), 75-86.
- Kirschner, P. A., Van Merriënboer, J. J. G. (2013), « Do learners really know best ? Urban legends in education », *Educational Psychologist* , 48 (3), 169-183.
- Kitamura, T., Ogawa, S. K., Roy, D. S., Okuyama, T., Morrissey, M. D., Smith, L. M., ..., Tonegawa, S. (2017), « Engrams and circuits crucial for systems consolidation of a memory », *Science* , 356 (6333), 73-78.
- Klingberg, T. (2010), « Training and plasticity of working memory », *Trends Cogn. Sci.* , 14 (7), 317-324.
- Knops, A., Thirion, B., Hubbard, E. M., Michel, V., Dehaene, S. (2009), « Recruitment of an area involved in eye movements during mental arithmetic », *Science* , 324 (5934), 1583-1585.
- Knops, A., Viarouge, A., Dehaene, S. (2009), « Dynamic representations underlying symbolic and nonsymbolic calculation : Evidence from the operational momentum effect », *Atten. Percept. Psychophys.* , 71 (4), 803-821.
- Knudsen, E. I., Knudsen, P. F. (1990), « Sensitive and critical periods for visual calibration of sound localization by barn owls », *J. Neurosci.*, 10 (1), 222-232.
- Knudsen, E. I., Zheng, W., DeBello, W. M. (2000), « Traces of learning in the auditory localization pathway », *Proc. Natl Acad. Sci. USA* , 97 (22), 11815-11820.
- Koechlin, E., Dehaene, S., Mehler, J. (1997), « Numerical transformations in five month old human infants », *Mathematical Cognition* , 3, 89-104.
- Koechlin, E., Ody, C., Kouneiher, F. (2003), « The architecture of cognitive control in the human prefrontal cortex », *Science* , 302 (5648), 1181-1185.
- Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., ..., Grasby, P. M. (1998), « Evidence for striatal dopamine release

- during a video game », *Nature* , 393, 266-268.
- Kolinsky, R., Morais, J., Content, A., Cary, L. (1987), « Finding parts within figures : A developmental study », *Perception* , 16 (3), 399-407.
- Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-Cabral, L., Morais, J. (2011), « Enantiomorphy through the Looking-Glass : Literacy effects on mirror-image discrimination », *Journal of Experimental Psychology : General* , 140 (2), 210-238.
- Kontra, C., Goldin-Meadow, S., Beilock, S. L. (2012), « Embodied learning across the life span », *Topics in Cognitive Science* , 4 (4), 731-739.
- Kontra, C., Lyons, D. J., Fischer, S. M., Beilock, S. L. (2015), « Physical experience enhances science learning », *Psychological Science* , 26 (6), 737-749.
- Kouider, S., Stahlhut, C., Gelskov, S. V., Barbosa, L. S., Dutat, M., Gardelle, V. de, Christophe, A., Dehaene, S., Dehaene-Lambertz, G. (2013), « A neural marker of perceptual consciousness in infants », *Science* , 340 (6130), 376-380.
- Krause, M. R., Zanos, T. P., Csorba, B. A., Pilly, P. K., Choe, J., Phillips, M. E., ..., Pack, C. C. (2017), « Transcranial direct current stimulation facilitates associative learning and alters functional connectivity in the primate brain », *Current Biology* , 27 (20), 3086-3096.
- Kropff, E., Treves, A. (2008), « The emergence of grid cells : Intelligent design or just adaptation ? », *Hippocampus* , 18 (12), 1256-1269.
- Krubitzer, L. (2007), « The magnificent compromise : Cortical field evolution in mammals », *Neuron* , 56 (2), 201-208.
- Kuhl, P. K., Tsao, F. M., Liu, H. M. (2003), « Foreign-language experience in infancy : Effects of short-term exposure and social interaction on phonetic learning », *Proc. Natl Acad. Sci. USA* , 100 (15), 9096-9101.
- Kurdziel, L., Duclos, K., Spencer, R. M. C. (2013), « Sleep spindles in midday naps enhance learning in preschool children », *Proc. Natl Acad. Sci. USA* ,

110 (43), 17267-17272.

- Kushnir, T., Xu, F., Wellman, H. M. (2010), « Young children use statistical sampling to infer the preferences of other people », *Psychological Science*, 21 (8), 1134-1140.
- Kutas, M., Federmeier, K. D. (2011), « Thirty years and counting : Finding meaning in the N400 component of the event-related brain potential (ERP) », *Annu. Rev. Psychol.* , 62, 621-647.
- Kutas, M., Hillyard, S. A. (1980), « Reading senseless sentences : Brain potentials reflect semantic incongruity », *Science* , 207 (4427), 203-205.
- Kwan, K. Y., Lam, M. M. S., Johnson, M. B., Dube, U., Shim, S., Rašin, M. R., ..., Sestan, N. (2012), « Species-dependent posttranscriptional regulation of NOS1 by FMRP in the developing cerebral cortex », *Cell* , 149 (4), 899-911.
- Lake, B. M., Salakhutdinov, R., Tenenbaum, J. B. (2015), « Human-level concept learning through probabilistic program induction », *Science* , 350 (6266), 1332-1338.
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., Gershman, S. J. (2016), « Building machines that learn and think like people », *Behav. Brain Sci.* , 1-101.
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., Gershman, S. J. (2017), « Building machines that learn and think like people », *Behav. Brain Sci.* , 40, e253.
- Landau, B., Gleitman, H., Spelke, E. (1981), « Spatial knowledge and geometric representation in a child blind from birth », *Science* , 213 (4513), 1275-1278.
- Lane, C., Kanjlia, S., Omaki, A., Bedny, M. (2015), « “Visual” cortex of congenitally blind adults responds to syntactic movement », *J. Neurosci.* , 35 (37), 12859-12868.
- Langston, R. F., Ainge, J. A., Couey, J. J., Canto, C. B., Bjerknes, T. L., Witter, M. P., ..., Moser, M. B. (2010), « Development of the spatial representation system in the rat », *Science* , 328 (5985), 1576-1580.

- LeCun, Y., Bengio, Y., Hinton, G. (2015), « Deep learning », *Nature* , 521 (7553), 436-444.
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998), « Gradient-based learning applied to document recognition », *Proceedings of the IEEE* , 86 (11), 2278-2324.
- Lefèvre, J., Mangin, J.-F. (2010), « A reaction-diffusion model of human brain development », *PLoS Comput. Biology* , 6 (4), e1000749.
- Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., Niv, Y. (2017), « Dynamic interaction between reinforcement learning and attention in multidimensional environments », *Neuron* , 93 (2), 451-463.
- Lerner, Y., Honey, C. J., Silbert, L. J., Hasson, U. (2011), « Topographic mapping of a hierarchy of temporal receptive windows using a narrated story », *J. Neurosci.*, 31 (8), 2906-2915.
- Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., ..., Dehaene-Lambertz, G. (2015), « New human-specific brain landmark : The depth asymmetry of superior temporal sulcus », *Proc. Natl Acad. Sci. USA* , 112 (4), 1208-1213.
- Li, P., Legault, J., Litcofsky, K. A. (2014), « Neuroplasticity as a function of second language learning : Anatomical changes in the human brain », *Cortex* , 58, 301-324.
- Li, S., Lee, K., Zhao, J., Yang, Z., He, S., Weng, X. (2013), « Neural competition as a developmental process : Early hemispheric specialization for word processing delays specialization for face processing », *Neuropsychologia* , 51 (5), 950-959.
- Lillard, A., Else-Quest, N. (2006), « The early years. Evaluating Montessori education », *Science* , 313 (5795), 1893-1894.
- Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Ranganath, C., Redish, A. D. (2017), « Viewpoints : How the hippocampus contributes to memory, navigation and cognition », *Nature Neuroscience* , 20, 1434-1447.

- Liu, S., Ullman, T. D., Tenenbaum, J. B., Spelke, E. S. (2017), « Ten-month-old infants infer the value of goals from the costs of actions », *Science* , 358 (6366), 1038-1041.
- Livingstone, M. S., Vincent, J. L., Arcaro, M. J., Srihasam, K., Schade, P. F., Savage, T. (2017), « Development of the macaque face-patch system », *Nature Communications* , 8, 14897.
- Loewenstein, G. (1994), « The psychology of curiosity : A review and reinterpretation », *Psychological Bulletin* , 116 (1), 75-98.
- Lømo, T. (2018), « Discovering long-term potentiation (LTP) – Recollections and reflections on what came after », *Acta Physiologica (Oxford)* , 222 (2).
- Louie, K., Wilson, M. A. (2001), « Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep », *Neuron* , 29 (1), 145-156.
- Lyons, I. M., Beilock, S. L. (2012), « When math hurts : math anxiety predicts pain network activation in anticipation of doing math », *PLoS ONE* , 7 (10), e48076.
- Lyons, K. E., Ghetti, S. (2011), « The development of uncertainty monitoring in early childhood », *Child Development* , 82 (6), 1778-1787.
- Ma, L., Xu, F. (2013), « Preverbal infants infer intentional agents from the perception of regularity », *Developmental Psychology* , 49 (7), 1330-1337.
- Mack, A., Rock, I. (1998), *Inattentive Blindness* , Cambridge (Mass.), MIT Press.
- Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., Frith, C. D. (2000), « Navigation-related structural change in the hippocampi of taxi drivers », *Proc. Natl Acad. Sci. USA* , 97 (8), 4398-4403.
- Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S., Burgess, N. (2003), « Navigation expertise and the human hippocampus : A structural brain imaging analysis », *Hippocampus* , 13 (2), 250-259.



- Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Godjil, S., Dubois, J., ..., Wallois, F. (2013), « Syllabic discrimination in premature human infants prior to complete formation of cortical layers », *Proc. Natl Acad. Sci. USA* , 110 (12), 4846-4851.
- Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., Caramazza, A. (2009), « Category-specific organization in the human brain does not require visual experience », *Neuron* , 63 (3), 397-405.
- Maloney, E. A., Beilock, S. L. (2012), « Math anxiety : Who has it, why it develops, and how to guard against it », *Trends Cogn. Sci.* , 16 (8), 404-406.
- Markman, E. M., Wachtel, G. F. (1988), « Children's use of mutual exclusivity to constrain the meanings of words », *Cognitive Psychology* , 20 (2), 121-157.
- Markman, E. M., Wasow, J. L., Hansen, M. B. (2003), « Use of the mutual exclusivity assumption by young word learners », *Cognitive Psychology* , 47 (3), 241-275.
- Marques, J. F., Dehaene, S. (2004), « Developing intuition for prices in euros : Rescaling or relearning prices ? », *J. Exp. Psychol. Appl.* , 10 (3), 148-155.
- Marshall, C. (2017), « Montessori education : A review of the evidence base », *NPJ Science of Learning* , 2 (1), 11.
- Marshall, L., Helgadottir, H., Molle, M., Born, J. (2006), « Boosting slow oscillations during sleep potentiates memory », *Nature* , 444 (7119), 610-613.
- Marti, S., King, J.-R., Dehaene, S. (2015), « Time-resolved decoding of two processing chains during dual-task interference », *Neuron* , 88 (6), 1297-1307.
- Marti, S., Sigman, M., Dehaene, S. (2012), « A shared cortical bottleneck underlying Attentional Blink and Psychological Refractory Period », *NeuroImage* , 59 (3), 2883-2898.

- Martin, S. L., Ramey, C. T., Ramey, S. (1990), « The prevention of intellectual impairment in children of impoverished families : Findings of a randomized trial of educational day care », *American Journal of Public Health* , 80 (7), 844-847.
- Maye, J., Werker, J. F., Gerken, L. (2002), « Infant sensitivity to distributional information can affect phonetic discrimination », *Cognition* , 82 (3), B101-11.
- Mayer, R. E. (2004), « Should there be a three-strikes rule against pure discovery learning ? The case for guided methods of instruction », *The American Psychologist* , 59 (1), 14-19.
- McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., McClelland, J. L. (2002), « Success and failure in teaching the [r]-[l] contrast to Japanese adults : Tests of a Hebbian model of plasticity and stabilization in spoken language perception », *Cognitive, Affective and Behavioral Neuroscience* , 2 (2), 89-108.
- McCrink, K., Wynn, K. (2004), « Large-number addition and subtraction by 9-month-old infants », *Psychological Science*, 15 (11), 776-781.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J., Amiel-Tison, C. (1988), « A precursor of language acquisition in young infants », *Cognition* , 29 (2), 143-178.
- Meyer, T., Olson, C. R. (2011), « Statistical learning of visual transitions in monkey inferotemporal cortex », *Proc. Natl Acad. Sci. USA* , 108 (48), 19401-19406.
- Millum, J., Emanuel, E. J. (2007), « Ethics. The ethics of international research with abandoned children », *Science* , 318 (5858), 1874-1875.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ..., Hassabis, D. (2015), « Human-level control through deep reinforcement learning », *Nature* , 518 (7540), 529-533.
- Mongelli, V., Dehaene, S., Vinckier, F., Peretz, I., Bartolomeo, P., Cohen, L. (2017), « Music and words in the visual cortex : The impact of musical

- expertise », *Cortex* , 86, 260-274.
- Mongillo, G., Barak, O., Tsodyks, M. (2008), « Synaptic theory of working memory », *Science* , 319 (5869), 1543-1546.
- Monzalvo, K., Fluss, J., Billard, C., Dehaene, S., Dehaene-Lambertz, G. (2012), « Cortical networks for vision and language in dyslexic and normal children of variable socioeconomic status », *NeuroImage* , 61 (1), 258-274.
- Morais, J. (2017), « Literacy and democracy », *Language, Cognition and Neuroscience* , 33 (3), 351-371.
- Morais, J., Bertelson, P., Cary, L., Alegria, J. (1986), « Literacy training and speech segmentation », *Cognition* , 24, 45-64.
- Morais, J., Kolinsky, R. (2005), « Literacy and cognitive change », in M. Snowling, C. Hulme (éd.), *The Science of Reading : A Handbook* , Oxford, Blackwell, 188-203.
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., Chau, T. (2011), « Short-term music training enhances verbal intelligence and executive function », *Psychological Science* , 22 (11), 1425-1433.
- Morrison, C. M., Ellis, A. W. (1995), « Roles of word frequency and age of acquisition in word naming and lexical decision », *J. Exp. Psychol. Learn. Mem. Cogn.* , 21 (1), 116.
- Morton, J., Johnson, M. H. (1991), « CONSPEC and CONLERN : A two-process theory of infant face recognition », *Psychol. Rev.* , 98 (2), 164-181.
- Mottint, O. (2018), « Faut-il renoncer aux pédagogies actives ? », *democratisation-scolaire.fr*, 29 mars, <https://www.democratisation-scolaire.-fr/spip.php?article287> .
- Moyer, R. S., Landauer, T. K. (1967), « Time required for judgements of numerical inequality », *Nature* , 215, 1519-1520.
- Muckli, L., Naumer, M. J., Singer, W. (2009), « Bilateral visual field maps in a patient with only one hemisphere », *Proc. Natl Acad. Sci. USA* , 106 (31), 13034-13039.

- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C., Weiller, C. (2003), « Broca's area and the language instinct », *Nature Neuroscience* , 6 (7), 774-781.
- Naatanen, R., Paavilainen, P., Rinne, T., Alho, K. (2007), « The mismatch negativity (MMN) in basic research of central auditory processing : A review », *Clin. Neurophysiol.* , 118 (12), 2544-2590.
- Nau, M., Navarro Schröder, T., Bellmund, J. L. S., Doeller, C. F. (2018), « Hexadirectional coding of visual space in human entorhinal cortex », *Nature Neuroscience* , 21 (2), 188-190.
- Nelson, C. A., Zeanah, C. H., Fox, N. A., Marshall, P. J., Smyke, A. T., Guthrie, D. (2007), « Cognitive recovery in socially deprived young children : The Bucharest Early Intervention Project », *Science* , 318 (5858), 1937-1940.
- Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., ..., Dehaene, S. (2017), « Neurophysiological dynamics of phrase-structure building during sentence processing », *Proc. Natl Acad. Sci. USA* , 114 (18), E3669-E3678.
- Nemmi, F., Helander, E., Helenius, O., Almeida, R., Hassler, M., Räsänen, P., Klingberg, T. (2016), « Behavior and neuroimaging at baseline predict individual response to combined mathematical and working memory training in children », *Developmental Cognitive Neuroscience* , 20, 43-51.
- Ngo, H.-V. V., Martinetz, T., Born, J., Mölle, M. (2013), « Auditory closed-loop stimulation of the sleep slow oscillation enhances memory », *Neuron* , 78 (3), 545-553.
- Nieder, A., Dehaene, S. (2009), « Representation of number in the brain », *Annu. Rev. Neurosci.* , 32, 185-208.
- Noble, K. G., Norman, M. F., Farah, M. J. (2005), « Neurocognitive correlates of socioeconomic status in kindergarten children », *Dev. Sci.* , 8 (1), 74-87.
- Norimoto, H., Makino, K., Gao, M., Shikano, Y., Okamoto, K., Ishikawa, T., ..., Ikegaya, Y. (2018), « Hippocampal ripples down-regulate synapses », *Science* , 359 (6383), 1524-1527.

- Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., ..., Iriki, A. (2001), « Functional brain mapping of monkey tool use », *NeuroImage* , 14 (4), 853-861.
- Oechslin, M. S., Gschwind, M., James, C. E. (2018), « Tracking training-related plasticity by combining fMRI and DTI : The right hemisphere ventral stream mediates musical syntax processing », *Cereb. Cortex*, 28 (4), 1209-1218.
- Olah, C., Mordvintsev, A., Schubert, L. (2017), « Feature visualization : How neural networks build up their understanding of images », Distill, 7 novembre, <https://distill.pub/2017/feature-visualization/> .
- Olesen, P. J., Westerberg, H., Klingberg, T. (2004), « Increased prefrontal and parietal activity after training of working memory », *Nature Neuroscience* , 7 (1), 75-79.
- Orbán, G., Berkes, P., Fiser, J., Lengyel, M. (2016), « Neural variability and sampling-based probabilistic representations in the visual cortex », *Neuron* , 92 (2), 530-543.
- Paller, K. A., McCarthy, G., Wood, C. C. (1988), « ERPs predictive of subsequent recall and recognition performance », *Biological Psychology* , 26 (1-3), 269-276.
- Pallier, C., Dehaene, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., Mehler, J. (2003), « Brain imaging of language plasticity in adopted adults : Can a second language replace the first ? », *Cereb. Cortex*, 13 (2), 155-161.
- Pallier, C., Devauchelle, A. D., Dehaene, S. (2011), « Cortical representation of the constituent structure of sentences », *Proc. Natl Acad. Sci. USA* , 108 (6), 2522-2527.
- Palminteri, S., Kilford, E. J., Coricelli, G., Blakemore, S.-J. (2016), « The computational development of reinforcement learning during adolescence », *PLoS Computational Biology* , 12 (6), e1004953.
- Pashler, H., McDaniel, M., Rohrer, D., Bjork, R. (2008), « Learning styles concepts and evidence », *Psychological Science in the Public Interest* , 9 (3), 105-119.

- Pegado, F., Comerlato, E., Ventura, F., Jobert, A., Nakamura, K., Buiatti, M., ..., Dehaene, S. (2014), « Timing the impact of literacy on visual processing », *Proc. Natl Acad. Sci. USA* , 111 (49), E5233-5242.
- Pegado, F., Nakamura, K., Braga, L. W., Ventura, P., Nunes Filho, G., Pallier, C., ..., Dehaene, S. (2014), « Literacy breaks mirror invariance for visual stimuli : A behavioral study with adult illiterates », *Journal of Experimental Psychology : General* , 143 (2), 887-894.
- Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., ..., Maquet, P. (2004) », « Are spatial memories strengthened in the human hippocampus during slow wave sleep ? », *Neuron* , 44 (3), 535-545.
- Pena, M., Werker, J. F., Dehaene-Lambertz, G. (2012), « Earlier speech exposure does not accelerate speech acquisition », *J. Neurosci.*, 32 (33), 11159-11163.
- Penn, D. C., Holyoak, K. J., Povinelli, D. J. (2008), « Darwin's mistake : Explaining the discontinuity between human and nonhuman minds », *Behav. Brain Sci.* , 31 (2), 109-130 ; discussion, 130-178.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., Frith, C. D. (2006), « Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans », *Nature* , 442 (7106), 1042-1045.
- Piantadosi, S. T., Tenenbaum, J. B., Goodman, N. D. (2012), « Bootstrapping in a language of thought : A formal model of numerical concept learning », *Cognition* , 123 (2), 199-217.
- Piantadosi, S. T., Tenenbaum, J. B., Goodman, N. D. (2016), « The logical primitives of thought : Empirical foundations for compositional cognitive models », *Psychological Review* , 123 (4), 392-424.
- Piazza, M., DeFeo, V., Panzeri, S., Dehaene, S. (2018), « Learning to focus on number », *Cognition* , sous presse.
- Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., ..., Zorzi, M. (2010), « Developmental trajectory of number acuity reveals a

- severe impairment in developmental dyscalculia », *Cognition* , 116 (1), 33-41.
- Piazza, M., Pica, P., Izard, V., Spelke, E. S., Dehaene, S. (2013), « Education enhances the acuity of the nonverbal approximate number system », *Psychological Science* , 24 (6), 1037-1043.
- Pica, P., Lemer, C., Izard, V., Dehaene, S. (2004), « Exact and approximate arithmetic in an Amazonian indigene group », *Science* , 306 (5695), 499-503.
- Pierce, L. J., Klein, D., Chen, J.-K., Delcenserie, A., Genesee, F. (2014), « Mapping the unconscious maintenance of a lost first language », *Proc. Natl Acad. Sci. USA* , 111 (48), 17314-17319.
- Pinheiro-Chagas, P., Dotan, D., Piazza, M., Dehaene, S. (2017), « Finger tracking reveals the covert stages of mental arithmetic », *Open Mind* , 1 (1), 30-41.
- Pittenger, C., Kandel, E. R. (2003), « In search of general mechanisms for long-lasting plasticity : Aplysia and the hippocampus », *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* , 358 (1432), 757-763.
- Poirel, N., Borst, G., Simon, G., Rossi, S., Cassotti, M., Pineau, A., Houdé, O. (2012), « Number conservation is related to children's prefrontal inhibitory control : An fMRI study of a piagetian task », *PLoS ONE* , 7 (7), e40802.
- Poo, M.-M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., ..., Stevens, C. (2016), « What is memory ? The present state of the engram », *BMC Biology* , 14, 40.
- Posner, M. I. (1994), « Attention : The mechanisms of consciousness », *Proc. Natl Acad. Sci. USA* , 91, 7398-7403.
- Posner, M. I., Rothbart, M. K. (1998), « Attention, self-regulation and consciousness », *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* , 353 (1377), 1915-1927.

- Prado, E. L., Dewey, K. G. (2014), « Nutrition and brain development in early life », *Nutrition Reviews* , 72 (4), 267-284.
- Prehn-Kristensen, A., Munz, M., Göder, R., Wilhelm, I., Korr, K., Vahl, W., ..., Baving, L. (2014), « Transcranial oscillatory direct current stimulation during sleep improves declarative memory consolidation in children with attention-deficit/hyperactivity disorder to a level comparable to healthy controls », *Brain Stimulation* , 7 (6), 793-799.
- Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., Menon, V. (2014), « Hippocampal-neocortical functional reorganization underlies children's cognitive development », *Nature Neuroscience* , 17 (9), 1263-1269.
- Quartz, S. R., Sejnowski, T. J. (1997), « The neural basis of cognitive development : A constructivist manifesto », *Behav. Brain Sci.* , 20 (4), 537-556 ; discussion, 556-596.
- Rakic, P., Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N., Goldman-Rakic, P. S. (1986), « Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex », *Science* , 232 (4747), 232-235.
- Ramanathan, D. S., Gulati, T., Ganguly, K. (2015), « Sleep-dependent reactivation of ensembles in motor cortex promotes skill consolidation », *PLoS Biology* , 13 (9), e1002263.
- Ramirez, S., Liu, X., Lin, P.-A., Suh, J., Pignatelli, M., Redondo, R. L., ..., Tonegawa, S. (2013), « Creating a false memory in the hippocampus », *Science* , 341 (6144), 387-391.
- Ramirez, S., Liu, X., MacDonald, C. J., Moffa, A., Zhou, J., Redondo, R. L., Tonegawa, S. (2015), « Activating positive memory engrams suppresses depression-like behaviour », *Nature* , 522 (7556), 335-339.
- Rankin, C. H. (2004), « Invertebrate learning : What can't a worm learn ? », *Current Biology* , 14 (15), R617-618.
- Rasch, B., Büchel, C., Gais, S., Born, J. (2007), « Odor cues during slow-wave sleep prompt declarative memory consolidation », *Science* , 315 (5817), 1426-1429.



- Rasmussen, A., Jirenhed, D. A., Hesslow, G. (2008), « Simple and complex spike firing patterns in Purkinje cells during classical conditioning », *Cerebellum* , 7 (4), 563-566.
- Rattan, A., Savani, K., Chugh, D., Dweck, C. S. (2015), « Leveraging mindsets to promote academic achievement : Policy recommendations », *Perspectives on Psychological Science* , 10 (6), 721-726.
- Reich, L., Szwed, M., Cohen, L., Amedi, A. (2011), « A ventral visual stream reading center independent of visual experience », *Current Biology* , 21 (5), 363-368.
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., Reissland, N. (2017), « The human fetus preferentially engages with face-like visual stimuli », *Current Biology* , 27 (12), 1825-1828.e3.
- Rescorla, R. A., Wagner, A. R. (1972), « A theory of Pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement », *Classical Conditioning II : Current Research and Theory* , 2, 64-99.
- Ribeiro, S., Goyal, V., Mello, C. V., Pavlides, C. (1999), « Brain gene expression during REM sleep depends on prior waking experience », *Learning and Memory* , 6 (5), 500-508.
- Ritchie, S. J., Tucker-Drob, E. (2017), « How much does education improve intelligence ? A meta-analysis », PsyArXiv Preprints, <https://psyarxiv.com/-kymhp/> .
- Rivera, S. M., Reiss, A. L., Eckert, M. A., Menon, V. (2005), « Developmental changes in mental arithmetic : Evidence for increased functional specialization in the left inferior parietal cortex », *Cereb. Cortex*, 15 (11), 1779-1790.
- Robey, A. M., Dougherty, M. R., Buttaccio, D. R. (2017), « Making retrospective confidence judgments improves learners' ability to decide what not to study », *Psychological Science* , 28 (11).
- Roediger, H. L., Karpicke, J. D. (2006), « Test-enhanced learning : Taking memory tests improves long-term retention », *Psychological Science* , 17

(3), 249-255.

- Rohrer, D., Taylor, K. (2006), « The effects of overlearning and distributed practise on the retention of mathematics knowledge », *Applied Cognitive Psychology* , 20 (9), 1209-1224.
- Rohrer, D., Taylor, K. (2007), « The shuffling of mathematics problems improves learning », *Instructional Science* , 35 (6), 481-498.
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., Gabrieli, J. D. E. (2018), « Beyond the 30-million-word gap : Children's conversational exposure is associated with language-related brain function », *Psychological Science* , 29 (5), 700-710.
- Rouault, M., Koechlin, E. (2018), « Prefrontal function and cognitive control : From action to language », *Current Opinion in Behavioral Sciences* , 21, 106-111.
- Rudoy, J. D., Voss, J. L., Westerberg, C. E., Paller, K. A. (2009), « Strengthening individual memories by reactivating them during sleep », *Science* , 326 (5956), 1079.
- Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W. J., Bick, A., Frost, S. J., ..., Frost, R. (2015), « Universal brain signature of proficient reading : Evidence from four contrasting languages », *Proc. Natl Acad. Sci. USA* , 112 (50), 15510-15515.
- Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., Posner, M. I. (2005), « Training, maturation, and genetic influences on the development of executive attention », *Proc. Natl Acad. Sci. USA* , 102 (41), 14931-14936.
- Rugani, R., Fontanari, L., Simoni, E., Regolin, L., Vallortigara, G. (2009), « Arithmetic in newborn chicks », *Proc. Biol. Sci.* , 276 (1666), 2451-2460.
- Rugani, R., Vallortigara, G., Priftis, K., Regolin, L. (2015), « Animal cognition. Number-space mapping in the newborn chick resembles humans' mental number line », *Science* , 347 (6221), 534-536.

- Sabbah, N., Authié, C. N., Sanda, N., Mohand-Saïd, S., Sahel, J.-A., Safran, A. B., ..., Amedi, A. (2016), « Increased functional connectivity between language and visually deprived areas in late and partial blindness », *NeuroImage* , 136, 162-173.
- Sackur, J., Dehaene, S. (2009), « The cognitive architecture for chaining of two mental operations », *Cognition* , 111 (2), 187-211.
- Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., ..., Batista, A. P. (2014), « Neural constraints on learning », *Nature* , 512 (7515), 423-426.
- Saffran, J. R., Aslin, R. N., Newport, E. L. (1996), « Statistical learning by 8-month-old infants », *Science* , 274 (5294), 1926-1928.
- Sakai, T., Mikami, A., Tomonaga, M., Matsui, M., Suzuki, J., Hamada, Y., ..., Matsuzawa, T. (2011), « Differential prefrontal white matter development in chimpanzees and humans », *Current Biology* , 21 (16), 1397-1402.
- Salimpoor, V. N., Van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., Zatorre, R. J. (2013), « Interactions between the nucleus accumbens and auditory cortices predict music reward value », *Science* , 340 (6129), 216-219.
- Samson, D. R., Nunn, C. L. (2015), « Sleep intensity and the evolution of human cognition », *Evolutionary Anthropology* , 24 (6), 225-237.
- Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A., Schonon, S. de (2005), « Reversibility of the other-race effect in face recognition during childhood », *Psychological Science* , 16 (6), 440-444.
- Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., ..., Gabrieli, J. D. E. (2013), « Tracking the roots of reading ability : White matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children », *J. Neurosci.* , 33 (33), 13251-13258.
- Saygin, Z. M., Osher, D. E., Koldewyn, K., Reynolds, G., Gabrieli, J. D., Saxe, R. R. (2012), « Anatomical connectivity patterns predict face selectivity in

- the fusiform gyrus », *Nature Neuroscience* , 15 (2), 321-327.
- Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., ..., Kanwisher, N. (2016), « Connectivity precedes function in the development of the visual word form area », *Nature Neuroscience* , 19 (9), 1250-1255.
- Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., Botvinick, M. M. (2016), « Statistical learning of temporal community structure in the hippocampus », *Hippocampus* , 26 (1), 3-8.
- Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F., Steinmetz, H. (1995), « Increased corpus callosum size in musicians », *Neuropsychologia* , 33 (8), 1047-1055.
- Schmidt, R. A., Bjork, R. A. (1992), « New conceptualizations of practice : Common principles in three paradigms suggest new concepts for training », *Psychological Science* , 3 (4), 207-217.
- Schoenemann, P. T., Sheehan, M. J., Glotzer, L. D. (2005), « Prefrontal white matter volume is disproportionately larger in humans than in other primates », *Nature Neuroscience* , 8 (2), 242-252.
- Schultz, W., Dayan, P., Montague, P. R. (1997), « A neural substrate of prediction and reward », *Science* , 275, 1593-1599.
- Schweinhart, L. J. (1993), *Significant Benefits : The High/Scope Perry Preschool Study through Age 27. Monographs of the High/Scope Educational Research Foundation, No. Ten* , High/Scope Educational Research Foundation, ERIC.
- Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., Madsen, J. R. (2003), « Theta and gamma oscillations during encoding predict subsequent recall », *J. Neurosci.* , 23 (34), 10809-10814.
- Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., McCarthy, D. C., Brandt, A., ..., Kahana, M. J. (2006), « Hippocampal and neocortical gamma oscillations predict memory formation in humans », *Cereb. Cortex*, 17 (5), 1190-1196.

- Seehagen, S., Konrad, C., Herbert, J. S., Schneider, S. (2015), « Timely sleep facilitates declarative memory consolidation in infants », *Proc. Natl Acad. Sci.* , 112 (5), 1625-1629.
- Seitz, A., Lefebvre, C., Watanabe, T., Jolicœur, P. (2005), « Requirement for high-level processing in subliminal learning », *Current Biology* , 15 (18), R753-755.
- Senghas, A., Kita, S., Özyürek, A. (2004), « Children creating core properties of language : Evidence from an emerging sign language in Nicaragua », *Science* , 305 (5691), 1779-1782.
- Shah, P. E., Weeks, H. M., Richards, B., Kaciroti, N. (2018), « Early childhood curiosity and kindergarten reading and math academic achievement », *Pediatric Research* , 1.
- Shatz, C. J. (1996), « Emergence of order in visual system development », *Proc. Natl Acad. Sci. USA* , 93 (2), 602-608.
- Sheese, B. E., Rothbart, M. K., Posner, M. I., White, L. K., Fraundorf, S. H. (2008), « Executive attention and self-regulation in infancy », *Infant. Behav. Dev.* , 31 (3), 501-510.
- Sheridan, M. A., Fox, N. A., Zeanah, C. H., McLaughlin, K. A., Nelson, C. A. (2012), « Variation in neural development as a result of exposure to institutionalization early in childhood », *Proc. Natl Acad. Sci. USA* , 109 (32), 12927-12932.
- Shi, R., Lepage, M. (2008), « The effect of functional morphemes on word segmentation in preverbal infants », *Dev. Sci.* , 11 (3), 407-413.
- Shipston-Sharman, O., Solanka, L., Nolan, M. F. (2016), « Continuous attractor network models of grid cell firing based on excitatory-inhibitory interactions », *Journal of Physiology* , 594 (22), 6547-6557.
- Shneidman, L. A., Arroyo, M. E., Levine, S. C., Goldin-Meadow, S. (2013), « What counts as effective input for word learning ? », *Journal of Child Language* , 40 (3), 672-686.

- Shneidman, L. A., Goldin-Meadow, S. (2012), « Language input and acquisition in a Mayan village : How important is directed speech ? », *Dev. Sci.* , 15 (5), 659-673.
- Shohamy, D., Turk-Browne, N. B. (2013), « Mechanisms for widespread hippocampal involvement in cognition », *Journal of Experimental Psychology : General* , 142 (4), 1159-1170.
- Siegler, R. S. (1989), « Mechanisms of cognitive development », *Annu. Rev. Psychol.* , 40, 353-379.
- Siegler, R. S., Opfer, J. E. (2003), « The development of numerical estimation : Evidence for multiple representations of numerical quantity », *Psychological Science*, 14 (3), 237-243.
- Siegler, R. S., Thompson, C. A., Schneider, M. (2011), « An integrated theory of whole number and fractions development », *Cognitive Psychology* , 62 (4), 273-296.
- Sigman, M., Dehaene, S. (2008), « Brain mechanisms of serial and parallel processing during dual-task performance », *J. Neurosci.*, 28 (30), 7585-7598.
- Sigman, M., Pan, H., Yang, Y., Stern, E., Silbersweig, D., Gilbert, C. D. (2005), « Top-down reorganization of activity in the visual pathway after learning a shape identification task », *Neuron* , 46 (5), 823-835.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van den Driessche, G., ..., Hassabis, D. (2016), « Mastering the game of Go with deep neural networks and tree search », *Nature* , 529 (7587), 484-489.
- Simons, D. J., Chabris, C. F. (1999), « Gorillas in our midst : Sustained inattention blindness for dynamic events », *Perception* , 28 (9), 1059-1074.
- Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., Macnamara, B. N. (2018), « To what extent and under which circumstances are growth mindsets important to academic achievement ? Two meta-analyses », *Psychological Science* , 29 (4), 549-571.

- Skaggs, W. E., McNaughton, B. L. (1996), « Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience », *Science*, 271 (5257), 1870-1873.
- Smaers, J. B., Gómez-Robles, A., Parks, A. N., Sherwood, C. C. (2017), « Exceptional evolutionary expansion of prefrontal cortex in great apes and humans », *Current Biology* , 27 (5), 714-720.
- Spelke, E. (2003), « What makes us smart ? Core knowledge and natural language », in D. Gentner, S. Goldin-Meadow (éd.), *Language in Mind* , Cambridge (Mass.), MIT Press.
- Spencer, S. J., Steele, C. M., Quinn, D. M. (1999), « Stereotype threat and women's math performance », *Journal of Experimental Social Psychology* , 35 (1), 4-28.
- Spencer-Smith, M., Klingberg, T. (2015), « Benefits of a working memory training program for inattention in daily life : A systematic review and meta-analysis », *PLoS ONE* , 10 (3), e0119522.
- Srihasam, K., Mandeville, J. B., Morocz, I. A., Sullivan, K. J., Livingstone, M. S. (2012), « Behavioral and Anatomical consequences of early versus late symbol training in macaques », *Neuron* , 73 (3), 608-619.
- Stahl, A. E., Feigenson, L. (2015), « Observing the unexpected enhances infants' learning and exploration », *Science* , 348 (6230), 91-94.
- Starkey, P., Cooper, R. G. Jr (1980), « Perception of numbers by human infants », *Science* , 210, 1033-1035.
- Starkey, P., Spelke, E. S., Gelman, R. (1990), « Numerical abstraction by human infants », *Cognition* , 36, 97-127.
- Steele, C. M., Aronson, J. (1995), « Stereotype threat and the intellectual test performance of African Americans », *Journal of Personality and Social Psychology* , 69 (5), 797.
- Steinhauer, K., Drury, J. E. (2012), « On the early left-anterior negativity (ELAN) in syntax studies », *Brain and Language* , 120 (2), 135-162.

- Stickgold, R. (2005), « Sleep-dependent memory consolidation », *Nature* , 437 (7063), 1272-1278.
- Strauss, M., Sitt, J. D., King, J.-R., Elbaz, M., Azizi, L., Buiatti, M., ..., Dehaene, S. (2015), « Disruption of hierarchical predictive coding during sleep », *Proc. Natl Acad. Sci. USA* , 112 (11), E1353-1362.
- Striem-Amit, E., Amedi, A. (2014), « Visual cortex extrastriate body-selective area activation in congenitally blind people “seeing” by using sounds », *Current Biology* , 24 (6), 687-692.
- Strnad, L., Peelen, M. V., Bedny, M., Caramazza, A. (2013), « Multivoxel pattern analysis reveals auditory motion information in MT+ of both congenitally blind and sighted individuals », *PLoS ONE* , 8 (4), e63198.
- Sun, T., Patoine, C., Abu-Khalil, A., Visvader, J., Sum, E., Cherry, T. J., ..., Walsh, C. A. (2005), « Early asymmetry of gene transcription in embryonic human left and right cerebral cortex », *Science* , 308 (5729), 1794-1798.
- Sun, Z. Y., Klöppel, S., Rivière, D., Perrot, M., Frackowiak, R., Siebner, H., Mangin, J.-F. (2012), « The effect of handedness on the shape of the central sulcus », *NeuroImage* , 60 (1), 332-339.
- Sur, M., Garraghty, P. E., Roe, A. W. (1988), « Experimentally induced visual projections into auditory thalamus and cortex », *Science* , 242 (4884), 1437-1441.
- Sur, M., Rubenstein, J. L. R. (2005), « Patterning and plasticity of the cerebral cortex », *Science* , 310 (5749), 805-810.
- Sutton, R. S., Barto, A. G. (1998), *Reinforcement Learning : An Introduction* , Cambridge (Mass.), MIT Press.
- Szpunar, K. K., Khan, N. Y., Schacter, D. L. (2013), « Interpolated memory tests reduce mind wandering and improve learning of online lectures », *Proc. Natl Acad. Sci. USA* , 110 (16), 6313-6317.
- Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabregue, R., Amadon, A., Cohen, L. (2011), « Specialization for written words over objects in the



- visual cortex », *NeuroImage* , 56 (1), 330-344.
- Szwed, M., Qiao, E., Jobert, A., Dehaene, S., Cohen, L. (2014), « Effects of literacy in early visual and occipitotemporal areas of Chinese and French readers », *J. Cogn. Neurosci.*, 26 (3), 459-475.
- Szwed, M., Ventura, P., Querido, L., Cohen, L., Dehaene, S. (2012), « Reading acquisition enhances an early visual process of contour integration », *Dev. Sci.* , 15 (1), 139-149.
- Takeuchi, T., Duzkiewicz, A. J., Morris, R. G. M. (2014), « The synaptic plasticity and memory hypothesis : Encoding, storage and persistence », *Philos. Trans. R. Soc. B. Biol. Sci.* , 369 (1633).
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., Goodman, N. D. (2011), « How to grow a mind : Statistics, structure, and abstraction », *Science* , 331 (6022), 1279-1285.
- Terrace, H. S., Petitto, L. A., Sanders, R. J., Bever, T. G. (1979), « Can an ape create a sentence ? », *Science* , 206 (4421), 891-902.
- Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., Dehaene, S. (2014), « Learning to read improves the structure of the arcuate fasciculus », *Cereb. Cortex* , 24 (4), 989-995.
- Thornton, A., McAuliffe, K. (2006), « Teaching in wild meerkats », *Science* , 313 (5784), 227-229.
- Todorovic, A., de Lange, F. P. (2012), « Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields », *J. Neurosci.*, 32 (39), 13389-13395.
- Tombu, M., Jolicœur, P. (2004), « Virtually no evidence for virtually perfect time-sharing », *J. Exp. Psychol. Hum. Percept. Perform.* , 30 (5), 795-810.
- Uhrig, L., Dehaene, S., Jarraya, B. (2014), « A hierarchy of responses to auditory regularities in the macaque brain », *J. Neurosci.* , 34 (4), 1127-1132.

- van Praag, H., Kempermann, G., Gage, F. H. (2000), « Neural consequences of environmental enrichment », *Nature Reviews. Neuroscience* , 1 (3), 191-198.
- van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., Roelfsema, P. R. (2018), « The threshold for conscious report : Signal loss and response bias in visual and frontal cortex », *Science* .
- Ventura, P., Fernandes, T., Cohen, L., Morais, J., Kolinsky, R., Dehaene, S. (2013), « Literacy acquisition reduces the influence of automatic holistic processing of faces and houses », *Neuroscience Letters* , 554, 105-109.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., Cohen, L. (2007), « Hierarchical coding of letter strings in the ventral stream : Dissecting the inner organization of the visual word-form system », *Neuron* , 55 (1), 143-156.
- Vinckier, F., Naccache, L., Papeix, C., Forget, J., Hahn-Barma, V., Dehaene, S., Cohen, L. (2006), « “What” and “where” in word reading : Ventral coding of written words revealed by parietal atrophy », *J. Cogn. Neurosci.*, 18 (12), 1998-2012.
- Viswanathan, P., Nieder, A. (2013), « Neuronal correlates of a visual “sense of number” in primate parietal and prefrontal cortices », *Proc. Natl Acad. Sci. USA* , 110 (27), 11187-11192.
- Viswanathan, P., Nieder, A. (2015), « Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons », *Current Biology* , 25 (10), 1259-1269.
- Vogel, E. K., Machizawa, M. G. (2004), « Neural activity predicts individual differences in visual working memory capacity », *Nature* , 428 (6984), 748-751.
- Voss, M. W., Vivar, C., Kramer, A. F., van Praag, H. (2013), « Bridging animal and human models of exercise-induced brain plasticity », *Trends Cogn. Sci.* , 17 (10), 525-544.
- Waelti, P., Dickinson, A., Schultz, W. (2001), « Dopamine responses comply with basic assumptions of formal learning theory », *Nature* , 412 (6842), 43-

48.

- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., ..., Buckner, R. L. (1998), « Building memories : Remembering and forgetting of verbal experiences as predicted by brain activity », *Science* , 281 (5380), 1188-1191.
- Wagner, U., Gais, S., Haider, H., Verleger, R., Born, J. (2004), « Sleep inspires insight », *Nature* , 427 (6972), 352-355.
- Walker, M. P., Brakefield, T., Hobson, J. A., Stickgold, R. (2003), « Dissociable stages of human memory consolidation and reconsolidation », *Nature* , 425 (6958), 616-620.
- Walker, M. P., Stickgold, R. (2004), « Sleep-dependent learning and memory consolidation », *Neuron* , 44 (1), 121-133.
- Walker, M. P., Stickgold, R., Alsup, D., Gaab, N., Schlaug, G. (2005), « Sleep-dependent motor memory plasticity in the human brain », *Neuroscience* , 133 (4), 911-917.
- Walker, S. P., Chang, S. M., Powell, C. A., Grantham-McGregor, S. M. (2005), « Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children : Prospective cohort study », *The Lancet* , 366 (9499), 1804-1807.
- Wang, L., Krauzlis, R. J. (2018), « Visual selective attention in mice », *Current Biology* , 28 (5), 676-685.e4.
- Wang, L., Uhrig, L., Jarraya, B., Dehaene, S. (2015), « Representation of numerical and sequential patterns in macaque and human brains », *Current Biology* , 25 (15), 1966-1974.
- Warneken, F., Tomasello, M. (2006), « Altruistic helping in human infants and young chimpanzees », *Science* , 311 (5765), 1301-1303.
- Watanabe, T., Nanez, J. E., Sasaki, Y. (2001), « Perceptual learning without perception », *Nature* , 413 (6858), 844-848.

- Weber-Fox, C. M., Neville, H. J. (1996), « Maturational constraints on functional specialization for language processing : ERP and behavioral evidence in bilingual speakers », *J. Cogn. Neurosci.*, 8, 231-256.
- Werker, J. F., Hensch, T. K. (2014), « Critical periods in speech perception : New directions », *Annu. Rev. Psychol.* , 66, 173-196.
- Werker, J. F., Tees, R. C. (1984), « Cross-language speech perception : Evidence for perceptual reorganization during the first year of life », *Infant Behavior and Development* , 7, 49-63.
- Whitlock, J. R., Heynen, A. J., Shuler, M. G., Bear, M. F. (2006), « Learning induces long-term potentiation in the hippocampus », *Science* , 313 (5790), 1093-1097.
- Widloski, J., Fiete, I. R. (2014), « A model of grid cell development through spatial exploration and spike time-dependent plasticity », *Neuron* , 83 (2), 481-495.
- Wilhelm, I., Rose, M., Imhof, K. I., Rasch, B., Büchel, C., Born, J. (2013), « The sleeping child outplays the adult's capacity to convert implicit into explicit knowledge », *Nature Neuroscience* , 16 (4), 391-393.
- Wills, T. J., Cacucci, F., Burgess, N., O'Keefe, J. (2010), « Development of the hippocampal cognitive map in preweanling rats », *Science* , 328 (5985), 1573-1576.
- Wilson, M. A., McNaughton, B. L. (1994), « Reactivation of hippocampal ensemble memories during sleep », *Science* , 265 (5172), 676-679.
- Windsor, J., Moraru, A., Nelson, C. A., Fox, N. A., Zeanah, C. H. (2013), « Effect of foster care on language learning at eight years : Findings from the Bucharest Early Intervention Project », *Journal of Child Language* , 40 (3), 605-627.
- Wynn, K. (1992), « Addition and subtraction by human infants », *Nature* , 358, 749-750.

- Xu, F., Garcia, V. (2008), « Intuitive statistics by 8-month-old infants », *Proc. Natl Acad. Sci. USA* , 105 (13), 5012-5015.
- Xu, F., Tenenbaum, J. B. (2007), « Word learning as Bayesian inference », *Psychol. Rev.* , 114 (2), 245-272.
- Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhutdinov, R., ..., Bengio, Y. (2015), « Show, attend and tell : Neural image caption generation with visual attention », <https://arxiv.org/abs/1502.03044> .
- Yang, C. (2013), « Ontogeny and phylogeny of language », *Proc. Natl Acad. Sci. USA* , 110 (16), 6324-6327.
- Yoncheva, Y. N., Blau, V. C., Maurer, U., McCandliss, B. D. (2010), « Attentional focus during learning impacts N170 ERP responses to an artificial script », *Dev. Neuropsychol.* , 35 (4), 423-445.
- Yoon, J. M. D., Johnson, M. H., Csibra, G. (2008), « Communication-induced memory biases in preverbal infants », *Proc. Natl Acad. Sci. USA* , 105 (36), 13690-13695.
- Yoon, K., Buice, M. A., Barry, C., Hayman, R., Burgess, N., Fiete, I. R. (2013), « Specific evidence of low-dimensional continuous attractor dynamics in grid cells », *Nature Neuroscience* , 16 (8), 1077-1084.
- Young, C. B., Wu, S. S., Menon, V. (2012), « The neurodevelopmental basis of math anxiety », *Psychological Science* , 23 (5), 492-501.
- Zaromb, F. M., Karpicke, J. D., Roediger, H. L. (2010), « Comprehension as a basis for metacognitive judgments : Effects of effort after meaning on recall and metacognition », *J. Exp. Psychol. Learn. Mem. Cogn.* , 36 (2), 552-557.
- Zaromb, F. M., Roediger, H. L. I. (2010), « The testing effect in free recall is associated with enhanced organizational processes », *Memory and Cognition* , 38 (8), 995-1008.
- Zhu, X., Wang, F., Hu, H., Sun, X., Kilgard, M. P., Merzenich, M. M., Zhou, X. (2014), « Environmental acoustic enrichment promotes recovery from

developmentally degraded auditory cortical processing », *The J. Neurosci.* , 34 (16), 5406-5415.

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., Spinelli, D. (2005), « Word length effect in early reading and in developmental dyslexia », *Brain and Language* , 93 (3), 369-373.

Zylberberg, A., Dehaene, S., Roelfsema, P. R., Sigman, M. (2011), « The human Turing machine : A neural framework for mental programs », *Trends Cogn. Sci.* , 15 (7), 293-300.

# Remerciements

---

La maturation de ce livre a été stimulée par de nombreuses rencontres. Dès les années 1990, Michael Posner et Bruce McCandliss, de l'Université d'Oregon, m'ont convaincu que les sciences cognitives pouvaient être pertinentes dans le domaine éducatif. Je dois beaucoup aux nombreuses rencontres scientifiques qu'ils ont organisées avec l'aide de l'OCDE et de Bruno Della Chiesa. Par la suite, Marcela Peña, Sidarta Ribeiro, Mariano Sigman, Alejandro Maiche et Juan Valle Lisboa ont formé une génération entière de jeunes scientifiques lors des inoubliables rencontres annuelles de la Latin American School for Education, Cognitive and Neural Sciences. J'ai eu la chance de participer à chacune d'elles, et je leur dois beaucoup.

Une autre personne qui a eu cette chance est mon épouse et collègue Ghislaine Dehaene-Lambertz. Cela fait trente-deux ans que nous discutons de développement cérébral et, accessoirement, d'éducation de nos enfants. Il va de soi que je lui dois tout, y compris une relecture attentive des pages qui précèdent.

Autre anniversaire : cela fait trente-trois ans que j'ai rejoint les laboratoires de Jacques Mehler et de Jean-Pierre Changeux. Leur influence sur ma pensée est immense, et ils reconnaîtront ici bon nombre de leurs thèmes de prédilection – comme le feront d'autres collègues et amis très proches tels que Lucia Braga, Laurent Cohen, Naama Friedmann, Véronique Izard, Régine Kolinsky, José

Morais, Lionel Naccache, Christophe Pallier, Mariano Sigman, Elizabeth Spelke et Josh Tenenbaum.

Mon ami Antonio Battro m'a constamment encouragé dans mes recherches sur le cerveau, l'éducation et l'apprentissage. Je lui suis également reconnaissant de m'avoir fait rencontrer Nico, un artiste à la personnalité remarquable, qui m'a très gentiment autorisé à reproduire ici quelques-uns de ses tableaux. Merci également à Yoshua Bengio, Alain Chédotal, Guillaume et David Dehaene, Molly Dillon, Jessica Dubois, György Gergely, Eric Knudsen, Leah Krubitzer, Bruce McCandliss, Josh Tenenbaum, Fei Xu et Robert Zatorre pour m'avoir autorisé à reproduire les nombreuses figures qui émaillent ce texte.

Je voudrais également remercier toutes les institutions qui soutiennent mes recherches depuis des années avec une fidélité à toute épreuve, et notamment l'Inserm, le CEA, le Collège de France, l'université Paris-Sud, l'European Research Council et la Fondation Bettencourt-Schueller. C'est grâce à eux que j'ai pu m'entourer d'étudiants et de collaborateurs brillants et énergiques. Ils sont trop nombreux pour être tous cités ici, mais ils se reconnaîtront dans la longue liste de publications qui accompagne ce livre. Une mention particulière va à Anna Wilson, Dror Dotan et Cassandra Potier-Watkins, avec qui j'ai développé des logiciels éducatifs et des interventions dans les classes.

Jean-Michel Blanquer, ministre de l'Éducation nationale, m'a honoré de sa confiance en me proposant de présider son premier Conseil scientifique – qu'il en soit ici vivement remercié. Je suis reconnaissant à tous les membres du Conseil, et notamment Esther Duflo, Michel Fayol, Marc Gurgand, Caroline Huron, Elena Pasquinelli, Franck Ramus, Elizabeth Spelke et Jo Ziegler, ainsi qu'à son secrétaire général Nelson Vallejo-Gomez, pour leur engagement et pour tout ce qu'ils m'ont appris.

La préparation de cette édition a bénéficié d'une relecture attentive de Marie-Lorraine Colas.

Enfin, c'est le sixième livre que je publie avec mon éditrice Odile Jacob et son mari Bernard Gotlieb. Leur amitié et leur confiance me touchent profondément.



Termas El Corazon, Chili, juin 2018.

# Crédits iconographiques

---

Figure 2 bas , figure 8 , figure 13 haut gauche et milieu , figure 14 bas droite , figure 29 haut , figure 34 : © Stanislas Dehaene .

Figure 1 haut et bas : © avec l'aimable autorisation d'Antonio Battro .

Figure 3 haut : d'après <https://ai.googleblog.com/2017/05/using-machine-learning-to-explore.html> ; bas : d'après Olah, C., Mordvintsev, A. & Schubert, L., 2017 .

Figure 4 : d'après Guerguiev, J., Lillicrap, T. P. & Richards, B. A., 2017 , figure 6 .

Figure 5 : © avec l'aimable autorisation de Josh Tenenbaum ; d'après Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D., 2011 , figure 1 .

Figure 6 : d'après Kemp, C. & Tenenbaum, J. B., 2008 , figures 2 et 3 .

Figure 7 haut : © avec l'aimable autorisation de Fei Xu ; bas : © avec l'aimable autorisation de Molly Dillon et Elizabeth Spelke .

Figure 9 haut : © avec l'aimable autorisation de G. Dehaene-Lambertz et J. Dubois ; bas : d'après Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M. & Dehaene, S., 2006.

Figure 10 : © avec l'aimable autorisation de Leah Krubitzer ; d'après Krubitzer, L., 2007 .

Figure 11 haut : © avec l'aimable autorisation d'Alain Chédotal ; d'après Belle, M., Godefroy, D., Couly, G., Malone, S. A., Collier, F., Giacobini, P. & Chédotal, A., 2017 ; bas : © avec l'aimable autorisation de G. Dehaene-Lambertz et J. Dubois .

Figure 12 : d'après Amunts, K., Lenzen, M., Friederici, A. D., Schleicher, A., Morosan, P., Palomero-Gallagher, N. & Zilles, K., 2010 , figures 1 et 7 .

Figure 13 haut droite : d'après <https://www.pexels.com/photo/animals-apiary-beehive-beekeeping-928978> ; bas : d'après Hafting, T., Fyhn, M., Molden, S., Moser, M. B. & Moser, E. I., 2005 .

Figure 14 gauche : © Aisa/Leemage ; haut droite : © avec l'aimable autorisation de Philip Buttery .

Figure 15 : d'après Muckli, L., Naumer, M. J. & Singer, W., 2009 , figure 2 .

Figure 16 : © 1939, 1941, 1947, 1951, 1955, 1959, 1963, 1967, by the President and Fellows of Harvard College. © renewed 1967, 1969, 1975, 1979, 1983, 1987, 1991 .

Figure 17 haut : d'après Flege J. E., Munro, M. J. & MacKay, I. R., 1995 ; Johnson, J. S. & Newport, E. L., 1989 ; Hartshorne, J. K. *et al* ., 2018 ; bas : d'après Pierce, L. J., Klein, D., Chen, J.-K., Delcenserie, A. & Genesee, F., 2014 , figure 3 .

Figure 18 haut : © avec l'aimable autorisation d'Eric Knudsen ; bas : d'après Knudsen, E. I., Zheng, W. & DeBello, W. M., 2000 , figures 2 et 3 .

Figure 19 haut : © Michael Carroll, d'après Berens, A. E. & Nelson, C. A., 2015 ; bas : d'après Almas, A. N., Degnan, K. A., Radulescu, A., Nelson, C. A., Zeanah, C. H. & Fox, N. A., 2012 , figure 1 .

Figure 20 : d'après Amalric, M. & Dehaene, S., 2016 .

Figure 21 bas : d'après Amalric, M., Denghien, I. & Dehaene, S., 2017 .

Figure 22 : d'après Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ..., Cohen, L., 2010 .

Figure 23 : d'après Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ..., Cohen, L., 2010 .

Figure 24 haut : © avec l'aimable autorisation de G. Dehaene-Lambertz ; bas : d'après Monzalvo, K., Fluss, J., Billard, C., Dehaene, S. & Dehaene-Lambertz,

G., 2012 .

Figure 25 : d'après Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S., 2018 .

Figure 26 : d'après Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhutdinov, R., ..., Bengio, Y. 2015 .

Figure 27 bas : d'après Kilgard, M. P. & Merzenich, M. M., 1998 .

Figure 28 : © avec l'aimable autorisation de Bruce McCandliss ; d'après Yoncheva, Y. N., Blau, V. C., Maurer, U. & McCandliss, B. D., 2010 .

Figure 29 bas : © avec l'aimable autorisation de Robert Zatorre, d'après Bermudez, P., Lerch, J. P., Evans, A. C. & Zatorre, R. J., 2009 .

Figure 30 : © avec l'aimable autorisation de Gyorgy Gergely ; haut : d'après Egyed, K., Király, I. & Gergely, G., 2013 ; bas : d'après Gergely, G., Bekkering, H. & Kiraly, I., 2002 .

Figure 31 : d'après Held, R. & Hein, A., 1963 , figure 1 .

Figure 32 : d'après Kaplan, F. & Oudeyer, P.-Y., 2007 , figure 3 .

Figure 33 : d'après Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L. & Naccache, L., 2009 et Strauss, M., Sitt, J. D., King, J.-R., Elbaz, M., Azizi, L., Buiatti, M., ..., Dehaene, S., 2015 .

Figure 35 gauche : d'après Dehaene-Lambertz, G., Monzalvo, K. & Dehaene, S., 2018 ; droite : d'après Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A. & Spinelli, D., 2005 .

Figure 36 : d'après Chen, Z. & Wilson, M. A., 2017 .

Les autres illustrations sont dans le domaine public et libres de droits.

# TABLE

---

## INTRODUCTION

*Pourquoi l'apprentissage ?*

*Homo docens*

*Apprendre à apprendre*

*Le défi des machines*

## Première partie - QU'EST-CE QU'APPRENDRE ?

### CHAPITRE 1 - Comment un réseau de neurones apprend

*Apprendre, c'est ajuster les paramètres d'un modèle mental*

*Apprendre, c'est exploiter le potentiel de la combinatoire*

*Apprendre, c'est minimiser ses erreurs*

*Apprendre, c'est explorer l'espace des possibles*

*Apprendre, c'est optimiser une fonction de récompense*

*Apprendre, c'est restreindre l'espace de recherche*

*Apprendre, c'est projeter des hypothèses a priori*

### CHAPITRE 2 - Pourquoi notre cerveau apprend mieux que les machines

*Ce qui manque encore à l'intelligence artificielle*

*Apprendre, c'est inférer la grammaire d'un domaine*

*Apprendre, c'est raisonner en bon statisticien*

## Deuxième partie - COMMENT NOTRE CERVEAU APPREND

### CHAPITRE 3 - Le savoir invisible : les étonnantes intuitions des bébés

*Le concept d'objet*

*Le sens du nombre*

*L'intuition des probabilités*

*La connaissance des animaux et des personnes*

*La perception des visages*

*Le don des langues*

### CHAPITRE 4 - Naissance d'un cerveau

*Le cerveau du bébé, déjà bien structuré*

*Les autoroutes du langage*

*L'auto-organisation du cortex*

### CHAPITRE 5 - La part de l'acquis

*Qu'est-ce que la plasticité cérébrale ?*

*Le portrait d'un souvenir*

*Vraies synapses et faux souvenirs*

*La nutrition, élément clé de l'apprentissage*

*Pouvoirs et limites de la plasticité synaptique*

*Qu'est-ce qu'une période sensible ?*

*Il faut qu'une synapse soit ouverte ou fermée*

*Un miracle à Bucarest*

## CHAPITRE 6 - Recyclez votre cerveau

*L'hypothèse du recyclage neuronal*

*Les mathématiques recyclent les circuits du nombre*

*La lecture recycle les circuits de la vision et du langage parlé*

*Partitions, équations et visages*

*Les bénéfices d'un environnement enrichi*

## Troisième partie - LES QUATRE PILIERS DE L'APPRENTISSAGE

### CHAPITRE 7 - L'attention

*Alerte : l'éveil du cerveau*

*Orientation : le filtre du cerveau*

*Contrôle exécutif : l'aiguilleur du cerveau*

*Apprendre à faire attention*

*Je fais attention si tu fais attention*

*Enseigner, c'est faire attention à l'attention de l'autre*

### CHAPITRE 8 - L'engagement actif

*Un organisme passif n'apprend pas*

*Approfondir pour mieux apprendre*

*L'échec des pédagogies de la découverte*

*Savoir piquer la curiosité*

*Vouloir savoir : le moteur de la motivation*

*Les trois façons dont l'école peut tuer la curiosité*

### CHAPITRE 9 - Le retour sur erreur

*La surprise, moteur de l'apprentissage*

*Le cerveau fourmille de messages d'erreur*

*Ne confondons pas erreur et sanction*

*La note, piètre retour sur erreur*

*Se tester pour mieux apprendre*

*Une règle d'or : espacer les apprentissages*

#### CHAPITRE 10 - La consolidation

*Libérer les ressources cérébrales*

*Le rôle clé du sommeil*

*Le cerveau endormi revit les épisodes de la veille*

*Peut-on faire des découvertes en dormant ?*

*Sommeil, enfance et école*

#### CONCLUSION - Réconcilier l'éducation avec les neurosciences

*Treize maximes pour l'épanouissement des enfants*

*Une alliance pour l'école de demain*

Notes

Bibliographie

Remerciements

Crédits iconographiques



# Éditions Odile Jacob

## Des idées qui font avancer les idées

Retrouvez tous nos livres disponibles en numérique  
sur [odilejacob.fr](http://odilejacob.fr)

Rejoignez-nous sur Facebook



Suivez-nous sur Twitter



# Stanislas Dehaene

## Apprendre !

### Les talents du cerveau, le défi des machines

« Notre cerveau possède, dès la naissance, un talent que les meilleurs logiciels d'intelligence artificielle ne parviennent pas encore à imiter : la faculté d'apprendre.

Même le cerveau d'un bébé apprend déjà plus vite et plus profondément que la plus puissante des machines actuelles. Et cette remarquable capacité d'apprentissage, l'humanité a découvert qu'elle pouvait encore l'augmenter grâce à une institution : l'école.

Au cours des trente dernières années, d'importants progrès ont été réalisés dans la compréhension des principes fondamentaux de la plasticité cérébrale et de l'apprentissage.

Il est temps que chaque enfant, chaque adulte prenne la pleine mesure du potentiel énorme de son propre cerveau – et aussi, bien sûr, de ses limites.

Le fonctionnement de la mémoire, le rôle de l'attention, l'importance du sommeil sont autant de découvertes riches de conséquences pour chacun d'entre nous. Des idées très simples sur le jeu, le plaisir, la curiosité, la socialisation, la concentration ou le sommeil peuvent augmenter encore ce qui est déjà le plus grand talent de notre cerveau : apprendre ! » S. D.



**Stanislas Dehaene** est professeur au Collège de France, titulaire de la chaire de psychologie cognitive expérimentale, membre de l'Académie des sciences. Il préside le Conseil scientifique de l'Éducation nationale. Il a publié *Les Neurones de la lecture*, *La Bosse des maths* et *Le Code de la conscience*, qui ont rencontré un très grand succès.