Sommaire

HISTOIRE DES NEUROSCIENCES

La place de la neurophysiologie dans les débuts de l'intelligence...
Par Jean-Gaël Barbara
Voir l'article

HISTOIRE DES NEUROSCIENCES

Histoire de la Société de Psychophysiologie et de Neurosciences Cognitives
Par Thomas Hinault, Henrique Sequeira, Aurélie Bidet-Caulet, Anne Caclin, Marie Gomot, Nathalie Georg
Giard, Marie Timsit-Be
Voir l'article

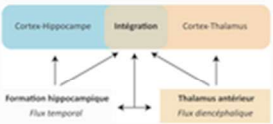
DOSSIER

Dialogues entre neurosciences et IA
Lire le dossier

TRIBUNE LIBRE - QUESTION D'ACTUALITÉ

Comment mener des recherches expérimentales en milieu scolaire ?
Par Sébastien Goudeau, Matthew J. Easterbrook, Marie-Pierre Fayant
Voir l'article

NOUVEAUTÉS EN NEUROSCIENCES



Le thalamus cognitif
Par Mathieu Wolff
Voir l'article

HOMMAGE

Hommage au Professeur Marie-Jo Besson
Par Jocelyne Caboche, Jean-Antoine Girault, Peter Vanhoutte, Emmanuel Valjent, Sylvie Retaux, Karen Brami-Cherrier, Marie-Lou Kemel
Voir l'article

Édito

L'Intelligence Artificielle (IA) est omniprésente dans notre quotidien. Elle effraie ou elle fascine selon les situations. Elle occupe une place particulière dans le domaine des neurosciences car d'un côté, les propriétés anatomiques et fonctionnelles du cerveau aident au développement de l'IA, et de l'autre, les méthodes issues de l'IA participent à une meilleure analyse ou modélisation de toute la complexité de l'anatomie et du fonctionnement du cerveau. Ne parle-t-on pas de neurones et de réseaux de neurones dans les deux cas ? Pour toutes ces raisons, nous avons choisi de consacrer une partie de ce numéro à la place de l'IA dans les neurosciences, à travers le dossier de ce numéro et la rubrique d'Histoire des neurosciences.

PAR YVES TILLET ET JEREMY DANNA

Aborder précisément les liens entre le l'IA et les neurosciences nécessiterait plusieurs ouvrages tant ces interactions sont nombreuses et diversifiées. Nous avons donc dû faire des choix et nous avons demandé aux experts de ces domaines d'aborder quelques exemples caractéristiques comme la place de l'IA dans les interactions cerveau-machines. Le contrôle du mouvement ne passe plus par une commande musculaire envoyée vers notre corps mais simplement par la pensée, ce qui était de la science-fiction dans les années 90 devient réalité. Les capacités de calcul de l'IA permettent d'analyser l'activité neuronale, d'induire des réponses adaptées aux différentes tâches, et d'apprendre. Ces travaux trouvent leurs applications dans la mise au point de neuroprothèses pour la restauration de la marche chez les tétraplégiques ou la récupération de mouvements après un accident. Au-delà des applications cliniques, sur un plan plus fondamental l'IA devient également un outil indispensable pour l'étude du cerveau. Vous lirez ainsi dans ce dossier comment des neurobiologistes et des informaticiens tourangeaux ont développé un nouvel outil d'analyse d'images obtenues en IRM sur différents modèles animaux. Les images sont analysées sans a priori anatomique. L'IA peut également être mise à profit dans l'étude et la modélisation des comportements sociaux tels que ceux décrits chez les fourmis et les termites. À l'aide de la modélisation mathématique, les chercheurs sont parvenus à décoder les interactions complexes entre les insectes, et notamment leur coordination lors de la construction du nid, incluant des comportements très sophistiqués. Un dossier à lire sans tarder.

Pour compléter ce dossier ne manquez pas l'article d'Histoire des Neurosciences dans lequel Jean-Gaël Barbara explique comment dans les années 50, à partir de quatre exemples aux États-Unis, en France, en Angleterre et au Canada, des cybernéticiens, des neurophysiologistes, des psychologues et des mathématiciens ont collaboré pour créer des dispositifs artificiels intelligents, capables de reproduire les capacités de l'intelligence humaine. On peut imaginer que les discussions n'ont pas toujours été simples entre les acteurs de ces différentes disciplines mais elles ont été fécondes et elles illustrent bien l'importance de l'interdisciplinarité.

D'interdisciplinarité il en est également question dans l'article de Thomas Hinault et son équipe qui nous expliquent comment la rencontre entre des

psychologues et des neurophysiologistes a donné naissance à la Société de Psychophysiologie et de Neurosciences Cognitives. Le but des chercheurs de cette Société, à savoir la compréhension du fonctionnement du cerveau et de ses capacités cognitives, était déjà au cœur des préoccupations des chercheurs en IA.

Dans la rubrique nouveauté, Mathieu Wolff nous livre les dernières données sur le thalamus, une structure longtemps délaissée car considérée comme un simple relais vers le cortex. Les plus récentes études montrent qu'en réalité le thalamus joue un rôle important – pour ne pas dire central – dans les fonctions intégratives et cognitives. Mathieu Wolff nous explique que le thalamus est impliqué notamment dans la mémoire, la flexibilité comportementale, la prise de décisions, autant de fonctions classiquement associées au cortex préfrontal. De même, un certain nombre de pathologies mentales sont associées à un dysfonctionnement thalamique. Toutes ces observations donnent une autre image du thalamus et son rôle mérite sans nul doute d'être revisité avec la considération que suggère sa position centrale dans le cerveau. Un article à ne surtout pas manquer.

Enfin, la recherche expérimentale sur les sujets humains soulève de nombreuses questions éthiques qui prennent peut-être encore plus d'importance dès lors qu'elles s'intéressent aux enfants. Ces recherches sont précieuses pour l'étude des mécanismes d'apprentissage et des processus cognitifs car elles servent aussi de références pour la construction et la mise en place des programmes scolaires par les pouvoirs publics. En regard de ces enjeux, Sébastien Goudeau et son équipe abordent un certain nombre de recommandations, notamment éthiques, pour faciliter et optimiser les collaborations entre les acteurs de cette recherche dans le milieu scolaire.

Nous espérons que cette Lettre saura retenir votre attention et assouvir votre curiosité, nous vous garantissons qu'elle a été écrite « avec I mais sans A ».

Bonne lecture.

Histoire des Neurosciences

La place de la neurophysiologie dans les débuts de l'intelligence artificielle au cours des années 1950

JEAN-GAËL BARBARA

Sorbonne Université, CNRS, Inserm, Institut de Biologie Paris Seine (IBPS), Centre de Neuroscience Sorbonne Université (NeuroSU), 75005, Paris, France.

Pendant la Seconde Guerre mondiale, puis lors de l'essor de la cybernétique dans les années 1950, le développement des recherches en intelligence artificielle s'est caractérisé par l'association assez naturelle avec les domaines de la neurophysiologie et de la psychologie, dans différents pays et dans différents contextes. On propose ici d'analyser quatre exemples de contextes parallèles, aux États-Unis, en France, en Angleterre, et au Canada, afin de montrer, dans chaque cas, comment la rencontre eut lieu et ce qu'elle put apporter soit aux deux domaines concernés, soit à l'un d'entre eux seulement.

Depuis les années 1950 les neurosciences ont progressivement établi des relations étroites avec le domaine de l'intelligence artificielle, avec comme conséquence la fondation d'un domaine de recherche interdisciplinaire fructueux, encore actuellement en plein développement (1). On se propose ici d'étudier les débuts de ce type de collaboration à travers quatre exemples concernant d'abord le fondateur de la cybernétique Norbert Wiener (1894-1964) aux États-Unis et sa collaboration très tôt avec les neurophysiologistes Warren S. McCulloch (1898-1969), Rafael Lorente de Nó (1902-1990), et William Grey Walter (1910-1977). Cet exemple permet de mieux comprendre l'intérêt réciproque qu'il y eut alors de rapprocher l'étude et la conception de dispositifs intelligents de visée antiaériens (missiles autoguidés pour abattre des avions ou des fusées) avec l'étude de l'intelligence humaine, telle qu'on commence alors à l'envisager, en opposition avec la psychologie behavioriste, par des modèles neurophysiologiques (c'est-à-dire des modèles qui prennent en compte la structure de circuits neuronaux locaux et les propriétés de propagation de l'influx nerveux).

Le deuxième exemple concerne le neurophysiologiste français Louis Lapicque (1866-1952) qui s'intéressa au début des années 1940 aux machines à calculer de son collègue mathématicien du CNRS Louis Couffignal (1902-1966), pour envisager l'étude des analogies entre de tels dispositifs et le cerveau humain, dont il imaginait des modèles neuronaux en partie mathématiques.

En trois, ce sont le psychologue Kenneth Craik (1914-1945) et le neurophysiologiste William Grey Walter (1910-1977), en Angleterre, qui imaginèrent des modèles électroniques d'apprentissage en vue de mieux comprendre les facultés psychiques, notamment celles dont l'étude était utile pour l'effort de guerre, et dans la perspective de s'en inspirer pour réaliser des « cerveaux électroniques ».

Enfin, le quatrième exemple est celui du psychologue canadien Donald Hebb (1904-1985) qui établit un modèle neuronal d'intelligence et d'apprentissage qui aiguisa la curiosité et l'intérêt de la société IBM dans le développement des premiers ordinateurs.

Norbert Wiener et son *AA predictor* en tant que modèle de cerveau humain

Au cours de la Seconde Guerre mondiale, le mathématicien américain de génie Norbert Wiener, alors âgé d'une quarantaine d'année, fut affecté dans une unité de recherche sur la défense antiaérienne. Il imagina la

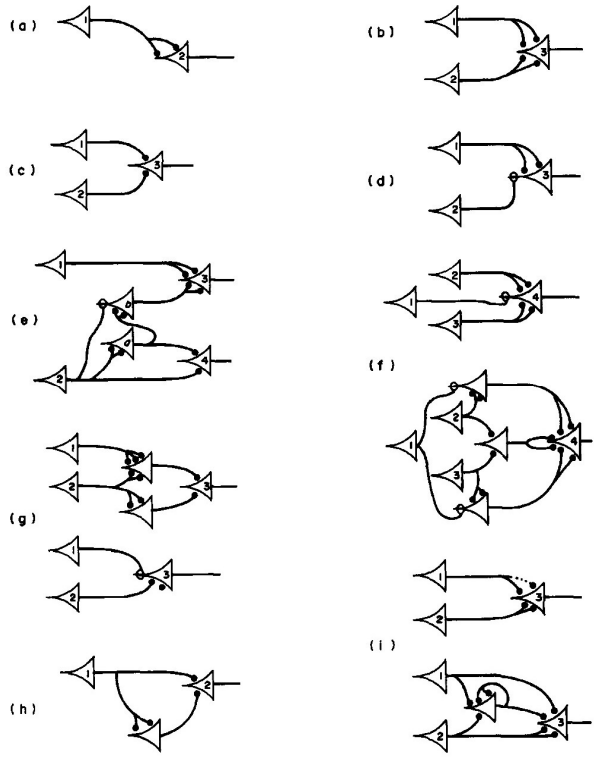
réalisation d'un dispositif électronique capable de prévoir la trajectoire d'un avion ou d'une fusée ennemis, en vue de l'abattre avec une batterie antiaérienne. Son dispositif était censé faire le calcul en quelques secondes et ordonner un tir d'une manière quasi immédiate automatique, sans action humaine. Si Wiener eut de réelles difficultés à mettre au point un tel dispositif et à convaincre ses supérieurs de la faisabilité rapide de son projet, Wiener se consola en imaginant que son projet représentait une nouvelle manière de concevoir l'intelligence. Sa conception fonctionnelle de l'intelligence avait l'intérêt d'être indépendante du dispositif évalué, qu'il soit artificiel ou vivant. Dès lors Wiener envisagea que sa conception d'un contrôle intelligent des comportements pouvait s'appliquer non seulement aux machines, mais également au comportement animal et à l'étude de la cognition humaine.

C'est ainsi que dès 1942, Wiener crée un club de discussion dans lequel il intégra des neurophysiologistes afin de mieux comprendre comment sa nouvelle manière de concevoir l'esprit humain – par analogie avec les dispositifs artificiels intelligents – pouvait être étayée par les connaissances les plus récentes portant sur le système nerveux. Puis de telles collaborations se poursuivirent dans le cadre des célèbres conférences interdisciplinaires de la fondation J. Macy.

Ces discussions aboutirent à l'idée que l'esprit humain pouvait être considéré, par analogie avec la machine, comme étant composé d'entités fonctionnelles ayant des fonctions psychologiques précises, comme par exemple la visée d'un objet, la reconnaissance d'une forme, le scan d'un paysage pour détecter la position d'un objet dans une scène. C'est là une idée fondamentale du cognitivisme qui émergea au cours des années 1950 et qui aboutit beaucoup plus tard par exemple à l'étude neurophysiologique de la reconnaissance des visages. Mais au début des années 1940, l'idée de tels modules était encore largement hypothétique, même si des aires fonctionnelles comme les aires du langage étaient déjà connues.

Le neurophysiologiste américain Lorente de Nó d'origine espagnole, du célèbre groupe des axonologistes, et l'un des derniers élèves de S. Ramón y Cajal, proposa lors des discussions avec Wiener qu'il était possible d'imaginer les bases anatomiques et physiologiques de tels modules en considérant ses travaux des années 1930 sur l'anatomie des circuits des noyaux oculo-moteurs. Il proposa que ces modules fonctionnels pouvaient reposer sur des « chaînes fermées de neurones internunciaux », c'est-à-dire des circuits locaux autonomes. De tels circuits pouvaient alors être vus comme capables de traitements et de stockage d'informations et comme analogues aux petits « cerveaux électroniques » qu'on implantait alors dans des fusées militaires. D'un point de vue épistémologique, il est intéressant de constater que les neurophysiologistes retournaient alors à une vision plutôt localisationniste du cerveau, alors qu'à la même époque la plupart des neurologues développaient des idées plutôt contraires, holistes, dans un contexte d'évaluation psychologique d'ablations corticales chez l'animal et de validation des techniques de psychochirurgie chez l'Homme.

Discutant également avec Wiener sur ces sujets, le neurophysiologiste et psychiatre Warren McCulloch envisagea alors d'établir des modèles logiques de neurones formels qu'il construisit sur le papier à la manière de circuits électroniques dédiés à des calculs logiques (Figure 1). Dans leur article publié en 1943, McCulloch et son collègue Pitts décrivirent comment il était possible de définir des conditions d'excitation d'un neurone dans des schémas de réseaux de neurones formels. Ces réseaux constituaient des modèles simplifiés s'apparentant à des machines de Turing.



(a) $N_2(t) . \equiv . N_1(t-1)$
(b) $N_3(t) . \equiv . N_1(t-1) \vee N_2(t-1)$
(c) $N_3(t) . \equiv . N_1(t-1) . N_2(t-1)$
(d) $N_3(t) . \equiv N_1(t-1) . \sim N_2(t-1)$
(e) $N_3(t) : \equiv : N_1(t-1) . \vee . N_2(t-3) . \sim N_2(t-2)$
$N_4(t) . \equiv . N_2(t-2) . N_2(t-1)$
(f) $N_4(t) : \equiv : \sim N_1(t-1) . N_2(t-1) \vee N_3(t-1) . \vee . N_1(t-1)$
$N_2(t-1) . N_3(t-1)$
$N_4(t) : \equiv : \sim N_1(t-2) . N_2(t-2) \vee N_3(t-2) . \vee . N_1(t-2)$
$N_2(t-2) . N_3(t-2)$
(g) $N_3(t) . \equiv . N_2(t-2) . \sim N_1(t-3)$
(h) $N_2(t) . \equiv . N_1(t-1) . N_1(t-2)$
(i) $N_3(t) : \equiv : N_2(t-1) . \vee . N_1(t-1) . (Ex)t-1 . N_1(x) . N_2(x)$

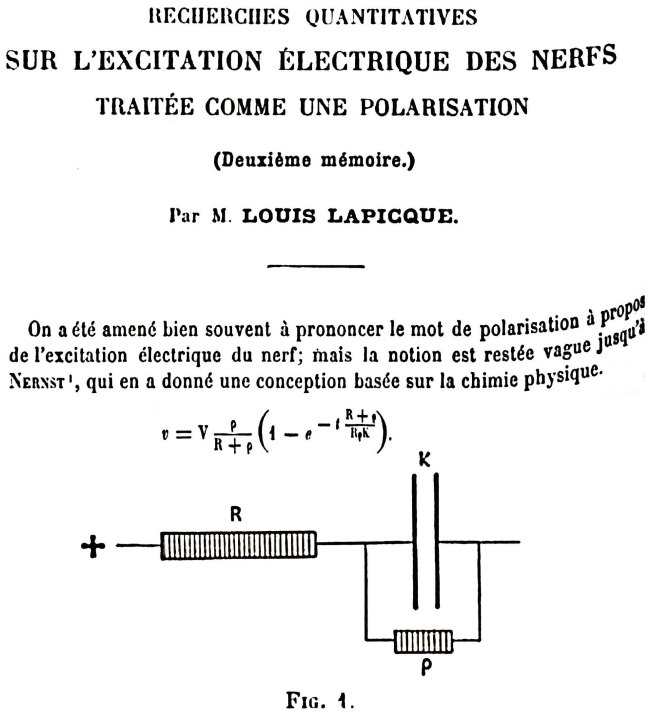
***Figure 1**. Les neurones formels de Warren Mc Culloch (1943)* (2).

Dans le cas le plus simple, un neurone (N3) n'est excité que si ses deux neurones présynaptiques (N1, N2) le sont simultanément. On peut alors dire que le neuroone N3 code pour la fonction logique « ET », dans la mesure où son excitation requiert celles des neurones N1 « ET » N2. Cette réflexion se plaçait dans la perspective philosophique de la logique propositionnelle et de l'algèbre de Boole, pour formaliser des fonctions logiques. Leur but était de faire comprendre que les réseaux de neurones pouvaient en principe réaliser des calculs logiques ou des opérations sur des signaux (les potentiels d'action), de la même manière que des dispositifs électroniques. Toutefois l'idée que des neurones pouvaient coder des informations n'était pas nouvelle, car le neurophysiologiste, Prix Nobel en 1932, Edgar Adrian (1889-1977) l'avait démontré à la fin des années 1920 avec le codage de l'intensité des sensations par la fréquence des potentiels d'action dans le nerf sensitif. Inversement des travaux comme ceux de McCulloch et Pitts aboutirent à la création d'intelligences artificielles par des réseaux de neurones artificiels.

Louis Lapicque, son modèle de neurone et l'intelligence artificielle

Le neurophysiologiste français Louis Lapicque, professeur en Sorbonne, ne participa qu'indirectement au développement de l'intelligence artificielle, lors de deux occasions très distantes. C'est dès 1907 que Lapicque publia un modèle de neurone assez simple dit « intègre et décharge » selon lequel un neurone peut émettre un potentiel d'action tout-ou-rien dès que son potentiel de base augmente au-delàs d'un certain seuil (Lapicque, 1907, Figure 2). Ce modèle était censé fournir un mécanisme de modulation de l'excitabilité d'un neurone en fonction d'une excitation de base (c'est-à-dire faible et de durée infinie). Il était prémonitoire tant le concept de potentiel d'action tout-ou-rien et le concept de potentiel de repos étaient encore très hypothétiques, compte tenu qu'aucun enregistrement électrophysiologique sur des nerfs n'avait pu encore étayer ces concepts. En effet on ne mesurait que l'effet des stimulations nerveuses presqu'exclusivement en mesurant l'intensité des contractions musculaires induites. Ce modèle resta largement oublié, puis il sera redécouvert et utilisé dans le développement de réseaux de neurones artificiels utilisant ce modèle pour générer de l'intelligence artificielle.

Au début des années 1940, Lapicque commença des discussions avec le mathématicien français Louis Couffignal, spécialiste du calcul logique et de grandes machines à calculer. Ce qui les intéressa alors furent les analogies possibles entre la structure de certaines machines et la structure microscopique du cervelet. Tous deux s'interrogeaient alors sur la possibilité d'analogies fonctionnelles entre ces structures qui présentent vue de très loin des similitudes d'organisations morphologiques, avec dans les deux cas des séries de structures parallèles (les feuillets du cervelet).


RECHERCHES QUANTITATIVES
SUR L'EXCITATION ÉLECTRIQUE DES NERFS
TRAITÉE COMME UNE POLARISATION
(Deuxième mémoire.)
Par M. LOUIS LAPICQUE.

On a été amené bien souvent à prononcer le mot de polarisation à propos de l'excitation électrique du nerf; mais la notion est restée vague jusqu'à Nernst[1], qui en a donné une conception basée sur la chimie physique.

$$v = V\frac{\rho}{R+\rho}\left(1 - e^{-t\frac{R+\rho}{R\rho K}}\right).$$

Fig. 1.

Figure 2. *Le modèle de neurone de Lapicque (1907)* (3).

Il est à présent facile d'imaginer comment Louis Lapicque aurait pu réaliser des modèles possédant des structures parallèles avec son modèle de neurone formel de 1907, afin d'imaginer quelles sortes d'opérations ces circuits étaient susceptibles de réaliser. Mais Lapicque et Couffignal s'en tinrent malheureusement à l'idée que des analogies superficielles de structures étaient sans aucun intérêt et qu'aucune base théorique solide n'était suffisante pour envisager de telles analogies fonctionnelles.

Si cette collaboration entre Lapicque et Couffignal n'aboutit pas dans les années 1940 à des spéculations sur l'intelligence artificielle, le modèle de 1907 de Lapicque fut quant à lui utilisé plus tard dans cette perspective de créer des réseaux de neurones artificiels capables d'opérations symboliques et d'une certaine forme d'intelligence. Le succès de ce modèle fut sa simplicité dès les années 1960, ce qui permit de l'employer comme unité de base des « spiking neural networks » de R. FitzHugh, J. Nagumo et W.J. Freeman. Puis il permit de construire, jusqu'à aujourd'hui, de grands circuits capables de comportements complexes avec les progrès des puissances de calcul.

Le psychologue visionnaire Kenneth Craik et les tortues intelligentes du neurophysiologiste William Grey Walter

Cette histoire concerne la création par le neurophysiologiste William Grey Walter d'animaux-robots à la fin des années 1940, ses célèbres « tortues électroniques » (Figure 3), capables de se mouvoir vers une source lumineuse (phototropisme) et d'associer deux stimuli comme déclencheur de leur déplacement (apprentissage par conditionnement). On peut dire que ces « animaux artificiels » appartiennent aussi bien à l'histoire de la neurophysiologie (théorique) qu'à l'histoire de l'intelligence artificielle. En effet, ils représentent aussi bien des modèles physiques du conditionnement pavlovien, avec lequel Grey Walter s'était familiarisé dans les années 1930, que la réalisation des premiers dispositifs électroniques intelligents.

Les réalisations de Grey Walter eurent pour origine des discussions avec le jeune psychologue Britannique Kenneth Craik rencontré pour une collaboration de recherche militaire au cours de la Seconde Guerre mondiale. Tous deux étaient alors vivement intéressés aussi bien par leurs domaines de recherche respectifs que par la micromécanique et l'électronique. Le bricolage électronique constituait alors déjà un hobby, bien que peu courant, pour certains jeunes savants imaginatifs.

Mais aussi bien pour Craik que pour Grey Walter ce « hobby » se montra bientôt d'une très grande utilité pour leurs domaines de recherche respectifs et ce de manière assez inattendue. En effet, Grey Walter s'était spécialisé dans l'étude de l'électroencéphalographie à la fin des années 1930 et ses connaissances en électronique l'avaient incité à construire un analyseur de fréquences utilisé pour l'analyse des électroencéphalogrammes. C'est ainsi que son analyseur électronique automatique, réalisant une sorte de transformée de Fourrier on line toutes les dix secondes, lui permit de mettre en évidence un rythme nouveau qu'il qualifia de rythme delta. Pendant l'effort de guerre, Grey Walter continua de travailler sur cet appareil qui pouvait être utile pour la surveillance militaire par l'analyse des sons émanant des dispositifs ennemis.

Craik rencontra Grey Walter pour lui demander d'utiliser son analyseur de fréquences sur des tracés obtenus par un dispositif de visée enregistrant des paramètres d'erreurs de tirs d'artilleurs. Il s'agissait là d'une étude de psychologie appliquée au tir dans laquelle Craik utilisait également ses talents de constructeurs de machines complexes. En marge de ce travail, et dans une perspective théorique et philosophique, Craik publia en 1943 un petit livre intitulé *The nature of explanation* (4). Il expliquait dans cet opuscule, comme Norbert Wiener au même moment, comment l'étude et la construction de dispositifs intelligents artificiels pouvaient révolutionner l'étude de la cognition humaine, en envisageant machines et cerveaux de la même manière.

Craik envisagea alors de construire des petits robots autonomes capables de se mouvoir librement en fonction d'un objectif et capables d'apprentissage. Ce petit livre, dont Einstein souligna l'importance, fut lu et pris très au sérieux par nombre de neurophysiologistes dont Grey Walter. Lorsqu'en 1945 Craik disparut prématurément dans un tragique accident de la circulation, Grey Walter poursuivit le développement de ses idées et il utilisa les ressources en électronique de son laboratoire pour construire des êtres électroniques intelligents autonomes, ses fameuses « tortues ». Ces machines firent alors la couverture de certains journaux et donnèrent lieu à des démonstrations publiques comme au colloque de Paris du CNRS de 1951 sur « Les machines à calculer et la pensée humaine ».

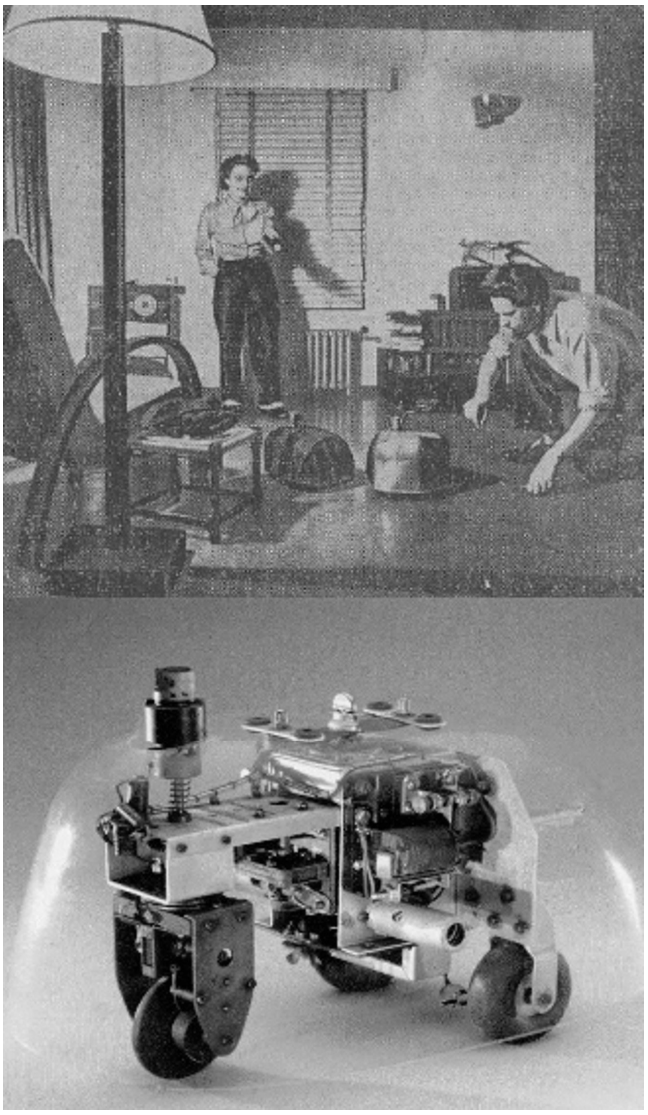


***Figure 3**. Coupures de presse montrant Grey Walter et deux tortues en déplacement dans son salon et leur cage. Bas, Reconstitution d'une tortue de Grey Walter avec une carapace transparente pour montrer la mécanique.*

Le projet de tortues de Grey Walter peut nous sembler appartenir aux aspects les plus spectaculaires et mondains de la presse s'intéressant alors à la cybernétique. Cependant, il est possible de mieux comprendre les motivations plus profondes de Grey Walter et la place que ces tortues tiennent dans ses réflexions.

Grey Walter était avant tout un neurologue, spécialiste d'électroencéphalographie et intéressé par les analyses des électroencéphalogrammes par des dispositifs électroniques, aussi bien en neurologie qu'en psychiatrie. Mais Grey Walter s'était également précédemment intéressé au conditionnement pavlovien (5). Grey Walter avait en effet étudié en 1935 l'effet de l'hypoxie sur l'apprentissage par conditionnement à Cambridge avec un élève de Pavlov. Il s'était alors montré très curieux de la typologie de Pavlov des « tempéraments » des chiens, utilisée pour expliquer les différences dans les résultats expérimentaux. Ces aspects psychologiques dans l'étude du conditionnement amenèrent plus tard Grey Walter à envisager la possibilité d'utiliser l'électroencéphalographie pour objectiver des différences individuelles. Grey Walter envisagea même de les corréler avec les capacités d'apprentissage de patients. Dans cette perspective, il voulut comprendre théoriquement comment des différences interindividuelles d'apprentissage pouvaient avoir des corrélats électroencéphalographiques. Pour s'en convaincre, il réalisa un modèle théorique neurophysiologique d'apprentissage par association basé sur les connaissances récemment acquises par les neurophysiologistes sur les voies thalamocorticales des années 1950. Ainsi il fut en mesure de proposer qu'il était en principe possible d'expliquer des aptitudes d'apprentissage par des mesures physiologiques.

Il est alors possible de mieux comprendre quelle place tiennent les tortues électroniques de Grey Walter dans le développement de ses travaux neurophysiologiques. Il semble en effet que Grey Walter ait souhaité donner une forme mathématique à son modèle neurophysiologique d'apprentissage, mais qu'il y renonça en raison de son manque de maîtrise des mathématiques (*little facility for mathematical representation*) (5). Les tortues de Grey Walter représentaient alors un modèle, non mathématique mais physique, celui d'un mécanisme d'apprentissage dont on pouvait étudier à la fois les performances et les indices de fonctionnement électrique. On comprend par-là comment la cybernétique et la construction de machines s'infiltra dans la neurophysiologie théorique à un stade très précoce de telles recherches.

En France, l'ingénieur et journaliste Albert Ducrocq (1921-2001) de la Société française d'électronique et de cybernétique construisit sur le modèle de ceux de Grey Walter ses propres robots au cours des années 1950, dont son « renard électronique » et son robot « Calliope » simulant l'imagination, en collaboration avec Louis Couffignal. Si la presse s'enthousiasma de manière naïve et si certains neurophysiologistes montrèrent les limites de tels modèles, il n'en demeure pas moins que tous ces animaux artificiels constituaient un premier pont entre la neurophysiologie et l'intelligence artificielle. On assista alors au déploiement de modèles cybernétiques et cognitifs de fonctions (comme la vision) à partir de données neurophysiologiques, et à l'émergence de nouveaux dispositifs artificiels intelligents inspirés des connaissances sur le vivant. Cette dernière tendance s'est alors considérablement développée depuis, jusqu'aux recherches actuelles en robotique et en intelligence artificielle.

Les « assemblées de neurones » de Donald Hebb, la plasticité de l'esprit et l'intelligence

C'est à la fin des années 1940 que le psychologue canadien publia son ouvrage célèbre, *The organization of behaviour* (1949) (6). Ce fut là le fruit de plus d'une décennie de réflexions sur le concept de plasticité du cerveau et sur celui d'intelligence. Alors jeune psychologue expérimentaliste, Hebb s'était mis au service du grand neurochirurgien canadien Wilder Penfield (1891-1976) de Montréal qui révolutionna la chirurgie des épileptiques. En psychologue clinicien, Hebb faisait des évaluations pré- et post-opératoires et étudiait les déficits et les gains de fonction que présentaient les patients dans certains cas.

En effet il arriva qu'un patient à qui l'on avait ôté une zone du cerveau présentât des performances d'intelligence supérieure à la situation pré-opératoire. Hebb imagina des modèles théoriques avec des zones différemment impliquées dans la mémoire et l'intelligence pour expliquer ses observations. Il lui arrivait alors de discuter avec Penfield qui n'acceptait pas toujours ses idées originales, mais Hebb rassembla au fil des années ses réflexions en adoptant des modèles de neurones formels ayant la capacité d'organiser leurs connexions selon des règles fixes imaginées (Figure 4). C'est par exemple le célèbre modèle de la « synapse de Hebb ».

Ce modèle de synapse représentait une formalisation du fonctionnement de chaînes fermées de neurones actifs (Lorente de Nó), d'une manière déjà utilisée par des neurologues de la fin du XIXe siècle, comme Exner ou le jeune Freud. Mais dans son modèle, Hebb imagina des conditions précises (comme dans le modèle de McCulloch et Pitts). Dans le modèle de Hebb, le renforcement d'une synapse établie entre deux neurones requérait la simultanéité (ou presque, à l'intérieur d'une

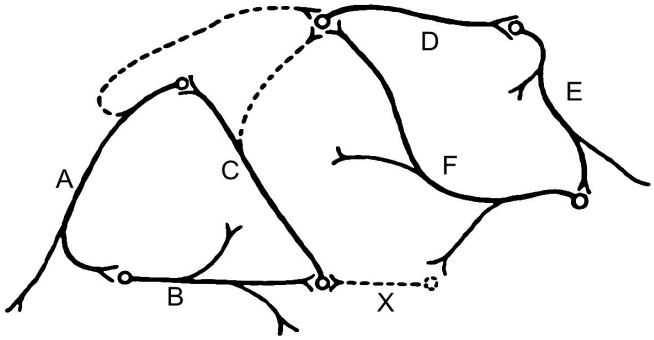


***Figure 4**. Ce schéma indique les modalités de création d'une « assemblée de neurones ». La figure montre deux groupes de neurones, le groupe A-B-C et le groupe D-E-F. Ces deux groupes de neurones sont localisés dans le cortex associatif et sont excités par les mêmes afférences sensorielles, mais pas en même temps. Les synapses sont renforcées à l'intérieur de chacun de ces groupes par la règle suivante : (on prend ici l'exemple des neurones A et B) Si le neurone A excite le neurone B alors que le neurone B est simultanément excité par l'afférence sensorielle, alors la synapse A-B sera renforcée. Cette règle explique la connectivité dans chaque assemblée de neurone, A-B-C et D-E-F. Dans chaque assemblée l'excitation d'un seul neurone engendre celle du circuit tout entier. Avec la même règle, il est alors possible que les deux assemblées fonctionnent comme un seul système ou une seule grande assemblée des neurones A-B-C-D-E-F s'il y a renforcement des synapses impliquant les afférences entre les deux groupes de neurones qui sont figurées en pointillés. Adapté de* (7).

petite fenêtre de temps) d'un influx présynaptique et d'une dépolarisation post-synaptique. C'était en quelque sorte une sorte de conditionnement pavlovien à l'échelle d'une synapse, si l'on imagine qu'une excitation synaptique inconditionnelle était nécessaire pour dépolariser le neurone post-synaptique et que l'excitation synaptique conditionnelle était renforcée par la simultanéité des deux stimulus (comme la nourriture et le son de la cloche du chien de Pavlov). Ce schéma de renforcement synaptique de Hebb eut et a encore une postérité incroyable, car il devint d'autant plus populaire et célèbre chez les neurophysiologistes qu'il anticipa les protocoles d'instauration de potentialisation et de dépression à long-terme, dans lesquels on réalise un appariement de ces deux types de stimulation pour obtenir un renforcement synaptique à long-terme.

Mais au début des années 1950, ce schéma de renforcement intéressa des ingénieurs travaillant chez IBM et qui tentèrent de l'instancier (d'en créer une nouvelle instance, une nouvelle catégorie) dans leurs schémas de dispositifs artificiels capables de stocker des informations dans leur projet de créer un ordinateur. En effet les modèles de neurones de Hebb ou assemblées de neurones (*cell ou neuronal assemblies*) fournirent un principe d'architecture de réseaux de neurones formels stockant des informations sous la forme de réseaux de neurones interconnectés pouvant présenter un certain patron électrique synchrone.

Conclusion

Ces quatre exemples démontrent combien localement et dans des contextes spécifiques certains neuro-physiologistes ont pu s'intéresser avant même l'essor de la cybernétique à un nouveau mode de pensée, dans lequel l'évaluation des capacités de dispositifs artificiels intelligents ou de modèles formels de neurones avait toute sa place dans leurs recherches.

Ce mode d'évaluation était en effet considéré comme directement applicable au vivant. Et en retour, l'organisation du vivant ou sa modélisation donnait lieu à de nouvelles formes d'organisation de dispositifs artificiels. Nul doute que ce mode de pensée fut bien l'un des fondements de l'essor de la cybernétique et de sa présentation auprès du grand public avec le débat autour de la pensée humaine et de la pensée artificielle.

Ce débat trouva peu à peu son chemin en se concentrant finalement sur l'idée que la cybernétique permettait d'offrir de nouveaux modèles de l'esprit humain et de ses facultés en incorporant le concept d'information et celui de commande, dans un esprit progressivement cognitiviste. Et d'une manière parallèle, il fut alors accepté que la création de dispositifs artificiels intelligents serait tôt ou tard capable de reproduire beaucoup de capacités humaines, y compris la lecture et la création de textes comme cela est actuellement disponible sur internet. On assista alors à la diffusion du vocabulaire cybernétique dans la neurophysiologie comme le terme de feed-back et l'utilisation de modèles informatiques pour l'analyse des comportements, d'une manière qui rapproche encore constamment les neurosciences du domaine de l'intelligence artificielle.

jean-gael.barbara @sorbonne-universite.fr

Références

(1) Surianarayanan, C.; Lawrence, J.J.; Chelliah, P.R.; Prakash, E.; Hewage, C. (2023). Convergence of Artificial Intelligence and Neuroscience towards the Diagnosis of Neurological Disorders – A Scoping Review. Sensors, 23, 3062.

(2) McCulloch W.S. and Pitts, W. (1943). "A Logical Calculus of the Ideas Immanent in Nervous Activity," The Bulletin of Mathematical Biophysics 5(4): 115–133.

(3) Lapicque L. (1907). « Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation », Journal de physiologie et de pathologie générale, 9, 620-635.

(4) Craik K.J.W. (1943). The Nature of Explanation, Cambridge University Press.

(5) Grey Walter, W. (1960). Conditioning Theories and Their Therapeutic Applications, Proceedings of the Royal Society of Medicine, 53, 1960, 495-503.

(6) Hebb, D. O. (1949). *The organization of behavior. a neuropsychological theory*. John Wiley & Sons.

(7) Hebb D.O. (1958). A textbook of Psychology, Philadelphie, Londres: W.B., Saunders Company.

Histoire des Neurosciences

Histoire de la Société de Psychophysiologie et de Neurosciences Cognitives[1]

THOMAS HINAULT[1], HENRIQUE SEQUEIRA[2], AURÉLIE BIDET-CAULET[3], ANNE CACLIN[4], MARIE GOMOT[5], NATHALIE GEORGE[6], MARIE-HÉLÈNE GIARD[7]*, MARINE TIMSIT-BERTHIER[8]*

[1]Université de Caen Normandie, Inserm, EPHE- PSL, CHU de Caen, GIP Cyceron, U1077, NIMH, Caen, France
[2] Université de Lille, CNRS, UMR 9193, SCALab - DEEP-Neuro, Villeneuve d'Ascq, France.
[3]Aix Marseille Univ, Inserm, INS, Inst Neurosci Syst, Marseille, France
[4]Université Claude Bernard Lyon 1, INSERM, CNRS, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, Bron, France
[5]Université de Tours, INSERM, Imaging Brain and Neuropsychiatry iBraiN U1253, 37032, Tours, France
[6]Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France
[7]Retraitée, Université Claude Bernard Lyon 1, INSERM, CNRS, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, Bron, France
[8]Laboratoire de Psychophysiologie Cognitive Appliquée. Faculté de Médecine, Liège, Belgique
**Co-dernier auteurs*

La Société de Psychophysiologie et de Neurosciences Cognitives (SPNC) s'intéresse au fonctionnement du cerveau humain en lien avec le fonctionnement cognitif. Cet article retrace le contexte et les origines de la société et son évolution jusqu'à nos jours.

1 De l'Electroencéphalographie à la Psychophysiologie

Les racines de la SPNC remontent au développement de l'électroencéphalographie (EEG, Figure 1), technique découverte en 1929 par le neurologue allemand Hans Berger, et du traitement des données électrophysiologiques avec des calculateurs permettant l'application de méthodes mathématiques de plus en plus sophistiquées. La SPNC a accompagné les avancées théoriques et méthodologiques de cette approche, ainsi que leurs applications cliniques, sur plusieurs décennies. Plus précisément, les travaux qui annoncent la perspective scientifique de la future SPNC débutent en 1934, par Alphonse Baudoin (1876-1957), Hermann Fischgold (1899-1982), Jean Scherrer (1917-2007), et par

***Figure 1** : Illustration d'un des premiers électroencéphalographes (Grass, modèle 8-10C).*

[1] *NDLR : La Société de Psychophysiologie et de Neurosciences Cognitives est un club affilié à la Société des Neurosciences (https://www.neurosciences.asso.fr/clubs-scientifiques/)*

Alfred Fessard (1900-1982), neurophysiologiste et chef de file des neurophysiologistes à partir des années 1950 et l'un des fondateurs du nouveau domaine des neurosciences, et de l'EEG à la suite d'Hans Berger. À noter aussi que Fessard contribuera, avec Henri Piéron, psychologue français et l'un des fondateurs de la psychologie scientifique en France, à la création d'un diplôme national de psychophysiologie dès 1944.

Dans ce contexte, le premier congrès international d'électroencéphalographie a eu lieu dès 1947 à Londres, alors que le développement de cette nouvelle technique en France prenait du retard par rapport aux pays anglo-saxons (Grande-Bretagne et USA). Ce retard est notamment à mettre en relation avec la lenteur de l'introduction des ordinateurs dans les programmes de recherche en neurosciences. Toutefois, deux laboratoires pionniers ont su dépasser ces difficultés pour développer l'usage de l'EEG: celui du professeur Antoine Rémond, fondateur à Paris en 1948 du Laboratoire d'Electrophysiologie et de Neurophysiologie Appliquée (LENA) (1) et le Laboratoire de Neurophysiologie Expérimentale et Clinique fondé quelques années plus tard par Michel Jouvet, professeur spécialiste du sommeil à Lyon.

Quelques années plus tard, en 1967, se tenait un congrès scientifique international à Liège (Belgique), sous la présidence de William Grey Walter. Pionnier de l'EEG, ce neurophysiologiste britannique est à l'origine de la découverte des ondes delta (rythme cérébral lent observé entre 1 et 3 Hz) et thêta (rythme cérébral observé entre 4 et 8 Hz) associées au sommeil. Il était question durant ce colloque d'un Potentiel Evoqué (PE) spécifique, la variation contingente négative (ou CNV pour *Contingent Negative Variations*), observé sur les électrodes antérieures environ 260–470 ms après la présentation d'un stimulus impliquant la préparation à une décision (Figure 2).

2. L'apport des études sur les potentiels évoqués

Dans ce contexte, l'histoire de la société débute véritablement en 1972, à Tours, avec une réunion intitulée « *Activités évoquées et leur conditionnement chez l'homme normal et en pathologie mentale* » organisée par Gilbert Lelord, sous la présidence d'Alfred Fessard. Parmi les nombreuses avancées qui sont attribuées à ce dernier, citons les mécanismes sous-jacents à la fatigabilité musculaire et les modulations du rythme alpha à la suite de conditionnements. Gilbert Lelord était quant à lui professeur à la Faculté de Médecine de Tours et co-fondateur d'une Unité INSERM. Il a grandement contribué à préciser les particularités cérébrales associées à l'autisme, étayant ainsi les conceptions neurodéveloppementales des troubles de ce spectre, s'opposant aux conceptions psychogènes alors en vigueur. Les activités ou PE sont obtenus par moyennage des signaux EEG liés à un événement observé à un instant temporel précis (à l'échelle de la milliseconde) et répété plusieurs fois (de quelques dizaines à plusieurs milliers). Leur lien avec les étapes élémentaires du traitement de l'information est au cœur des études de psychophysiologie cognitive.

La CNV n'est plus le thème central du congrès de Tours. Il s'agit alors de préciser l'usage de l'électroencéphalographie centrée sur les potentiels évoqués comme outil permettant de distinguer le fonctionnement « normal » du cerveau de celui relevant de la pathologie. L'objectif était double : d'une part, mieux comprendre le fonctionnement cognitif humain et les mécanismes cérébraux sous-jacents ; d'autre part, contribuer à la détection des atteintes neurologiques et psychiatriques ainsi qu'à l'identification des marqueurs associés aux effets des interventions thérapeutiques. Ces préoccupations sont restées centrales dans l'histoire de la société, et plus généralement des méthodes d'exploration fonctionnelle du cerveau humain, jusqu'à ce jour.

En 1982, se crée un *Groupe de travail et de réflexion sur les composantes tardives des potentiels évoqués, de langue française* porté par Nicole Lesèvre, alors directrice du LENA à l'Hôpital de la Pitié-Salpêtrière de Paris, Jacques Paty de Bordeaux et Martine Timsit de Liège. Ce groupe se réunit régulièrement et informellement à Paris, avec pour but d'expliciter les concepts et les modèles utilisés en psychophysiologie, et de présenter et discuter

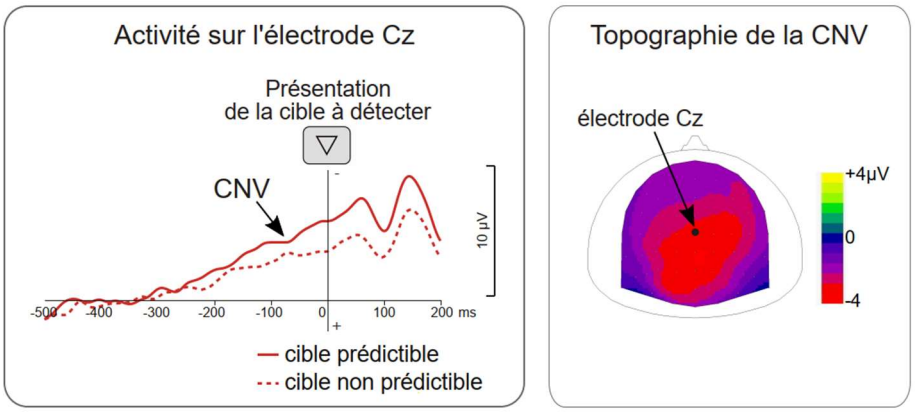


Figure 2 : *Illustration de la variation contingente négative (CNV) observée lors de la présentation d'une cible à détecter, ainsi que la topographie de cette composante. La négativité est plus importante lors que le stimulus est présenté à un emplacement prédictible à l'écran que non prédictible.*

les avancées méthodologiques du domaine. Nicole Lesèvre a publié des travaux majeurs sur l'utilisation des PE pour établir la chronométrie des opérations mentales, et a précisé les corrélats EEG des processus d'attention visuelle. Jacques Paty s'est intéressé aux fluctuations attentionnelles en fonction du niveau d'éveil, ainsi qu'à la composante P300 (PE positif survenant à partir d'environ 300 ms après la présentation d'un stimulus, associé à la détection de stimuli inattendus). Enfin, Martine Timsit, neuropsychiatre, a précisé l'interprétation des composantes P300 et CNV et leur variabilité inter-individuelle.

Au cours des premières réunions de ce groupe de travail, les corrélats électrophysiologiques des paradigmes de conditionnement furent abordés. L'accent sur le fonctionnement cognitif était central : qualifier et quantifier les processus mentaux qui surviennent lors d'une stimulation et d'une réponse (encodage, maintien des informations pertinentes, suppression des informations distractrices, préparation de la réponse, etc.), là où les paramètres comportementaux ne fournissent qu'une information plus limitée (temps de réponse ou taux d'erreur). Ces réunions ont précisé les notions de stades successifs de traitement de l'information et « d'états fonctionnels » : le cerveau peut fonctionner sous différents types de régimes qui déterminent le type d'attention, d'apprentissage et de mémoire. Ces états correspondraient à des organisations neuronales particulières, reflétées par les activités EEG. En cette année 1982, le groupe fut très actif, organisant trois rencontres, centrées sur les différents PE identifiés. Ceci répondait au besoin de clarifier ce champ de recherche en pleine expansion avec l'identification rapide de composantes toujours nouvelles : quels paradigmes donnent lieu à quelles composantes, quels sont les processus cognitifs associés ?

Par la suite, 22 réunions eurent lieu jusqu'en 1989, puis des séminaires annuels, abordant plusieurs grandes thématiques. Tout d'abord, il s'agissait de suivre les avancées méthodologiques concernant le traitement du signal enregistré en EEG, la cartographie des potentiels évoqués, et avec la technique nouvelle de la magnétoencéphalographie (MEG, mesurant le versant magnétique de l'activité neuronale). L'analyse en composantes principales de l'activité enregistrée a notamment permis de séparer efficacement l'activité cérébrale des artefacts physiologiques (clignements des yeux, activité cardiaque) ainsi que de décomposer les contributions respectives des différents rythmes cérébraux et PE pendant la réalisation de tâches cognitives. Une autre avancée méthodologique majeure a concerné l'identification des sources cérébrales de l'activité mesurée (dite « problème inverse »), avec l'objectif d'améliorer la résolution spatiale des mesures électrophysiologiques. Un très grand nombre de configurations de sources peuvent donner lieu à la même topographie d'activité EEG telle qu'enregistrée à la surface du scalp. L'intégration d'informations sur l'anatomie cérébrale individuelle, la position des électrodes sur le scalp ou encore les *a priori* sur les sources recherchées sont essentiels pour déterminer la configuration la plus probable. Ces points méthodologiques restent centraux aujourd'hui.

Sur le plan expérimental, de nombreux facteurs sont étudiés, les PE permettant d'observer l'effet des stimulations et des contextes. Sont ainsi abordés l'occurrence ou la disparition de ces composantes en fonction des conditions expérimentales, comme la disparition de la CNV après satiété, la distinction entre traitement automatique et volontaire, les effets du stress ou l'association de la CNV avec les capacités d'estimation temporelle. Les limites du moyennage du signal EEG au cours de la réalisation des tâches et entre individus dans l'étude de la variabilité intra- et interindividuelle sont aussi un thème récurrent.

Comme indiqué plus haut, un des objectifs des études était d'examiner les données psychophysiologiques dans le domaine clinique ; de ce fait, les réunions du groupe de travail puis de la société ont notamment soulevé les difficultés associées à l'utilisation de l'EEG auprès des populations de patients. Il s'agissait de caractériser les altérations de PE associées à diverses atteintes neurodéveloppementales, psychiatriques et neurologiques (par exemple, autisme, schizophrénie, maladie d'Alzheimer), leurs spécificités, et leurs caractéristiques communes.

Ces réunions ont fait progresser les connaissances théoriques en psychophysiologie. Les PE permettent de faire la chronométrie des processus cognitifs, inaccessibles par l'introspection. Leur étude ainsi que celle des rythmes cérébraux a permis de faire avancer les connaissances sur l'adaptation du fonctionnement cognitif en fonction du contexte, des expériences précédentes et sur les liens entre processus cognitifs, comme entre attention et mémoire. Les résultats ont aussi soulevé des hypothèses concernant les bases neurochimiques des potentiels lents et leur association avec d'autres mesures. Les réunions visaient aussi à échanger au sein de la communauté pour harmoniser les termes utilisés : composantes tardives des PE ou potentiels « cognitifs » ? Ces termes seront finalement englobés plus tard sous l'appellation internationale *Event-related potentials*. On voit ici la difficulté à préciser les corrélats cérébraux du fonctionnement cognitif, sans réductionnisme, tout en essayant de rendre compte de l'expérience subjective des individus.

Durant ces années, si des événements européens ont été organisés (Mannheim, Liège, Lausanne), les réunions ont

principalement eu lieu à Paris, Marseille, Tours et Lyon. En plus des trois fondateurs du groupe de travail, elles ont été organisées par plusieurs figures de l'électrophysiologie française : Bernard Renault, chercheur au CNRS et directeur du LENA à la suite de Nicole Lesèvre, a fortement contribué au développement de l'électroencéphalographie comme méthode d'imagerie cérébrale et a été une personnalité centrale du groupe de travail ; Gilbert Lelord, déjà cité, fut un pionnier des recherches dans le domaine de l'autisme en France; Françoise Macar, chercheuse au CNRS à Marseille, et Franck Vidal, professeur à l'Université d'Aix-Marseille ont précisé les mécanismes physiologiques associés au traitement et à la représentation du temps ; Marie-Hélène Giard, directrice de recherche INSERM à Lyon, a étudié les bases neurophysiologiques de l'attention auditive et l'intégration des informations visuelles et auditives dans la perception; Bernard Claverie, professeur et directeur de l'École nationale supérieure de cognitique de Bordeaux, a étudié l'électroencéphalographie intra-cérébrale et les mécanismes associés à l'épilepsie ; Nicole Bruneau, neurophysiologiste et chercheuse INSERM à Tours, a étudié l'intégration auditive corticale au cours du développement neurotypique et dans l'autisme.

3. La création de la société

La société, nommée « Société de Psychophysiologie cognitive » en 1991, prendra l'appellation actuelle « Société de Psychophysiologie et de Neurosciences Cognitives » en 2010, lors de la réunion annuelle, à Paris ; cette appellation visait à intégrer l'essor nouveau des neurosciences cognitives à travers diverses méthodes d'imagerie cérébrale. À cette occasion, un vif débat a eu lieu sur l'idée d'ajouter à la nouvelle appellation, celle de « Neurosciences Affectives ». Finalement, le terme « Neurosciences Cognitives » a été considéré comme intégrant à la fois les processus sensoriels, ceux de plus haut niveau et les aspects affectifs. Pendant cette période, les séminaires annuels ont continué de garder une ouverture vers d'autres pays francophones (Bruxelles) tout en impliquant de nouveaux laboratoires en France (Strasbourg, Toulouse, Lille). Sur le plan technologique, la société continue de suivre les avancées de l'électroencéphalographie cognitive, tout en intégrant celles de l'électrophysiologie neurovégétative (1999, Lille). Les PE restent une thématique centrale de la société ainsi que les rythmes cérébraux, avec notamment une emphase sur l'activité cérébrale à haute fréquence (dans la bande dite gamma, au-delà de 40 Hz). Les avancées concernent également l'imagerie cérébrale multimodale du fonctionnement cognitif, combinant en particulier imagerie électrophysiologique et hémodynamique, telle que l'imagerie fonctionnelle par résonance magnétique (IRMf).

Les thématiques traitées au sein de la SPNC accompagnent l'évolution de nos connaissances sur les processus cognitifs et émotionnels associés au fonctionnement cérébral. Dans ce contexte, de nouveaux champs font leur apparition, comme la cognition sociale (processus cognitifs permettant les interactions sociales), les corrélats cérébraux des processus psycho-linguistiques, ou encore le lien entre pathologies et émotions.

La volonté de développer la contribution de l'électroencéphalographie à la caractérisation des atteintes pathologiques, et son usage en clinique, reste très présente lors de ces réunions. Les séminaires de la période 1993-2010 porteront notamment sur les applications cliniques des potentiels évoqués identifiés (P300, *mismatch negativity* [MMN] et CNV), les corrélats cérébraux associés à l'autisme, ou encore aux perturbations du traitement sémantique. Les thématiques de la société intègrent aussi la contribution de l'EEG dans le contexte des réhabilitations sensorielles et motrices.

Durant cette période, la société, toujours affiliée à la Société des Neurosciences, développe également d'importants liens avec d'autres groupes de recherche et sociétés savantes. Elle devient membre de l'*International Organization of Psychophysiology* (IOP, créée en 1982 et regroupant les recherches en psychophysiologie sur le plan international) en 2014 et a également un représentant à l'*European Society for Cognitive and Affective Neuroscience* (ESCAN). La société a par ailleurs rejoint en 2024 le Collège des Sociétés Savantes Académiques de France.

Depuis 2010, les réunions annuelles de la SPNC se poursuivent avec l'implication de laboratoires issus de nouvelles villes comme Chambéry ou Caen, et le développement de champs d'étude supplémentaires en lien direct avec les développements théoriques en psychologie cognitive comme l'étude de la cognition incarnée (prise en compte de l'influence des expériences sensorielles et de la position corporelle sur le fonctionnement cognitif), ou l'exploration des fonctions cognitives ou émotionnelles à l'aide de l'EEG intracérébrale, dans le contexte pré-chirurgical de l'épilepsie. L'association des mesures EEG avec de nouvelles méthodologies comme la fNIRS (imagerie spectroscopique en proche infrarouge) sont également abordées.

En conclusion, la SPNC, sous ses différentes appellations, a accompagné le développement francophone de l'électrophysiologie et de la magnétoencéphalographie depuis plus de quatre décennies. Elle maintient de fortes associations avec la communauté internationale : en

2023, la société faisait partie de la conférence *Cutting Gardens*, regroupant 21 institutions de 15 pays autour d'un programme scientifique commun. Par ailleurs, à l'occasion de plusieurs réunions annuelles de la SPNC, de nombreux collègues européens ont été invités à présenter leurs contributions majeures. La SPNC continue ainsi de contribuer au développement des recherches en psychophysiologie et en neurosciences cognitives, dans un cadre toujours informel et amical (*voir https://spnc.fr/*).

thomas.hinault@inserm.fr
henrique.sequeira@univ-lille.fr
aurelie.bidet-caulet@inserm.fr
anne.caclin@inserm.fr
marie.gomot@univ-tours.fr
nathalie.george@sorbonne-universite.fr
mh.steiner@orange.fr
timsit.berthier@wanadoo.fr

Références

(1) C. Chérici, Barbara J.G., 2007, « EEG, trois lettres pour percer les mystères du cerveau », La revue pour l'Histoire du CNRS, 19, 21-25

Dossier

Dialogues entre neurosciences et IA

PAR MAURICE GARRET

L'Intelligence Artificielle (IA) est omniprésente dans notre quotidien. Elle effraie ou elle fascine selon les situations. Elle occupe une place particulière dans le domaine des neurosciences car d'un côté, les propriétés anatomiques et fonctionnelles du cerveau aident au développement de l'IA, et de l'autre, les méthodes issues de l'IA participent à une meilleure analyse ou modélisation de toute la complexité de l'anatomie et du fonctionnement du cerveau. Ne parle-t-on pas de neurones et de réseaux de neurones dans les deux cas ? Pour toutes ces raisons, nous avons choisi de consacrer une partie de ce numéro à la place de l'IA dans les neurosciences, à travers le dossier de ce numéro et la rubrique d'Histoire des neurosciences.

Aborder précisément les liens entre le l'IA et les neurosciences nécessiterait plusieurs ouvrages tant ces interactions sont nombreuses et diversifiées. Nous avons donc dû faire des choix et nous avons demandé aux experts de ces domaines d'aborder quelques exemples caractéristiques comme la place de l'IA dans les interactions cerveau-machines. Le contrôle du mouvement ne passe plus par une commande musculaire envoyée vers notre corps mais simplement par la pensée, ce qui était de la science-fiction dans les années 90 devient réalité. Les capacités de calcul de l'IA permettent d'analyser l'activité neuronale, d'induire des réponses adaptées aux différentes tâches, et d'apprendre. Ces travaux trouvent leurs applications dans la mise au point de neuroprothèses pour la restauration de la marche chez les tétraplégiques ou la récupération de mouvements après un accident. Au-delà des applications cliniques, sur un plan plus fondamental l'IA devient également un outil indispensable pour l'étude du cerveau. Vous lirez ainsi dans ce dossier comment des neurobiologistes et des informaticiens tourangeaux ont développé un nouvel outil d'analyse d'images obtenues en IRM sur différents modèles animaux. Les images sont analysées sans a priori anatomique. L'IA peut également être mise à profit dans l'étude et la modélisation des comportements sociaux tels que ceux décrits chez les fourmis et les termites. À l'aide de la modélisation mathématique, les chercheurs sont parvenus à décoder les interactions complexes entre les insectes, et notamment leur coordination lors de la construction du nid, incluant des comportements très sophistiqués. Un dossier à lire sans tarder.

Pour compléter ce dossier ne manquez pas l'article d'Histoire des Neurosciences dans lequel Jean-Gaël Barbara explique comment dans les années 50, à partir de quatre exemples aux États-Unis, en France, en Angleterre et au Canada, des cybernéticiens, des neurophysiologistes, des psychologues et des mathématiciens ont collaboré pour créer des dispositifs artificiels intelligents, capables de reproduire les capacités de l'intelligence humaine. On peut imaginer que les discussions n'ont pas toujours été simples entre les acteurs de ces différentes disciplines mais elles ont été fécondes et elles illustrent bien l'importance de l'interdisciplinarité.

D'interdisciplinarité il en est également question dans l'article de Thomas Hinault et son équipe qui nous expliquent comment la rencontre entre des

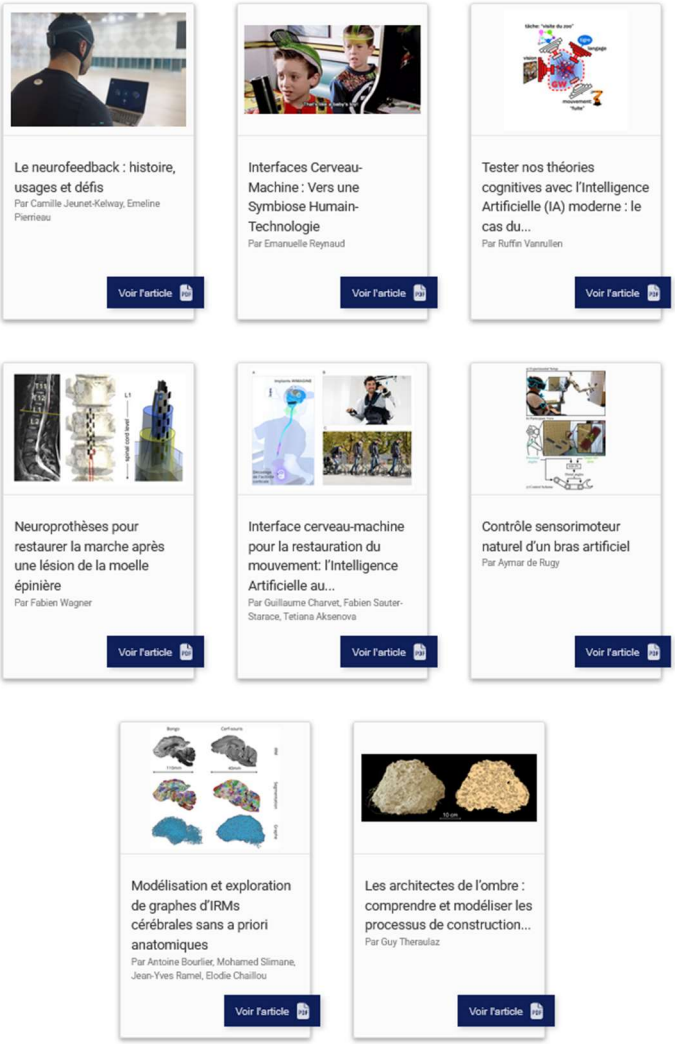
psychologues et des neurophysiologistes a donné naissance à la Société de Psychophysiologie et de Neurosciences Cognitives. Le but des chercheurs de cette Société, à savoir la compréhension du fonctionnement du cerveau et de ses capacités cognitives, était déjà au cœur des préoccupations des chercheurs en IA.

Dans la rubrique nouveauté, Mathieu Wolff nous livre les dernières données sur le thalamus, une structure longtemps délaissée car considérée comme un simple relais vers le cortex. Les plus récentes études montrent qu'en réalité le thalamus joue un rôle important – pour ne pas dire central – dans les fonctions intégratives et cognitives. Mathieu Wolff nous explique que le thalamus est impliqué notamment dans la mémoire, la flexibilité comportementale, la prise de décisions, autant de fonctions classiquement associées au cortex préfrontal. De même, un certain nombre de pathologies mentales sont associées à un dysfonctionnement thalamique. Toutes ces observations donnent une autre image du thalamus et son rôle mérite sans nul doute d'être revisité avec la considération que suggère sa position centrale dans le cerveau. Un article à ne surtout pas manquer.

Enfin, la recherche expérimentale sur les sujets humains soulève de nombreuses questions éthiques qui prennent peut-être encore plus d'importance dès lors qu'elles s'intéressent aux enfants. Ces recherches sont précieuses pour l'étude des mécanismes d'apprentissage et des processus cognitifs car elles servent aussi de références pour la construction et la mise en place des programmes scolaires par les pouvoirs publics. En regard de ces enjeux, Sébastien Goudeau et son équipe abordent un certain nombre de recommandations, notamment éthiques, pour faciliter et optimiser les collaborations entre les acteurs de cette recherche dans le milieu scolaire.

Nous espérons que cette Lettre saura retenir votre attention et assouvir votre curiosité, nous vous garantissons qu'elle a été écrite « avec I mais sans A ».

Bonne lecture.



Dossier

Le neurofeedback : histoire, usages et défis

CAMILLE JEUNET-KELWAY & EMELINE PIERRIEAU

Univ. Bordeaux, CNRS, EPHE, INCIA, UMR5287 F-33000 Bordeaux

Introduction

L'histoire du neurofeedback débute lorsque H. Berger enregistre en 1924 une activité électrophysiologique chez l'humain, par le biais d'électrodes placées à la surface du cuir chevelu (1). Cette innovation, aujourd'hui appelée électroencéphalographie (EEG), lui permet d'étudier le rôle fonctionnel des activités cérébrales. Il décrira notamment les ondes *alpha* (8-12Hz) et la modulation de leur amplitude durant le sommeil, mais aussi l'altération de certaines activités cérébrales lors de crises d'épilepsie. Dès les années 1930, l'usage de l'EEG à visée clinique et de recherche va largement se développer, et la compréhension des liens entre activité cérébrale, cognition et comportement s'approfondir. En 1969, J. Kamiya réussira à entraîner, grâce à un protocole basé sur le conditionnement opérant, des personnes souffrant de syndromes anxieux et dépressifs à auto-réguler l'amplitude de leurs ondes alpha et ainsi diminuer leurs symptômes cliniques (1). M.B. Sterman rapportera quant à lui les effets anticonvulsivants de l'auto-régulation des rythmes sensori-moteurs (2). C'est la naissance du neurofeedback (NF), qui consiste ainsi à apprendre, grâce à un retour sensoriel fourni en temps réel, à contrôler des activités cérébrales spécifiques sous-tendant une capacité ou habileté (e.g., cognitive, motrice) que l'on souhaite améliorer ou restaurer.

Plus tard, les années 1970 voient se développer l'informatique. J. Vidal imagine alors un système permettant aux humains de communiquer avec un ordinateur uniquement grâce à leur activité cérébrale (3) : une Interface Cerveau-Ordinateur (ICO). Les ICO sont aujourd'hui définies comme « *des systèmes qui mesurent l'activité cérébrale et la convertissent en (quasi) temps réel en outputs fonctionnels utiles pour remplacer, restaurer, augmenter, supplémenter, et/ou améliorer les outputs naturels du cerveau [..].* » selon la *BCI Society*, sur la base d'un questionnaire complété en 2024 par près de 150 scientifiques travaillant dans le domaine des ICO. Cette définition, très large, recoupe à la fois le contrôle d'applications et l'entraînement cognitif, historiquement associés aux champs des ICO et du NF, respectivement. Tous deux reposent sur le même principe d'apprentissage de l'auto-régulation d'activités cérébrales spécifiques, et diffèrent principalement par l'objectif de cette auto-régulation. Dans le champ du NF, l'objectif est l'amélioration ou la restauration de capacités/habiletés. Il s'agit donc d'identifier *a priori* des activités cérébrales d'intérêt (approche *knowledge-based*) et d'entraîner la personne à les auto-réguler. Dans le champ des ICO, l'objectif est de contrôler une application de manière fiable. Il s'agit donc d'identifier des activités cérébrales stables et distinctes, chacune permettant l'envoi d'une commande spécifique à l'application. Une approche *data-driven* est donc privilégiée : pour chaque utilisateur, des algorithmes (p.ex., d'apprentissage automatique) seront utilisés pour identifier les activités cérébrales qui permettront d'envoyer les commandes les plus fiables.

Par la suite, nous aborderons les trois applications principales de l'entraînement à l'auto-régulation d'activités cérébrales à l'heure actuelle : en premier l'étude de mécanismes neurophysiologiques, puis le contrôle d'applications et enfin l'entraînement cognitif. Nous nous concentrerons sur la mesure d'activités oscillatoires par EEG parce que ce sont les protocoles les plus utilisés aujourd'hui, en raison de l'accessibilité et de la portabilité de l'EEG, et des connaissances acquises sur le rôle fonctionnel de ces activités oscillatoires.

Le neurofeedback comme outil d'étude de mécanismes neurophysiologiques

Entraîner une personne à auto-réguler une activité cérébrale en vue de modifier un comportement nécessite d'avoir établi l'existence d'une relation causale entre les

deux. Une telle démonstration peut être réalisée de manière non invasive par l'intermédiaire de quatre outils : la stimulation magnétique transcrânienne répétée, la stimulation transcrânienne à courant alternatif, la stimulation transcrânienne par ultrasons et le NF. Ce dernier constitue ainsi l'unique méthode de modulation endogène (c'est-à-dire sans stimulation externe). Le NF peut néanmoins indirectement engendrer des modulations d'autres activités que celle ciblée, si bien que la relation de causalité pourrait en réalité ne pas s'appliquer spécifiquement à l'activité cérébrale ciblée mais plutôt à un réseau d'activités cérébrales modulées avec le NF. Ce phénomène est qualifié de causalité indirecte et révèle l'importance d'étudier les changements globaux de l'activité cérébrale engendrés par le NF, alors que seules les modulations locales de l'activité cérébrale ciblée sont souvent rapportées. A défaut de démonstrations de causalité, le NF a néanmoins pour intérêt de permettre de dissocier l'impact de la modulation d'une activité cérébrale sur différentes variables comportementales. Par exemple, alors qu'une hausse de l'activité beta (13-30 Hz) au niveau des régions motrices a été associée à un ralentissement du mouvement dans la maladie de Parkinson, une étude faisant usage du NF nous a récemment permis de démontrer que l'activité beta était en réalité prédictive de la flexibilité motrice et non de la vitesse du mouvement, sa réduction permettant aussi bien d'accélérer que de ralentir le mouvement selon l'instruction donnée (4) (voir Fig. 1).

Le neurofeedback pour permettre le contrôle d'applicationss

Le contrôle d'applications représente l'usage historique des ICO. Originellement, l'objectif était de permettre à des personnes dans un état « locked-in » de regagner en autonomie grâce à l'envoi de commandes mentales par le biais de l'auto-régulation d'activités sensori-motrices (2). Bien que des progrès importants aient été faits, les performances actuelles de ces ICO demeurent incompatibles avec un usage en vie réelle. Il existe en réalité une forte variabilité intra- et inter-individuelle en termes de performances de contrôle ICO. Etudier cette variabilité afin d'en comprendre les origines et les déterminants (neurophysiologiques, cognitifs et psychologiques) permettrait d'adapter les procédures d'entraînement et d'améliorer l'efficience des ICO. Malheureusement, il n'existe à l'heure actuelle aucune base de données issues d'expérimentations avec neurofeedback dont la taille et la qualité permettraient de capturer et de comprendre cette variabilité. La communauté européenne a pris ce problème à bras-le-corps et s'est unie en consortium afin de concevoir un protocole standardisé, de collecter et de partager une telle base de données (5). En attendant, des solutions sont possibles pour améliorer la fiabilité des ICO. La première consiste à utiliser des méthodes d'imagerie cérébrale plus précises que l'EEG. Il s'agit notamment d'avoir recours à des méthodes semi-invasives du type électro-corticographie, p.ex. pour contrôler un exosquelette ou restaurer la marche via la stimulation de la moelle épinière (6); ou encore d'avoir recours à des méthodes invasives telles que des électrodes *Utah arrays*, p.ex. pour décoder le langage (7). Bien que présentant un meilleur rapport signal/bruit et une moindre sensibilité aux artefacts, ces méthodes nécessitent une opération chirurgicale. Elles sont donc associées à un risque accru et une acceptabilité potentiellement moindre. La seconde solution consiste à modifier le type de d'activité ciblée, et de s'intéresser aux potentiels évoqués. Il s'agit par exemple de créer un fauteuil roulant intelligent autonome et d'utiliser la détection de potentiels d'erreur pour en ajuster la

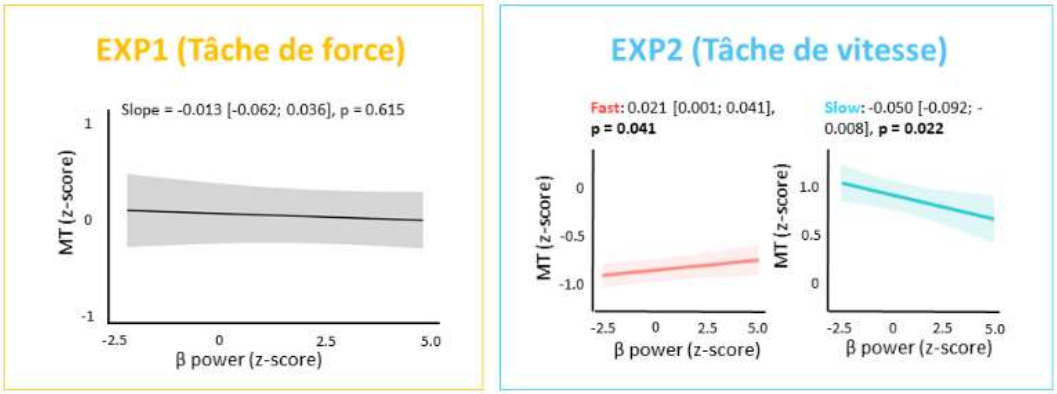


Figure 1. *Illustration des résultats de (4). Une association significative entre temps de mouvement (MT) et activité beta (β power) a été observée lors d'une tâche de vitesse (encadré bleu), mais pas lors d'une tâche de force (encadré jaune). Le signe de la relation linéaire liant temps de mouvement et activité beta était inversé en fonction de l'instruction de vitesse : positif lors de mouvements rapides et négatif lors de mouvements lents. En d'autres termes, réduire l'activité beta se traduisait par une meilleure performance à la tâche, que celle-ci implique d'augmenter ou de réduire la vitesse des mouvements.*

trajectoire ; ou encore d'avoir recours à un P300-speller pour communiquer. Ces deux types d'ICO sont qualifiées d'ICO passive (parce que l'utilisateur n'envoie pas de commande volontaire) et réactive (parce que la commande est générée par la réponse à un stimulus), respectivement, en opposition aux ICO actives. Ces ICO ont l'avantage de nécessiter des ressources cognitives et un temps d'entraînement moindres, et sont associées, pour les ICO réactives du moins, à des performances bien supérieures à celles des ICO actives. Néanmoins, les ICO passives ne sont pas encore très fiables, étant donné la nécessité de détection des états en essais uniques. Les ICO réactives peuvent quant à elles générer un inconfort.

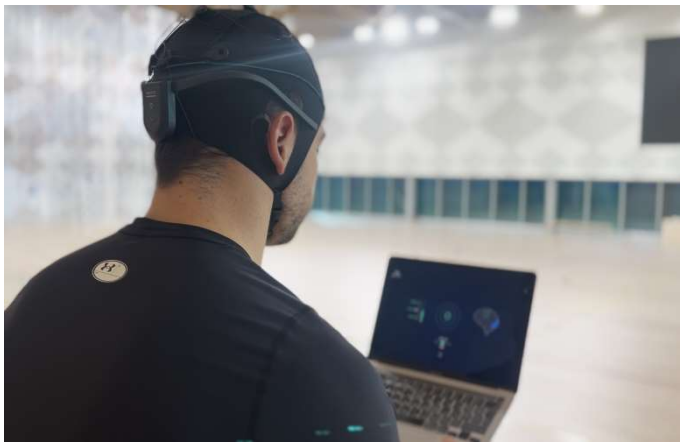


***Figure 2**. Photo représentant un athlète lors d'une séance de neurofeedback. L'athlète est équipé d'un bonnet EEG et s'entraîne à moduler ses rythmes sensori-moteurs afin d'améliorer son geste.*

Le neurofeedback comme outil d'entrainement cognitif

L'entraînement cognitif constitue le principal usage du NF à l'heure actuelle. Son champ d'application s'étend de l'optimisation des performances cognitives et motrices chez le sujet sain (p.ex., pilote, athlète – voir Fig. 2) à la récupération motrice suite à un accident vasculaire cérébral (AVC), et à la remédiation cognitive dans le trouble du déficit de l'attention avec ou sans hyperactivité (TDAH). Chez des patients souffrant de séquelles motrices suite à un AVC, l'entraînement par NF renforçant la réduction des activités sensorimotrices au niveau du cortex moteur ipsilésionnel a démontré de manière consistante une amélioration significative du score clinique moteur (8) (voir Fig. 3). La réduction des activités theta par l'intermédiaire d'un NF a quant à elle été associée à une réduction des symptômes liés au TDAH (9). Cependant, l'intérêt du NF pour le TDAH a été questionné en raison de la demande attentionnelle inhérente à l'entraînement par NF. Est-ce réellement la modulation de l'activité cérébrale ciblée, ou bien le fait d'effectuer plusieurs sessions d'un entraînement cognitif qui explique ces résultats positifs ? Thibault et al. (10) suggèrent ainsi que l'efficience du NF relève d'un effet placebo, et soulève la question importante de la condition contrôle. Bien que l'utilisation d'une condition *sham* semble constituer l'option la plus robuste, elle présente le risque d'un désengagement potentiel du participant du fait que le feedback soit perçu comme *faux* ou *incontrôlable*, ou encore par les problèmes éthiques causés par son utilisation en clinique. Au-delà de la nécessité de paradigmes plus « contrôlés », l'optimisation du NF dans l'objectif de faciliter sa modulation par le sujet constitue un point central d'amélioration.

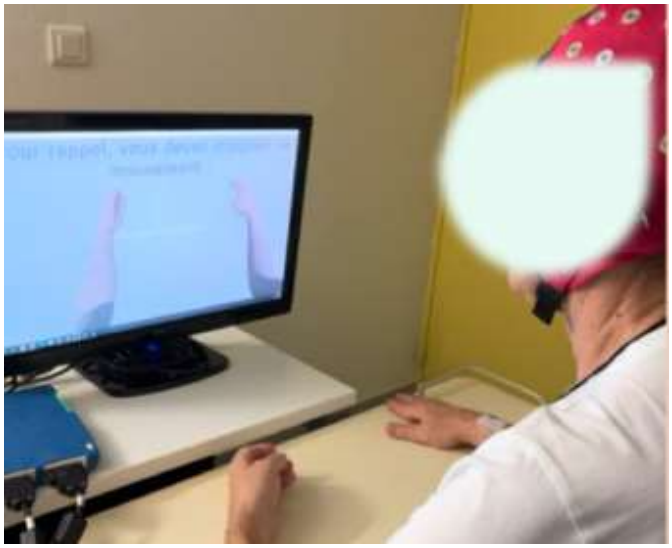


***Figure 3**. Photo d'une personne ayant subi un AVC et prenant part à l'étude clinique que nous menons dans le cadre du projet ANR ABCIS, visant à évaluer la pertinence d'une procédure de rééducation fonctionnelle personnalisée du membre supérieur en phase subaiguë en termes de récupération motrice et d'acceptabilité. Au signal, la personne imaginera un mouvement de son membre plégique. Son activité cérébrale sera alors analysée. Si la modulation de ses rythmes sensorimoteurs au niveau controlatéral dépasse un certain seuil, un neurofeedback sera fourni sous forme de mouvements de la main virtuelle à l'écran (dont l'amplitude est proportionnelle au taux de désynchronisation). Ce neurofeedback synchronisé avec l'imagerie motrice permet de clore la boucle sensori-motrice et ainsi de favoriser la plasticité synaptique.*

Cette optimisation ne pourra se faire qu'à condition, dans un premier temps, de comprendre les mécanismes psychologiques, cognitifs et neurophysiologiques impactant l'apprentissage par NF. C'est dans cet objectif que nous avons conçu un modèle (11) mettant en évidence l'implication majeure de facteurs *spécifiques*, p.ex., les capacités d'imagerie mentale pour les entraînements faisant appel à l'imagerie motrice, et de facteurs *non-spécifiques*, influents quel que soit le paradigme de neurofeedback, et divisés en deux catégories : i) les facteur cognitifs et motivationnels (p.ex., attention, engagement) et ii) les facteurs liés à la relation à la technologie (p.ex., auto-efficacité, anxiété, agentivité). Ces facteurs *non spécifiques* pourraient expliquer qu'un apprentissage puisse avoir lieu dans des groupes contrôle pour qui le neuromarqueur cible n'est pas celui récompensé. Plusieurs méta-analyses (8,9) montrent cependant qu'il existe bien aussi des effets spécifiques et

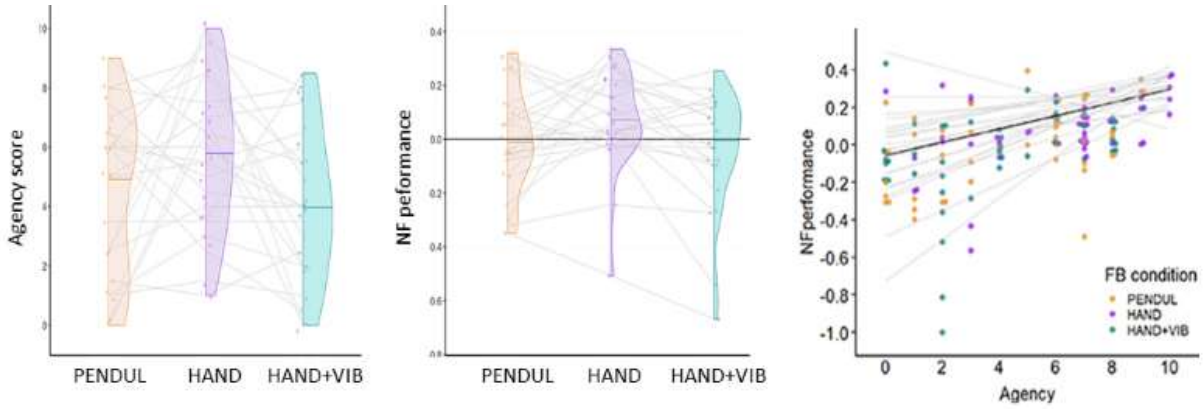


***Figure 4**. Illustration des résultats de (12). Le neurofeedback était présenté sous 3 formes différentes : un pendule (PENDUL), une main virtuelle seule (HAND) et une main virtuelle associée à des vibrations (HAND+VIB). La condition de main virtuelle seule (HAND) était associée à un score d'agentivité et une performance au neurofeedback plus élevés que les autres conditions. La performance au neurofeedback s'est avérée positivement corrélée au score d'agentivité au travers des différentes conditions expérimentales*

que le NF ne se résume pas à un effet placebo ou *non spécifique*. Par ailleurs, il est intéressant, pour maximiser l'apprentissage, de maximiser aussi les effets *non spécifiques*. Dans le cadre d'une étude menée chez 25 sujets sains, nous avons observé que le sentiment d'agentivité était affecté par la nature du feedback présenté et influençait significativement la performance au NF et la capacité des participants à moduler l'activité cérébrale visée (12) (voir Fig. 4). Ces résultats confirment qu'adapter les protocoles NF pour optimiser l'attention, la motivation, l'auto-efficacité, l'agentivité et réduire l'anxiété (c-à-d les facteurs non spécifiques) permettrait d'améliorer l'efficacité *spécifique* du NF.

Conclusion

Bien qu'ayant déjà plus de 50 ans d'histoire, le NF et les ICO connaissent un développement exponentiel récent. Plus de 50% des documents relatant de la recherche et de l'innovation dans le domaine ont été publiés au cours de ces 5 dernières années (~6 600 publications scientifiques). Ce développement se fait à la faveur d'une meilleure compréhension des liens entre cerveau, cognition et comportement, permise i) par une fiabilité et une accessibilité accrues des outils d'imagerie cérébrale et ii) par des méthodes de modélisation, de traitement du signal et d'IA de plus en plus robustes. C'est d'ailleurs sur ces deux axes que se concentrent aujourd'hui la majorité des investissements (de Neuralink à Meta en passant par InclusiveBrains, la start-up Marseillaise qui a permis, grâce à son exosquelette contrôlé par une ICO, à Nathalie Lebregere de porter la Flamme Olympique en 2024). La place de *l'humain dans la boucle* est quant à elle souvent négligée. Or, si l'on souhaite pouvoir mesurer (avec des capteurs) et décoder (avec de l'IA) des signaux cérébraux d'intérêt, il est nécessaire en amont que la personne soit capable de les générer. En réalité, un double apprentissage est mis en jeu : l'humain doit apprendre à auto-réguler les activités cérébrales ciblées alors que la machine doit apprendre à les reconnaître, malgré leur variabilité. Optimiser ce double apprentissage, et plus largement l'efficience du NF et des ICO, représente un défi interdisciplinaire dans lequel les neuroscientifiques, encore trop peu représentés, ont un rôle majeur à jouer :

- Neurophysiologie : clarifier les mécanismes sous-tendant les activités cérébrales ciblées, la manière dont elles évoluent dans le temps (non-stationnarité), les déterminants de cette évolution, et leur l'influence sur les fonctions perceptuelles, cognitives et motrices.
- Neurologie et psychiatrie : indiquer les activités cérébrales impliquées dans la physiopathologie et la symptomatologie des troubles neuropsychiatriques.
- Neuropsychologie : concevoir des procédures de NF prenant en considération les facteurs psychologiques propres à la symptomatologie des patients.
- Neuroréhabilitation : renforcer l'arsenal clinique grâce aux neurotechnologies, améliorer leur acceptabilité et leur efficience.
- Neurosciences cognitives et neuroéducation : identifier les déterminants des capacités d'apprentissage par NF, concevoir des procédures d'entraînement permettant d'optimiser cet apprentissage.
- Neuroergonomie : étudier la manière dont des états mentaux au sens large (p.ex., charge cognitive, fatigue) sont reflétés dans l'activité cérébrale et influent sur les capacités d'auto-régulation.
- Neuroingénierie : innover avec des algorithmes d'IA *brain-inspired*, améliorer les performances de décodage

Le *Gartner Hype Cycle* 2024, qui fournit une représentation graphique des attentes liées à différentes technologies émergeantes et de leur maturité, montre que, dans 5 à 10 ans, nous pourrions connaître une phase de désillusion liée au fait que le NF et les ICO ne répondent

pas aux attentes élevées de la population. Nous sommes en effet dans une période de forte hausse de ces attentes, stimulée par la couverture médiatique des ICO qui laisse à penser que des personnes avec des handicaps moteurs sévères pourront à nouveau communiquer et se déplacer de manière naturelle, en toute autonomie. Il est donc de notre responsabilité, i) d'ajuster notre communication pour éviter de générer des attentes déraisonnables et ii) d'orienter notre recherche afin de nous assurer de répondre aux besoins et attentes des utilisateurs finaux. Pour ce faire, il est essentiel de remettre l'humain au centre de la boucle et d'adopter une approche centrée sur l'utilisateur

camille.jeunet@u-bordeaux.fr

Références

(1) Kübler A. (2025). *Proceedings of the 13th International Winter Conference on Brain-Computer Interfaces*.

(2) Sterman M.B. (1984) In: Elbert, T., Rockstroh, B., Lutzenberger, W., Birbaumer, N. (eds) *Self-Regulation of the Brain and Behavior*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-69379-3_8

(3) Vidal J. (1973) *Annu Rev Biophys Bioeng*, 2:157-80. doi: 10.1146/annurev.bb.02.060173.001105.

(4) Pierrieau E., et al. (2025) *bioRxiv,* 2025.01.23.634491

(5) Jeunet, C., et al. (2020). *CHIST-ERA*

(6) Wagner FB. et al. (2018) *Nature,* 563(7729):65–71.

(7) Vansteensel L.J. & Jarosiewicz B. (2020) *Handb Clin Neurol,* 168:67-85. doi: 10.1016/B978-0-444-63934-9.00007-X.

(8) Bai Z. et al. (2020) *J Neuro Engineering Rehabil,* 17, 57. https://doi.org/10.1186/s12984-020-00686-2

(9) Van Doren, J. et al. (2019) *Eur Child Adolesc Psychiatry*, 28(3):293-305. doi: 10.1007/s00787-018-1121-4.

(10) Thibault R.T. Lifshitz M., Raz A. (2016) *Cortex*, 74:247-261.

(11) Jeunet C. et al. (2016) *Prog Brain Res,* 228:3-35. doi: 10.1016/bs.pbr.2016.04.002.

(12) Dussard C. et al. (2024) *J Neural Eng*, 21(5). doi: 10.1088/1741-2552/ad7f88.

.

Dossier

Interfaces Cerveau-Machine : Vers une Symbiose Humain-Technologie

EMANUELLE REYNAUD

Laboratoire Étude des Mécanismes Cognitifs - Université Lumière Lyon II

Les Interfaces Cerveau-Machines (ICM) : à l'aube d'une nouvelle interaction

L'histoire de l'humanité est celle d'une relation toujours plus intime avec la technologie. Des premiers outils en pierre taillée aux machines industrielles, des premiers ordinateurs aux smartphones d'aujourd'hui, chaque avancée a repoussé les limites de nos capacités naturelles. Mais en parallèle, un autre défi émerge : réduire toujours plus l'effort physique requis pour interagir avec ces technologies, jusqu'à, pourquoi pas, abolir cette contrainte. En effet, autrefois l'homme façonnait le monde de ses mains ; chaque outil exigeait un geste, une pression, un mouvement. Puis vinrent les machines, les interfaces graphiques, les écrans tactiles et la commande vocale, réduisant peu à peu l'effort nécessaire pour communiquer avec la technologie.

Aujourd'hui, une nouvelle révolution se profile : celle des interfaces cerveau-machine (ICM) (1). Ces dispositifs veulent établir un lien direct entre l'activité cérébrale et un système informatique. Plus besoin de levier, de manette, de bouton ni même de clic : la pensée seule suffirait à commander la machine. Le lien entre l'homme et la technologie ne passera plus par son corps, mais par sa pensée. L'effort physique disparaîtrait, ouvrant la voie à une interaction purement mentale avec nos outils, l'humain intégrant plus ou moins aisément le rôle du Commandant-en-chef du système humain-machine (Figure 1).

De la science-fiction à la réalité

Longtemps reléguées aux récits de science-fiction, les ICM sont aujourd'hui en plein essor, portées par les avancées en neurosciences et en intelligence artificielle (IA). En 2023, la revue *Nature Electronics* (2) les a consacrées technologie de l'année, soulignant leur potentiel disruptif dans des domaines aussi variés que la médecine, la communication et le divertissement.

***Figure 1**- Scène de "**Retour vers le Futur**", Part II (1989), dialogue entre deux enfants du futur et Marty, qui joue à un jeu d'arcades des années 80 manuellement :*

Enfant #1 : Tu veux dire qu'il faut se servir de ses mains

Enfant #2 : Ça ressemble à un jouet pour bébés !

Comment fonctionne une interface cerveau-machine ?

Les ICM reposent sur un principe fascinant et en apparence simple : convertir l'activité cérébrale en signaux exploitables par une machine. En réalité ces interfaces dépendent de fait de deux processus par essence extrêmement complexes et techniquement exigeants : l'enregistrement et l'interprétation des signaux cérébraux (3).

Le cerveau génère des signaux électriques mesurés à différents niveaux :

- ICM non-invasives : électrodes placées à la surface du crâne (ex. : électroencéphalographie).
- ICM semi-invasives : électrodes positionnées à la surface du cortex (ex. : électrocorticographie).

- ICM invasives : implants insérés directement dans le cerveau, offrant une meilleure résolution.

Une fois ces signaux enregistrés, ils sont filtrés pour éliminer le bruit puis analysés pour en extraire des caractéristiques pertinentes. Des algorithmes d'IA dits de Machine Learning apprennent ensuite à les traduire en instructions exploitables. Enfin, ces instructions sont transmises à un dispositif externe : bras robotique, ordinateur, interface de communication ou même exosquelette.
Ce principe en apparence simple cache une réalité beaucoup moins triviale : les signaux cérébraux sont en effet très bruités, changeants dans le temps, parfois inconsistants, et uniques à chaque individu. La phase d'apprentissage, chargée d'accorder la technologie avec les spécificités cognitives de l'utilisateur doit donc se dérouler avec des signaux par nature hétérogènes. Pour contrer cette hétérogénéité, elle doit se baser sur des données les moins bruitées possibles et doit employer des algorithmes d'apprentissage spécifiques et adaptés à chaque individu.

Applications et bénéfices concrets des ICM

Autrefois laborieux, le décodage des signaux cérébraux gagne aujourd'hui en fluidité grâce aux progrès du Machine Learning, des neurosciences, de la neuroimagerie et de la neuro-ingénierie. L'amélioration des capteurs, la puissance croissante des algorithmes et l'émergence de dispositifs biocompatibles ouvrent désormais la voie à des usages jusque-là seulement fantasmés.
L'une des applications les plus prometteuses des ICM concerne les personnes en situation de handicap. Les progrès dans ce domaine ont été spectaculaires, avec des avancées qui permettent aujourd'hui de restaurer une partie précieuse de leur autonomie. Elles permettent notamment à des patients atteints de paralysie de contrôler de façon fluide des prothèses robotiques, par exemple un bras artificiel, par la seule pensée (4). Une interface cerveau-machine bidirectionnelle a récemment permis d'établir un pont artificiel entre le cerveau et la moelle épinière. Grâce à cette avancée spectaculaire, l'activité cérébrale d'un patient paralysé a été enregistrée, puis convertie en signaux de stimulation épidurale contournant la lésion de la moelle. Ce processus a permis au patient de contrôler ses muscles des membres inférieurs par la pensée et de retrouver la marche (5).
Au-delà de la mobilité, les ICM offrent des perspectives prometteuses dans le domaine de la communication, en particulier pour les personnes atteintes de maladies neurodégénératives comme la sclérose latérale amyotrophique (SLA), qui peuvent perdre toute capacité de communication verbale ou physique. En interprétant leur activité cérébrale, ces interfaces permettraient de rétablir l'interaction avec leurs proches ou soignants en permettant de générer du texte, des commandes ou des phrases. Par exemple, grâce aux signaux émis par le cerveau (aire 6v) pour commander l'appareil articulatoire d'un patient paralysé incapable de parler, une ICM a permis de convertir ses pensées en mots à un rythme d'environ un mot par seconde. Le taux d'erreur était inférieur à 10 % pour un vocabulaire appris de 50 mots et inférieur à 25 % pour un lexique de 125 000 mots (6).
Si les premières ICM ont été conçues pour compenser des handicaps, leur champ d'application s'étend désormais bien au-delà du domaine médical. Dans la sphère du divertissement par exemple, les ICM peuvent être utilisées pour proposer de nouvelles façons, plus immersives, d'interagir avec un environnement de jeu virtuel, par exemple en contrôlant la navigation d'un avatar par la pensée (7).

Perspectives d'avenir

- *Des ICM plus performantes et acceptables pour tous*

Les progrès rapides en IA, en neurosciences et en nanotechnologies laissent présager un avenir où les ICM seront plus précises, moins invasives et accessibles à un plus grand nombre de personnes. En effet, aujourd'hui, les ICM non-invasives sont limitées en termes de résolution et de fiabilité. Le rapport signal sur bruit est souvent très défavorable, et les algorithmes fournissent des résultats par essence imparfaits, ce qui conduit à des erreurs. Le développement de capteurs plus sensibles et d'algorithmes toujours plus sophistiqués permettront de mieux décoder des signaux cérébraux de meilleure qualité, et d'atteindre des performances supérieures en termes de précision et de rapidité des décodages effectués. Les dispositifs de recueil de l'activité cérébrale vont se miniaturiser pour devenir plus discrets, plus confortables, davantage biocompatibles, et de ce fait plus facilement acceptés par l'utilisateur.
Alors même que l'enregistrement et l'interprétation des données progressent, un défi majeur subsiste : la production même du signal par l'utilisateur. En effet, certaines personnes éprouvent des difficultés à contrôler une ICM en raison d'une variabilité très élevée de leurs signaux cérébraux à travers le temps et l'espace cérébral. Ce phénomène, qualifié d'« illettrisme des ICM », est actuellement étudié afin d'être surmonté (8), car il pourrait exclure environ 20 % de la population de l'accès à cette technologie.

- *Augmentation cognitive*

Si le domaine médical continuera d'être un moteur majeur du développement des ICM, à plus long terme, on peut imaginer des ICM utilisés pour améliorer les capacités cognitives et physiques des individus, ouvrant ainsi la voie au concept de l'humain augmenté, au cœur du courant transhumaniste. Les travaux entrepris en ce sens restent modestes et surtout soumis à des questions de (neuro)éthique pour l'instant limitantes (9).

- *Interactions Humain-Machine directes, fusion avec l'Intelligence Artificielle ?*

Dans le secteur des nouvelles technologies, quelques travaux étudient des dispositifs capables de faire évoluer nos interactions quotidiennes avec les machines, nous dispensant des fastidieux moyens d'interaction physiques existants et rendant les interactions Humain-Machine plus fluides et intuitives. Naviguer sur internet sans écran, contrôler son environnement ou diriger des véhicules par la pensée : autant d'exemples dignes des classiques de la Science-Fiction qui pourraient devenir une réalité. Toujours à la frontière de la Science-Fiction, l'intégration de l'IA avec les ICM pourrait permettre une communication fluide entre la pensée humaine et les machines, voire une anticipation des besoins et des intentions de l'utilisateur par une IA totalement adaptée et couplée à son utilisateur. Ces couplages seraient à même de révolutionner des domaines comme la vie domestique, ou le travail (10).

- *Défis éthiques et risques*

Ces éléments prospectifs nous amènent à comprendre que le développement des ICM doit inévitablement s'accompagner d'une réflexion approfondie sur leurs implications éthiques et sociétales, qui doit être menée par l'ensemble des acteurs concernés. L'évolution des ICM n'est pas seulement une affaire d'ingénieurs et de neuroscientifiques, mais elle concerne l'ensemble de la société : la communauté scientifique, mais aussi le monde politique et la société civile doivent se saisir de questions essentielles sur le respect de la vie privée, la souveraineté cognitive et l'avenir de l'identité humaine. La question de la protection et de l'utilisation des données cérébrales dans un cadre éthique strict suscite des préoccupations croissantes. Ces inquiétudes s'intensifient à mesure que la recherche se développe au sein d'organisations étatiques ou privées dont la transparence reste limitée. La question de la possibilité de la manipulation et du contrôle mental peut également se poser : une utilisation malveillante des ICM pourrait-elle permettre d'hacker le cerveau ? La question de la responsabilité partagée entre l'utilisateur et l'ICM est également fondamentale, comme elle se pose aujourd'hui avec les voitures autonomes. Celle de la sureté sanitaire de ces dispositifs est un prérequis à l'utilisation des ICM. Et enfin, la question plus large de l'évolution de l'humanité dans son ensemble devrait être au centre des débats.

Conclusion : Vers une symbiose Humain-Technologie ?

Les ICM représentent l'un des tournants technologiques majeurs du XXIe siècle, à la croisée des neurosciences, de l'IA et de la technologie. Bien qu'elles posent des défis importants, leur potentiel pour transformer la médecine, la communication et le divertissement est immense. Ce qui semblait autrefois fantaisiste est en passe de devenir une réalité tangible, promettant d'améliorer la qualité de vie de millions de personnes. L'idée d'une fusion entre l'homme et la machine, longtemps explorée par la science-fiction, devient de plus en plus tangible. Alors que des dystopies comme *Matrix* (1999) ou *Black Mirror* (2011) nous ont mis en garde contre une soumission incontrôlée de l'humain à la machine, d'autres récits, comme *Neuromancien* (1984) ou *Ghost in the Shell* (1995), explorent la possibilité d'un partenariat entre l'homme et la machine, où chacun compenserait les limites de l'autre. La technologie ne viendrait pas remplacer ou dominer l'humain, mais collaborer de manière transparente avec lui dans une symbiose Humain-Technologie. La coopération entre la pensée biologique et la puissance de calcul artificiel, fusionnant le raisonnement intuitif de l'homme et la rapidité d'exécution des machines pourrait permettre aux ICM de devenir l'outil ultime pour dépasser nos propres limites, sans jamais perdre ce qui fait de nous des êtres humains.

Emanuelle.Reynaud@univ-lyon2.fr

Références

(1) Wolpaw J. et al. (2002) *Clin Neurophysiol,* 113, 767–791.

(2) An interface connects (2023) *Nat Electron,* 6, 89 https://doi.org/10.1038/s41928-023-00938-8 .

(3) Nicolelis M. A. L. (2001) *Nature* 409, 403–407.

(4) Robinson N. et al. (2021) *Curr Opin Biomed Eng,* 20, 100354.

(5) Lorach H. et al. (2023) *Nature* 618, 126–133.

(6) Willett F. R. et al. (2023) *Nature* 620, 1031–1036.

(7) Kapgate D. D. (2024) *Behav Brain Res,* 472, 115154.

(8) Kim D.-H. et al (2023) *Front Hum Neurosci*, 17. https://doi.org/10.3389/fnhum.2023.1194751

(9) Gordon E. C. & Seth A. K. (2024) *PLOS Biol*, 22 e3002899.

(10) Maiseli B. et al. (2023) *Brain informatics* 10, 20.

Dossier

Tester nos théories cognitives avec l'Intelligence Artificielle (IA) moderne : le cas du Global Workspace

RUFIN VANRULLEN

CerCo, CNRS UMR 5549, Université de Toulouse, Pavillon Baudot, Hôpital Purpan, 31052 Toulouse Cedex

Comment de nouveaux réseaux de neurones profonds, inspirés de la théorie cognitive du Global Workspace, nous permettent-ils de vérifier l'apport fonctionnel de cette architecture computationnelle ?

La théorie du Global Workspace.

Les neurosciences cognitives regorgent de théories promettant d'expliquer l'émergence de phénomènes complexes tels que la perception, la mémoire, la cognition ou même la conscience. Prenons l'exemple de la théorie du « Global Workspace » proposée par B. Baars pour décrire la cognition et la conscience humaines et animales (1). Un ensemble de « modules » de traitement neuronal spécialisés entrent en compétition pour l'accès à un espace de travail global (le Global Workspace ou GW). Cette compétition est arbitrée par un processus attentionnel tenant compte de l'état du système et de ses objectifs. L'information qui entre dans le Global Workspace est ensuite diffusée à l'ensemble des modules, donnant lieu à une représentation unifiée pour le système entier (c'est le « broadcast »).

Comment savoir si le cerveau fonctionne vraiment de cette manière ? En pratique, on teste plutôt des prédictions expérimentales qui découlent—plus ou moins directement—de la théorie. Il faut d'abord postuler quels composants, structures et processus neuronaux remplissent les différentes fonctions supposées : modules, compétition, workspace, broadcast… C'est ce que propose, par exemple, l'hypothèse du Global *Neuronal* Workspace de S. Dehaene et collègues (2). Il devient ainsi possible de tester expérimentalement les prédictions attenantes, comme dans l'exemple récent d'une collaboration « adversairielle[1] » entre partisans du Global Workspace et ceux d'une théorie concurrente, la théorie de l'Information Intégrée ou IIT (3). Les seuls résultats expérimentaux, pourtant, ne suffisent pas à valider ou invalider l'une ou l'autre théorie : même en cas d'échec, peut-être est-ce la correspondance entre théorie et processus neuronaux qui serait à revoir, plutôt que la théorie elle-même ?

Comment tester *directement* l'idée centrale d'une théorie ? Pour le GW : comment savoir si une architecture cognitive combinant des modules spécialistes, connectés dynamiquement à une représentation centrale partagée, pourrait suffire à engendrer des capacités cognitives avancées ? Il faudrait pouvoir produire artificiellement de telles architectures, afin de les évaluer sur des tâches complexes - une idée techniquement irréalisable, jusqu'à très récemment…

IA et apprentissage profond

Au cours des dernières années, les progrès de l'IA et notamment de l'apprentissage profond (ou « deep learning ») ont constamment redéfini les frontières de ce que l'on pensait être techniquement réalisable. Le deep learning consiste à entrainer des réseaux de neurones artificiels comportant des dizaines de couches, des millions de neurones et jusqu'à plusieurs milliards de connexions, afin de résoudre une tâche prédéfinie. Classifier une image, reconnaître un visage ou une voix,

[1] *Ce terme est emprunté aux GANs, les réseaux génératifs adversairiels qui améliorent leurs performances en opposant deux sous-réseaux « adversaires », entraînés conjointement.*

maîtriser un jeu vidéo ou de stratégie, mais aussi générer de toutes pièces une image, une vidéo réaliste ou un morceau de musique, ou simplement répondre à des questions et requêtes farfelues en langage naturel—combien de ces capacités aurait-on prédit que l'IA pourrait atteindre, il y a seulement 10 ans ?

Une conséquence de ces avancées de l'IA pour les neurosciences est la possibilité inédite de déployer nos théories cognitives sur des systèmes artificiels à grande échelle, et ainsi les mettre à l'épreuve (4).

Une implémentation IA du Global Workspace

C'est ce que nous avons entrepris de faire pour la théorie du Global Workspace (5). La procédure consiste à interconnecter différents réseaux préexistants qui jouent le rôle des modules spécialistes. Ceux-ci peuvent traiter des entrées visuelles, audio, ou du langage, mais également produire des images, du son, du texte, voire même contrôler les actions d'un robot. Un système attentionnel dédié décide du ou des modules qui pourront accéder au GW à chaque instant, en fonction de l'état du système, de ses objectifs et du contenu actuel de chaque module. Lorsqu'un module est sélectionné par l'attention, il implante son contenu dans le GW ; pour cela, un encodeur spécifique de chaque module convertit son activité en une représentation commune, une sorte de *lingua franca* ou langue véhiculaire du Global Workspace. Il s'ensuit une diffusion de cette représentation commune à l'ensemble des modules, appelée « broadcast » ; pour cela, chaque module dispose également d'un décodeur pour retranscrire le langage commun du GW dans son langage propre. On voit dans cette métaphore des langages propres et communs que l'opération du GW consiste principalement en une forme de traduction neuronale. Ainsi on peut faire appel à des outils d'apprentissage non supervisé [2], traditionnellement utilisés en traitement du langage (6) : pour entraîner les encodeurs et décodeurs qui permettent d'échanger l'information entre modules et GW, on s'appuie sur un cycle de traduction suivie de la traduction inverse. S'assurer qu'un tel cycle reste cohérent revient à aligner entre eux les espaces de représentation des modules et du GW, sans pour autant connaître la « vraie » traduction (si tant est qu'elle existe). Ceci implique que l'entraînement d'une architecture de Global Workspace nécessitera moins de données et moins de supervision explicite qu'un apprentissage classique (entièrement supervisé).

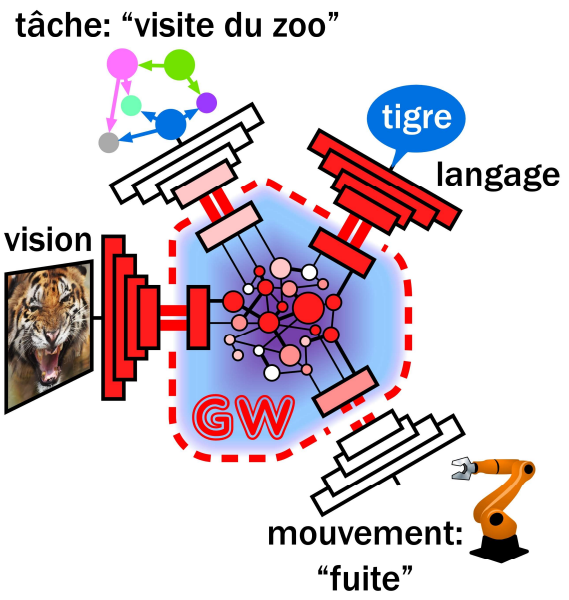


Figure 1: *Schéma illustratif du Global Workspace (GW) dans nos implémentations IA. Dans cet exemple, 4 modules spécialisés sont connectables au GW (en bleu). Ces modules pré-entraînés pour différentes fonctions (classification d'images, génération de texte, contrôle d'actions robotiques ou représentations de tâches) sont capables d'opérer indépendamment du GW. Cependant, un système attentionnel (non représenté ici) peut choisir de connecter chaque module au GW. Dans ce cas, l'information traitée par le module accède au GW, et ce faisant, devient elle-même accessible à l'ensemble des autres modules via la diffusion ou « broadcast ». Même un module qui n'est pour l'instant pas connecté au GW (dans cet exemple, le module de mouvement) a ainsi connaissance de la présence du tigre, traduit dans un format pertinent pour ce module (ici, la préparation d'une réponse motrice de fuite). Si le module était connecté, cette connaissance donnerait lieu à une réponse (de fuite) effective. Ainsi, le système entier a la capacité de contrôler son comportement et de l'adapter à la situation de manière flexible.*

Premiers résultats

En effet, nous avons pu confirmer l'intérêt d'apprendre une représentation multimodale via le GW, dans des scénarios simplifiés où seuls deux modules de représentation (vision, langage) sont connectés au GW (7,8). Par rapport à un entraînement supervisé classique, l'utilisation des cycles de traduction (reflétant le phénomène de broadcast) permet de diviser par un facteur de 5 à 10 le nombre de paires d'exemples [image, texte] requis, pour un niveau de performance équivalent (7). De plus, la représentation multimodale apprise au sein du GW permet d'atteindre de meilleures

[2] *L'apprentissage non supervisé s'oppose à l'apprentissage supervisé, où l'on indique directement au système la sortie désirée pour chaque entrée.*

performances quand on l'utilise ensuite pour des tâches de classification, par rapport à un système multimodal conventionnel tel que le célèbre modèle CLIP d'OpenAI (9) (CLIP : Contrastive Language-Image Pretraining).

De même, nous avons pu utiliser la représentation du GW pour entrainer un agent d'apprentissage par renforcement (« reinforcement learning ») à contrôler un robot dans un environnement industriel simulé (8). Grâce à sa méthode d'entrainement basée sur le broadcast, le système GW s'avère robuste à des changements de modalité d'entrée : par exemple, si le système apprend initialement à contrôler le robot sur la base d'informations numériques décrivant sa position dans l'environnement, il pourra ensuite généraliser la stratégie apprise lorsqu'on lui donnera des images en entrée au lieu des informations numériques (on parle de transfert « zero-shot »). Un système équivalent mais entraîné à la manière de CLIP (9) ne démontre pas ces capacités de généralisation et de robustesse.

Perspectives

Ces premiers éléments contribuent à établir que l'architecture GW apporte bien un avantage fonctionnel aux systèmes cognitifs artificiels. Cependant, de tels résultats nécessiteront d'être confirmés et étendus dans plusieurs directions. Tout d'abord, il faut considérer d'autres modalités d'entrée/sortie que l'image et le texte : par exemple, l'audio, la parole, et l'action ; mais aussi des processus internes comme l'émotion, la motivation, la mémoire… Apprendre à aligner ces diverses modalités via le GW (grâce aux cycles de traduction et au broadcast) reviendrait à appréhender les analogies qui existent naturellement entre elles et, ce faisant, à ancrer les représentations du monde dans leurs manifestations internes à l'agent ; et, symétriquement, à ancrer les représentations symboliques de l'agent dans un référentiel extérieur, i.e. dans l'environnement. On peut penser que c'est cet ancrage (ou « grounding ») qui fait défaut aux systèmes d'IA actuels, même les plus performants, comme les grands modèles de langues (LLMs ou Large Language Models). Ainsi, la modélisation du GW pourrait aussi avoir des conséquences pratiques en IA, en rapprochant les LLMs de la cognition humaine et en réduisant leurs erreurs invraisemblables ou « hallucinations ».

Un autre axe de recherche est l'exploration du mécanisme de sélection attentionnelle. Un module peut être sélectionné parce qu'il porte une information intrinsèquement saillante, ou parce qu'il semble bénéfique à un instant donné, ou encore, parce qu'il participe à une séquence d'opérations planifiée pour résoudre un problème complexe. Dans tous ces cas, il faut pouvoir entrainer (par apprentissage profond) un contrôleur attentionnel qui prendra correctement en charge cette sélection. Nous développons déjà des versions d'un tel contrôleur permettant de sélectionner parmi plusieurs modalités celle qui porte l'information la moins bruitée, ou encore de recruter itérativement un module d'incrément numérique pour résoudre une tâche d'addition. À terme, cette architecture générique devrait pouvoir adapter le comportement et les réponses de l'IA à sa situation et ses objectifs, à des horizons temporels plus ou moins lointains, c'est-à-dire manifester une aptitude cognitive de « système 2 » (par opposition à une cognition plus automatique et réflexive de « système 1 »).

Améliorer les capacités de l'IA est une manière claire de démontrer la viabilité de la théorie du Global Workspace. Une autre manière serait de vérifier que l'activité et le comportement de nos systèmes artificiels correspondent bien à ce qu'il se passe dans le cerveau. Nous collectons actuellement des données d'IRMf dans des conditions multimodales (image+texte) afin de comparer l'activité cérébrale à celle du GW (ainsi qu'à d'autres systèmes d'IA multimodale comme CLIP). Au niveau comportemental, nous tentons d'évaluer dans quelles conditions l'apprentissage multimodal du GW favoriserait l'émergence d'associations cross-modales non véridiques, ressemblant à des formes de synesthésie. Ces travaux pourraient engendrer des prédictions expérimentales, vérifiables par les Neurosciences.

Enfin, cette ligne de recherche pose également des questions philosophiques et éthiques non négligeables, dès lors qu'on se rappelle que le Global Workspace n'est pas uniquement une théorie de la cognition, mais est aussi considérée aujourd'hui comme une théorie majeure de la conscience. Doit-on penser qu'une implémentation IA donnerait lieu à une réelle expérience phénoménologique artificielle—probablement pas dans les systèmes rudimentaires que nous avons présentés ici, mais peut-être chez leurs descendants ? La question mérite considération (10).

rufin.vanrullen@cnrs.fr

Références

(1) Baars, B. J. (1993). *A cognitive theory of consciousness,* Cambridge University Press.

(2) Mashour G. A. et al. (2020). *Neuron*, 105(5), 776-798.

(3) Melloni L. et al. (2023) *PLoS ONE*, 18(2), e0268577.

(4) Hassabis D. et al. (2017) *Neuron*, 95(2), 245-258.

(5) VanRullen R., & Kanai R. (2021). *Trends Neurosci*, 44(9), 692-704.

(6) Artetxe M. et al. (2018). *Unsupervised neural machine translation. 6th International Conference on Learning Representations*, ICLR 2018.

(7) Devillers B. et al. (2025) *IEEE Trans Neural Netw Learn Syst*, 36(5):7843-7857. doi: 10.1109/TNNLS.2024.3416701

(8) Maytié L. et al. (2024). *Reinforcement Learning Journal* 3 (170), 1410-1426

(9) Radford A. et al. (2021). *Proceedings of the 38th International Conference on Machine Learning,* PMLR 139:8748-8763.

(10) Butlin P. et al. (2023) *arXiv preprint arXiv*:2308.08708.

Dossier

Neuroprothèses pour restaurer la marche après une lésion de la moelle épinière

FABIEN WAGNER

Équipe NeuroDTx, Institut des Maladies Neurodégénératives (CNRS UMR 5293), Université de Bordeaux

Introduction : les neuroprothèses motrices

Les lésions de la moelle épinière, ou lésions médullaires, perturbent la communication entre le cerveau et les circuits spinaux qui contrôlent le mouvement et intègrent les retours sensoriels, généralement situés en dessous du niveau de la lésion. La perturbation des différentes sources anatomiques du contrôle moteur descendant et des afférences sensorielles ascendantes peut entraîner une paralysie motrice totale ou partielle, mais permanente. Pendant des décennies, la récupération des fonctions motrices après une lésion médullaire chronique a été considérée comme impossible en raison de l'altération sévère de ces voies de communication bidirectionnelles.

Ces limitations ont mené au développement de technologies visant à restaurer la fonction motrice par stimulation électrique des structures impliquées dans la génération du mouvement (muscles, nerfs périphériques, moelle épinière ou cerveau). De telles technologies voient une application soit en tant que dispositif d'assistance (neuroprothèse), soit comme outil de réadaptation dans la cadre d'un parcours de soins.

La toute première neuroprothèse motrice a été inventée en 1961 par Liberson et collègues : il s'agissait d'un stimulateur du nerf péronier permettant de traiter le syndrome du pied tombant après une hémiplégie. Le terme neuroprothèse lui-même a été utilisé pour la première fois dans la littérature scientifique en 1971 pour désigner un implant intraspinal qui permettait la vidange de la vessie après une paraplégie. Depuis lors, la définition des neuroprothèses motrices a également été étendue aux technologies qui extraient les commandes motrices de signaux cérébraux afin de contrôler des dispositifs externes, également appelées interfaces cerveau-machine ("brain-machine interface" - BMI). Dans ce dernier cas, des algorithmes d'apprentissage automatique et d'intelligence artificielle (IA) sont utilisés afin de décoder les intentions motrices à partir des signaux cérébraux (voir également l'article de Guillaume Charvet, Fabien Sauter-Starace et Tetiana Aksenova dans ce dossier).

Dans cet article, nous passons brièvement en revue deux types de neuroprothèses motrices : la stimulation électrique fonctionnelle ("functional electrical stimulation" - FES) et la stimulation électrique de la moelle épinière ("spinal cord stimulation" - SCS) épidurale. Pour une revue plus détaillée du sujet, voir (1).

La stimulation électrique fonctionnelle

La stimulation électrique fonctionnelle est une méthode de neurostimulation cliniquement approuvée qui active les axones efférents innervant des muscles spécifiques pour produire un mouvement souhaité. La stimulation peut être délivrée à proximité du muscle ciblé ou sur un nerf moteur l'innervant, auquel cas on parle de stimulation nerveuse périphérique ("peripheral nerve stimulation" - PNS), en utilisant des électrodes non invasives, percutanées ou entièrement implantées. Ces électrodes sont ensuite connectées à un stimulateur électrique, qui peut généralement contrôler jusqu'à 16 canaux indépendants. Ces systèmes de stimulation peuvent simplement être utilisés pour renforcer la force musculaire, une approche nommée stimulation électrique neuromusculaire, ou pour assister dans des tâches fonctionnelles. De plus, la stimulation électrique fonctionnelle peut servir de technologie d'assistance, réduisant l'incapacité à exécuter un mouvement donné, ou faire partie d'une thérapie de réadaptation pouvant conduire à de la neuroplasticité et à des améliorations fonctionnelles, en fonction du trouble sous-jacent et de sa gravité.

La stimulation électrique fonctionnelle a été utilisée au cours des 60 dernières années pour faciliter des tâches motrices des membres supérieurs et inférieurs, telles que le maintien en position debout, la marche, l'atteinte d'objets et la préhension. Plusieurs systèmes approuvés par la “Food and Drug Administration” (FDA) aux Etats-Unis ou avec un marquage de Conformité Européenne (CE) en Europe sont désormais disponibles à des fins commerciales pour ces applications.

Cliniquement, la stimulation électrique fonctionnelle a été principalement utilisée pour l'hémiplégie résultant d'un accident vasculaire cérébral (AVC), par exemple comme stimulateur dans le syndrome du pied tombant, et pour la rééducation de la motricité du membre supérieur. Dans le contexte des lésions médullaires, la stimulation électrique fonctionnelle a également été largement testée dans des études de recherche, mais elle n'est pas encore devenue une pratique clinique standard. Dans l'ensemble, les systèmes de stimulation électrique fonctionnelle représentent la principale classe de neuroprothèses ayant été cliniquement acceptée, mais ils restent principalement réservés à la rééducation post-AVC. Cependant, ce type de neuroprothèse souffre d'une limitation conceptuelle, car elle recrute les axones moteurs dans un ordre non physiologique. Les axones moteurs de gros diamètre, qui produisent des forces importantes mais sont peu résistants à la fatigue, sont en effet recrutés en premier, ce qui rend difficile la génération de forces importantes pendant de longues périodes. Pour contourner cette limitation, des méthodes ciblant les afférences sensorielles et activant les circuits spinaux, telles que la stimulation médullaire, sont nécessaires.

Origine et redécouverte de la stimulation médullaire

Bien établie comme traitement de la douleur chronique, la stimulation médullaire épidurale, qui consiste à délivrer un courant électrique dans l’espace épidural de la moelle épinière, a récemment reçu une attention significative pour ses applications dans le contrôle moteur. La stimulation médullaire peut être délivrée via des électrodes implantées ou percutanées contenant jusqu'à 16 contacts et placées dans l'espace épidural au niveau postérieur de la moelle épinière. Ces électrodes sont ensuite connectées à un générateur de pulses qui peut être soit externe, soit implantable. Les effets moteurs de la stimulation médullaire épidurale sont principalement médiés par l'activation de fibres afférentes de gros diamètre situées dans les racines postérieures, qui recrutent à leur tour les groupes de motoneurones dans le segment spinal innervé.

La première observation selon laquelle la stimulation médullaire pourrait être utilisée comme neuroprothèse motrice après une lésion médullaire a été faite par Barolat et collègues en 1986 (2). Ceux-ci ont montré qu'un sujet avec une lésion médullaire incomplète (c'est-à-dire avec des fibres sensorielles ou motrices résiduelles traversant le site de la lésion) avait retrouvé un contrôle moteur volontaire avec la stimulation médullaire après plusieurs mois de stimulation (Figure 1). Une décennie plus tard, Dimitrijevic et collègues ont montré qu’une stimulation médullaire appliquée au niveau lombaire de la moelle épinière pouvait produire des réponses électromyographiques rythmiques associée à des mouvements de flexion-extension des jambes similaires à la marche (3). Ces résultats ont été obtenus chez six sujets ayant une lésion médullaire complète (c'est-à-dire sans fibres résiduelles traversant la lésion) et ont fourni des preuves indirectes de l'existence d'un réseau locomoteur spinal (“central pattern generator” - CPG) chez l'humain. Des études ultérieures ont montré que la variation de la fréquence de stimulation pouvait également produire une extension bilatérale des membres inférieurs.

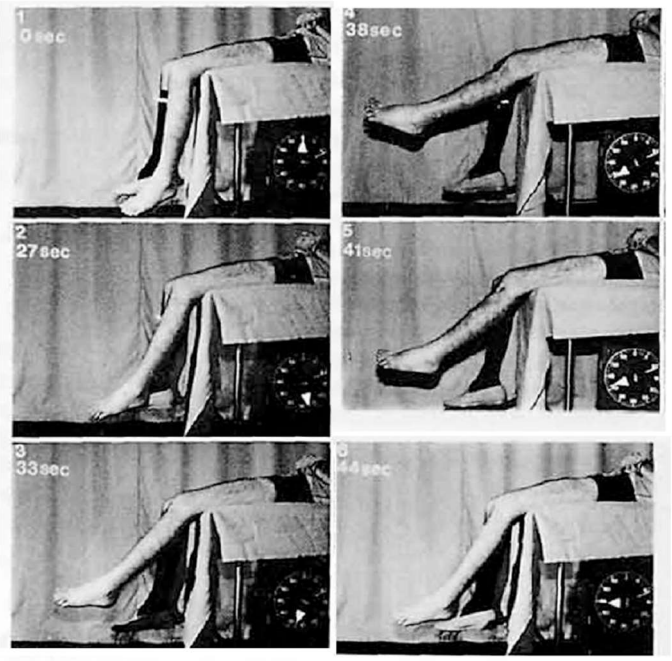


***Figure 1** : Photographies d'un individu ayant une lésion médullaire incomplète lors d'une tentative d'extension volontaire du genou gauche en présence de stimulation médullaire en 1986 (extrait de (2)).*

La première combinaison de la stimulation médullaire avec un entraînement locomoteur a été réalisée en 2002 chez un sujet ayant une lésion médullaire incomplète placé dans un système de support de poids corporel. La stimulation médullaire a entraîné une facilitation immédiate de la marche, qui s'est améliorée au cours de l'entraînement mais n'a pas été maintenue sans stimulation médullaire. La combinaison de la stimulation médullaire et de la réadaptation a été relancée en 2011 par Harkema et collègues, qui ont montré que la stimulation médullaire permettait de maintenir une

position debout indépendante après 80 sessions d'entraînement intensif chez un sujet avec une lésion médullaire complète (4). Cette étude a redécouvert que la stimulation médullaire peut permettre des mouvements volontaires des muscles paralysés, tout comme l'avait observé Barolat en 1986. Des résultats similaires ont ensuite été confirmés chez trois autres participants et ont été ultérieurement reproduits à la Mayo Clinic (Rochester, USA). Un participant ayant continué de s'entraîner avec la stimulation médullaire pendant presque 4 ans a même observé une récupération partielle des mouvements volontaires de ses jambes sans stimulation, indiquant le potentiel de cette technique pour promouvoir des mécanismes de neuroplasticité.

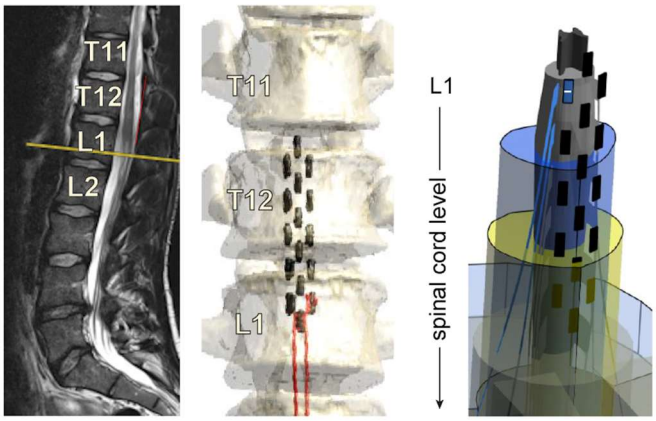


***Figure 2** : Champ de 16 électrodes implantées au niveau épidural postérieur de la moelle épinière lombaire chez un participant à l'étude (7), et reconstruction des différents volumes (jaune : graisse épidurale, bleu foncé : liquide céphalo-rachidien, gris : matière grise et blanche, bleu clair : racines postérieures).*

L'essor de la stimulation médullaire pour la marche

Un jalon dans l'application de cette technologie chez les blessés médullaires a été atteint en 2018 (Figure 2). Pour la première fois, trois groupes indépendants aux Etats-Unis et en Suisse ont démontré chez un total de six sujets que la stimulation médullaire combinée à une rééducation intensive pouvait restaurer une marche au sol indépendante (Figure 3). Au "Kentucky SCI center" (Louisville, USA), quatre blessés médullaires complets au niveau moteur ont utilisé une stimulation délivrée en continu (c'est-à-dire sans modulation des paramètres de stimulation au cours du temps) afin d'entraîner le maintien de la position debout, la marche sur tapis roulant avec soutien de poids, et enfin la marche au sol (5). Tous les participants ont réussi à atteindre une position debout assistée, et ont amélioré la stabilité du tronc en position assise en présence de stimulation. De plus, les deux participants ayant une lésion médullaire complète motrice et incomplète sensorielle ont acquis la capacité de marcher au sol avec stimulation médullaire continue et des dispositifs d'assistance après une période de 15 et 85 semaines respectivement. De même, à la "Mayo Clinic", un blessé médullaire sensorimoteur complet s'est entraîné à « marcher » en position allongée sur le côté, puis sur un tapis roulant et enfin au sol, le tout en présence de stimulation (6). Après 43 semaines d'entraînement et avec la stimulation médullaire en continu, ce participant a pu se tenir debout, faire des pas sur le tapis roulant et marcher au sol avec un déambulateur et l'aide d'un kinésithérapeute.

En parallèle, l'équipe de Courtine et Bloch à Lausanne a ouvert la voie à un nouveau paradigme appelé stimulation médullaire spatiotemporelle (7). Sur la base de travaux précliniques antérieurs chez le rat et le macaque, ils ont utilisé un générateur de pulses implantable (Activa RC, Medtronic) avec des capacités de contrôle en temps réel pour développer un protocole de stimulation alternant entre les phases de balancement, d'acceptation du poids et de propulsion au cours du cycle de marche. Ces fonctions ont été ciblées par des configurations d'électrodes spécifiques, déclenchées soit à un rythme prédéfini, soit en temps réel par les capacités motrices résiduelles. Cette stimulation médullaire dite spatiotemporelle a permis une facilitation immédiate de la

***Figure 3** : Applications de la stimulation médullaire dans un contexte de rééducation motrice, afin de faciliter différentes tâches fonctionnelles (position assise, debout, marche avec assistance et / ou déambulateur).*

marche et une récupération à long terme de la fonction motrice chez trois sujets avec une lésion médullaire incomplète chronique. Cette approche a ensuite été étendue à un total de neuf participants, dont trois patients avec une lésion médullaire motrice complète, implantés avec une nouvelle matrice d'électrodes adaptée à la fois aux fonctions motrices des jambes et du tronc (8). La stimulation médullaire a permis de restaurer non seulement la marche mais également la stabilité du tronc.

Autres applications et perspectives

Dans l'ensemble, ces résultats ont suscité un vif intérêt pour l'intégration de la stimulation de la moelle épinière dans les protocoles de rééducation après une lésion médullaire. D'autres applications ont également vu le jour, comme l'utilisation de la stimulation médullaire au niveau lombaire dans la maladie de Parkinson pour réduire les symptômes de « freezing » ("freezing of gait"), caractérisés par un blocage de la marche temporaire et involontaire, qui entraîne des chutes fréquentes. L'application de la stimulation médullaire a ainsi permis, chez un premier patient parkinsonien atteint de « freezing », de réduire ces troubles et de promouvoir une démarche fluide (9).

Récemment, la stimulation de la moelle épinière a également été couplée à une interface cerveau-machine afin que les profils de stimulation soient contrôlés directement par les signaux corticaux, donnant au patient un contrôle plus naturel et davantage d'autonomie (10). Dans ce cas, un algorithme d'IA est utilisé pour décoder jusqu'à 7 états discrets : la flexion de la hanche, du genou et de la cheville de manière bilatérale, ainsi que l'état de repos. Le résultat de ce décodage est ensuite utilisé pour contrôler les configurations d'électrodes et paramètres de stimulation optimaux pour faciliter le mouvement décodé (voir également l'article de Guillaume Charvet, Fabien Sauter-Starace et Tetiana Aksenova dans ce dossier). La stimulation médullaire, par sa capacité à activer la moelle épinière elle-même et pas seulement les fibres efférentes et les muscles, se situe donc à un point stratégique pour promouvoir la récupération motrice dans plusieurs pathologies. Ce domaine de recherche se dirige actuellement vers des études à plus grande échelle démontrant l'efficacité de ces technologies invasives, ainsi que vers le développement d'alternatives non-invasives telles que la stimulation spinale transcutanée ("transcutaneous spinal cord stimulation" - tSCS). Le futur de la réadaptation s'avère donc particulièrement stimulant.

fabien.wagner@u-bordeaux.fr

Références

(1) Seáñez I. et al. (2022) In: *Reinkensmeyer DJ, Marchal-Crespo L, Dietz V, editors. Neurorehabilitation Technology* pp 369–399. https://doi.org/10.1007/978-3-031-08995-4_18

(2) Barolat G. et al. (1986) *Appl Neurophysiol* 49(6):307–314.

(3) Dimitrijevic MR. et al. (1998) *Ann N Y Acad Sci* 860:360–376.

(4) Harkema S. (2011) *Lancet* 377:1938–1947.

(5) Angeli CA. et al. (2018) *N Engl J Med* 379(13):1244–1250.

(6) Gill ML. et al. (2018) *Nat Med* 24(11):1677–1682.

(7) Wagner FB. et al. (2018) *Nature* 563(7729):65–71.

(8) Rowald A. et al. (2022) *Nat Med* 28(2):260–271.

(9) Milekovic T. et al. (2023) *Nat Med* 29(11):2854–2865.

(10) Lorach H. et al. (2023) *Nature* 618(7963):126–33. doi: 10.1038/s41586-023-06094-5.

Dossier

Interface cerveau-machine pour la restauration du mouvement: l'Intelligence Artificielle au cœur du décodage de l'activité corticale

GUILLAUME CHARVET, FABIEN SAUTER-STARACE, TETIANA AKSENOVA

Université Grenoble Alpes, CEA, Leti, Clinatec, Grenoble

Introduction

Fortes de premières preuves de concept cliniques validées, les interfaces cerveau-machine (ICM) constituent un espoir pour des millions de personnes porteuses d'un handicap moteur sévère et chronique. En effet, les ICM motrices portent l'espoir de restaurer la mobilité des membres, offrant ainsi aux patients une plus grande autonomie et une amélioration significative de leur qualité de vie.

Les ICM motrices permettent ainsi de proposer une solution alternative de contrôle des membres pour les patients atteints de handicap moteur sévère comme la lésion partielle ou complète de la moelle épinière (ex : tétraplégie et paraplégie), la sclérose latérale amyotrophique ou d'accidents vasculaires cérébraux (Voir également l'article de Fabien Wagner dans ce dossier). Ce concept d'ICM remonte à 1973 (1), lorsque des chercheurs ont proposé de contrôler divers systèmes en décodant l'activité cérébrale. L'ICM permet ainsi de créer une liaison directe entre le cerveau et un système de suppléance fonctionnelle motrice permettant ainsi à un individu de le contrôler sans passer par le système nerveux périphérique ou les muscles. Ainsi, lorsque le handicap moteur n'est pas lié à une dégénérescence ou un traumatisme des neurones du cortex impliqués dans le contrôle moteur des membres, l'activité cérébrale motrice est intacte. De plus, le fait d'imaginer un mouvement ou l'exécuter provoque une activité électrique similaire au niveau du cortex moteur. Cette activité cérébrale peut ainsi être mesurée, et analysée pour décoder les intentions de mouvement de l'individu. Celui-ci pourra ainsi, après réalisation de tâches mentales d'imagination de mouvement, contrôler un effecteur comme par exemple un bras robotique, un exosquelette ou un stimulateur électrique des muscles ou de la moelle épinière.

Une ICM doit ainsi avoir la capacité de décoder en temps réel l'activité cérébrale mesurée afin de prédire les intentions de mouvements d'un individu pour piloter un effecteur. De plus, l'usage dans le cadre de la vie quotidienne de ces ICM nécessite le développement de méthodes de décodage précises, robustes et stables sur plusieurs mois. Pour cela, de nombreuses méthodes de décodage à base d'Intelligence Artificielle (IA) ont été développées par de nombreuses équipes de recherche dans le monde. Ces méthodes dépendent de la qualité du signal mesuré (via des électrodes d'électroencéphalographie (EEG), corticale (ECoG) ou intracérébrales (*Micro Electrode Array* - MEA) et de l'application visée.

En particulier, l'utilisation de l'électrocorticographie (ECoG) couplée à des algorithmes d'apprentissage automatique permet de décoder avec une précision croissante l'activité corticale liée aux mouvements. Ces technologies ouvrent des perspectives prometteuses pour la restauration de la motricité chez les patients en situation de handicap moteur. Cet article explore les avancées récentes des ICM basées sur l'ECoG et leur potentiel pour redonner de la mobilité aux patients.

L'électrocorticographie : une technologie clé pour l'interface cerveau-machine

L'ECoG est une technique de mesure de l'activité cérébrale qui repose sur l'implantation d'électrodes directement sur la surface du cortex. Contrairement à l'électroencéphalographie (EEG), qui enregistre l'activité cérébrale depuis le cuir chevelu, l'ECoG offre une

meilleure résolution spatiale et une stabilité des signaux sur plusieurs mois, voire plusieurs années.

L'EEG, bien que non invasif, impose une contrainte majeure : la pose d'un casque avant chaque utilisation entraîne un temps d'installation long et un entrainement fréquent des modèles de décodage, limitant ainsi son emploi régulier. À l'opposé, les MEA sont des matrices d'électrodes intracorticales qui, bien qu'elles offrent une résolution extrêmement fine permettant ainsi de contrôler de manière précise des systèmes robotiques d'assistance tel qu'un bras robotisé (2), de restaurer des sensations tactiles (3), ou de faciliter la communication (4), présentent des problèmes de stabilité signaux dans le temps provoquant des problèmes de stabilité des modèles de décodage, nécessitant un re-apprentissage fréquent. L'ECoG représente donc un compromis entre ces deux méthodes, alliant une précision satisfaisante à une stabilité accrue dans le temps.

Les électrodes ECoG captent les signaux électriques générés par les neurones à la surface du cortex moteur, qui sont ensuite traités par des algorithmes d'apprentissage automatique basés sur des méthodes de *machine learning* ou *deep learning*. Grâce à la stabilité des signaux ECoG, les modèles de décodage nécessitent moins de phases d'entrainement régulières, facilitant ainsi le contrôle efficace et pérenne de dispositifs dans le cadre d'un usage par les patients au quotidien.

L'apport de l'intelligence artificielle dans le décodage de l'activité cérébrale

L'IA joue un rôle essentiel dans le décodage des signaux cérébraux captés par l'ECoG. Les algorithmes de *deep learning*, notamment les réseaux de neurones convolutifs et récurrents, ainsi que les méthodes de *machine learning* basées sur des modèles récursifs, permettent de décoder avec précision les intentions motrices tout en garantissant une latence faible. En effet, pour une utilisation efficace en conditions réelles, le patient ne doit pas percevoir de décalage entre son intention de mouvement et l'exécution de celui-ci.

Les méthodes d'apprentissage supervisé sont souvent utilisées pour entraîner l'ICM (entrainement de modèles de décodage) à reconnaître les caractéristiques de l'activité neuronale associées à des intentions de mouvements spécifiques. Par exemple, un patient peut être invité à imaginer un mouvement de la main, tandis que l'algorithme analyse les signaux cérébraux correspondants et ajuste les paramètres du modèle de décodage pour améliorer sa capacité de prédiction, le tout en temps réel.

Des recherches récentes ont également exploré l'utilisation de techniques d'apprentissage non-supervisé qui permettent de réduire la dépendance aux données annotées et d'accroître la robustesse des modèles face à l'évolution de l'activité cérébrale au cours du temps, tout en optimisant la vitesse de traitement pour assurer une interaction fluide avec les effecteurs de suppléance fonctionnelle.

Application et cas concret : évaluations cliniques et usages d'une interface cerveau-machine chronique (WIMAGINE)

Les ICM basées sur l'ECoG offrent ainsi une solution prometteuse pour le contrôle direct d'effecteurs par l'activité cérébrale. Parmi ces avancées, le système WIMAGINE, développé par le CEA, s'est illustré par des preuves de concept cliniques majeures (Figure 1). Il a notamment permis à un patient tétraplégique de contrôler un exosquelette à quatre membres, un bras robotisé ou un fauteuil roulant (5,6) et a contribué à la restauration de la marche chez un patient paraplégique grâce au contrôle cérébral d'un stimulateur électrique de la moelle épinière (7), ouvrant ainsi de nouvelles perspectives pour la suppléance fonctionnelle et la neuro-réhabilitation

L'implant du système WIMAGINE (8) est un dispositif sans fil de mesure de signaux corticaux. L'implant est usuellement placé au-dessus du cortex moteur et sensitif, et plus précisément sur la dure-mère. Le système WIMAGINE présente ainsi la capacité d'enregistrer l'activité cérébrale ECoG avec une résolution spatiale de l'ordre de 4 mm (64 électrodes) et une excellente stabilité dans le temps (9), mais également il permet de décoder un grand nombre de degrés de liberté avec une latence faible entre l'intention de mouvement et l'exécution. Près de sept ans après les premières implantations, les systèmes WIMAGINE implantés présentent toujours des performances stables d'enregistrement des signaux ECoG.

Des algorithmes d'IA basés sur des méthodes de *machine learning* ont été développés (10,11) afin de réaliser un décodage en temps réel de l'activité cérébrale ECoG. Pour adresser le défi de l'usage à terme de cette technologie par les patients dans leur vie quotidienne, les algorithmes à base d'IA ainsi développés possèdent un certain nombre de caractéristiques. La création du modèle de décodage se fait de manière adaptive et en temps réel, ce qui raccourci considérablement les périodes d'apprentissage du modèle. Le décodage est réalisé de manière asynchrone, c'est à dire que le système est capable de décoder pendant les périodes de contrôle et des périodes de non-contrôle (ou état de repos) permettant ainsi un fonctionnement avec le minimum

d'assistance. Ces algorithmes permettent également de réaliser un décodage multi-membres (par exemple pour le contrôle des 2 bras et des jambes d'un exosquelette, ou le contrôle du mouvement des deux jambes grâce à un stimulateur au niveau de la moelle épinière). La qualité des signaux ECoG mesurés grâce aux implants WIMAGINE (sans aucun artefact lié aux mouvements) permet d'assurer la robustesse du décodage. Enfin, la stabilité dans le temps des signaux ECoG mesurés et la robustesse des algorithmes permettent de conserver un modèle de décodage sur plusieurs mois, ce qui représente un critère important pour l'usage d'une ICM dans le cadre de la vie quotidienne.

Enfin, cette technologie est utilisée actuellement de manière autonome par un patient dans le cadre de sa vie quotidienne (7), offrant un retour d'expérience précieux pour identifier et anticiper les axes d'améliorations. Les futurs développements viseront à optimiser son utilisation au quotidien en renforçant la précision et la vitesse du décodage, en améliorant la portabilité et l'autonomie du système, et en garantissant un usage sans assistance sur plusieurs années. Cela nécessitera notamment le développement de modèles de décodage capables de s'adapter à l'évolution de l'activité cérébrale des patients au cours du temps, assurant ainsi une performance stable et durable.

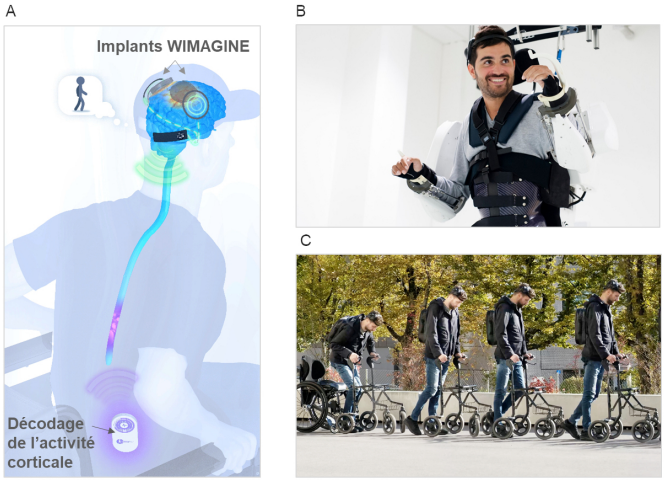


***Figure 1**: Technologie d'ICM WIMAGINE et évaluations cliniques : A) Implants WIMAGINE d'enregistrement de signaux corticaux et décodage des intentions de mouvement B) Evaluation clinique (CEA, CHUGA) de contrôle cérébral d'un exosquelette 4-membres C) Evaluation clinique (EPFL, CHUV, CEA, ONWARD) d'une interface cerveau-moelle épinière pour la restauration de la marche chez un patient paraplégique. © photos: LaBrêche/Juliette Treillet, Jimmy Ravier/EPFL*

Défis et perspectives

Les ICM ont connu des avancées majeures au cours des deux dernières décennies, démontrant leur potentiel pour la suppléance fonctionnelle motrice chez les patients atteints de handicaps moteurs sévères. Toutefois, leur intégration dans la vie quotidienne reste un défi, car les solutions technologiques actuelles nécessitent encore de nombreuses améliorations, tant en termes de performances techniques que de coût, de facilité d'usage et d'acceptabilité par les patients.

Parmi les principaux défis à relever, l'optimisation des algorithmes de décodage constitue un enjeu central. Il est essentiel d'améliorer leur précision, leur robustesse et leur stabilité afin de permettre la réalisation de tâches complexes du quotidien telles que manger, se laver ou travailler.

Sur le plan matériel, le développement de dispositifs implantables moins invasifs, offrant une résolution optimale de mesure de l'activité cérébrale et des capacités avancées de stimulation électrique, est un axe de recherche clé. Plusieurs équipes à travers le monde, dont *Neuralink, Synchron, Paradromics, Blackrock Neurotech, Precision Neuroscience, Cortec, Onward medical,* travaillent déjà à la mise au point de ces nouvelles générations d'implants. Par ailleurs, la miniaturisation des systèmes constitue un levier essentiel pour améliorer l'ergonomie et l'accessibilité des ICM. Le développement d'algorithmes de décodage embarqués directement sur des puces électroniques permettrait de réduire la taille des systèmes, d'optimiser la consommation énergétique et de faciliter l'utilisation des dispositifs dans un cadre quotidien. Cela représente en particulier l'un des défis adressés dans le cadre du projet européen NEMO-BMI coordonné par le CEA.

Enfin, l'avenir des ICM ne se limite pas à la suppléance fonctionnelle. En effet, elles pourraient jouer un rôle déterminant en neuro-réhabilitation dans la récupération motrice grâce à des solutions innovantes combinant effecteurs de rééducation motrice et interfaces cerveau-machine. Pour répondre aux besoins des patients, il sera crucial de proposer des effecteurs adaptés, qu'il s'agisse d'exosquelettes ou de technologies avancées de stimulation des muscles et de la moelle épinière.

Malgré de premières avancées significatives, de nombreux défis restent à relever pour faire des ICM une solution véritablement viable et accessible au quotidien. Leur succès reposera non seulement sur les progrès en neurosciences, en intelligence artificielle et en ingénierie biomédicale, mais aussi sur la capacité à transférer ces innovations vers l'industrie afin de garantir leur diffusion à grande échelle et leur adoption par le plus grand nombre.

Guilaume.charvet@cea.fr
Fabien.sauter@cea.fr
Tetiana.aksenova@cea.fr

Références

(1) Vidal J. J. (1973) *Annu Rev Biophys Bioeng* 2, 157–180. doi: 10.1146/annurev.bb.02.060173.001105

(2) Collinger J. L. et al. (2013). *Lancet*, 381(9866), 557-564.

(3) Flesher S. N. et al. (2021) *Science*, 372(6544), 831-836.

(4) Willett F. R. et al. (2023) *Nature*, 620(7976), 1031-1036.

(5) Benabid A. L. et al. (2019) *Lancet Neurol*, 18(12), 1112-1122.

(6) Bellicha A. et al. (2025) *J Neural Eng*, 22(1). doi: 10.1088/1741-2552/adae36. PMID: 39854845.

(7) Lorach H. et al. (2023). *Nature*, 618(7963):126-133. doi: 10.1038/s41586-023-06094-5.

(8) Mestais C. et al. (2015) *IEEE Trans Neural Syst Rehabil Eng,* 23(1):10-21. doi: 10.1109/TNSRE.2014.2333541.

(9) Larzabal C. et al. (2021) *J Neural Eng*, 18(5). doi: 10.1088/1741-2552/ac2003. PMID: 34425566.

(10) Eliseyev A. et al. (2017) *Sci Rep* 7(1), 16281.

(11) Moly A. et al. (2022). *J Neural Eng*, 19(2), 026021.

Dossier

Contrôle sensorimoteur naturel d'un bras artificiel

AYMAR de RUGY

Univ. Bordeaux, CNRS, INCIA, UMR 5287, F-33000 Bordeaux, France

Un contrôle naturel de prothèse en réalité virtuelle

Lorsqu'ils sont entrainés sur des données suffisamment riches, les réseaux de neurones artificiels sont capables de reproduire les comportements humains les plus probables, une vertu que certains pourront qualifier d'intelligence. Nous avons déjà montré (1, 2) qu'une intelligence artificielle de ce type est capable de reproduire la configuration des articulations manquantes chez une personne amputée du bras, d'une manière qui permet à cette dernière de saisir et déplacer des objets aussi facilement qu'avec un bras naturel dans un environnement virtuel (Figure 1, Vidéo 1 et 2). En effet, un réseau de neurones entrainé sur une base de mouvements naturels parvient à prédire les articulations du coude jusqu'au poignet à partir des seuls mouvements d'épaule d'un utilisateur et de la position et orientation dans l'espace de l'objet que celui-ci souhaite saisir. Le passage de cette preuve de concept en situation de réalité virtuelle à un monde physique soulève cependant de nombreuses difficultés. Dans les sections suivantes, nous illustrons comment certaines de ces difficultés peuvent être surmontées par des approches bio-inspirées.

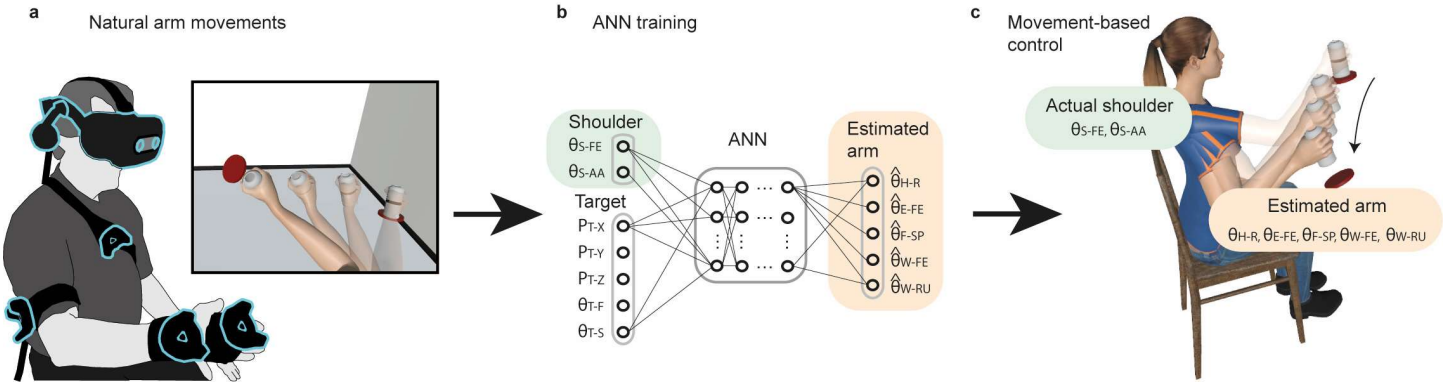


***Figure 1 : Contrôle à base de mouvements résiduels et d'information contextuelle**. **a**. Des mouvements naturels de bras sont enregistrés sur des participants impliqués dans la prise et le placement d'une bouteille dans un large espace en réalité virtuelle. **b**. Un réseau de neurones artificiels (Artificial Neural Network, ANN) est entraîné sur ces mouvements pour reconstruire les degrés de liberté distaux (orange) à partir des degrés de liberté proximaux (vert) et des informations sur la cible du mouvement (position et orientation de la bouteille). **c**. Contrôle de prothèse basé sur le mouvement : le participant exécute la même tâche à l'aide d'un bras hybride reproduisant en temps réel ses propres mouvements d'épaule (angles verts), et en utilisant les prédictions du réseau pour les articulations restantes de bras (angles orange). D'après (1).*

***Vidéo1** : https://www.youtube.com/watch?v=XPIlkrjwTtc Exemple de participant effectuant la tâche de prise et déplacement d'objet avec son bras valide en environnement virtuel, en téléopération simple.*

***Vidéo2** : https://www.youtube.com/watch?v=Utoa9aYWRK0 Exemple de patient amputé transhuméral effectuant la même tâche sans difficulté avec notre nouveau contrôle de prothèse*

Un premier test en robotique de téléopération

Le robot humanoïde Reachy, développé par la compagnie Pollen Robotics pour pouvoir être commandé en vision égocentré, nous a permis de tester un nouveau contrôle de prothèse dans le monde physique. Sur ce dispositif (Figure 2, Vidéo 3), l'opérateur voit à travers les yeux du robot dont la vue est projetée dans le casque de réalité virtuelle qu'il porte, et auquel les mouvements de la tête du robot sont asservis. Après avoir identifié la position de l'objet que l'opérateur souhaite saisir à l'aide d'outils de vision artificielle guidée par le regard (3), celui-ci contrôle l'épaule de Reachy par ses propres mouvements d'épaule, alors que le reste du bras est contrôlé par les prédictions du réseau de neurones entrainé préalablement sur mouvements naturels. L'efficacité du contrôle est ensuite démontrée par des performances et des évaluations d'utilisabilités comparables obtenues chez des participants valides et amputés (4).

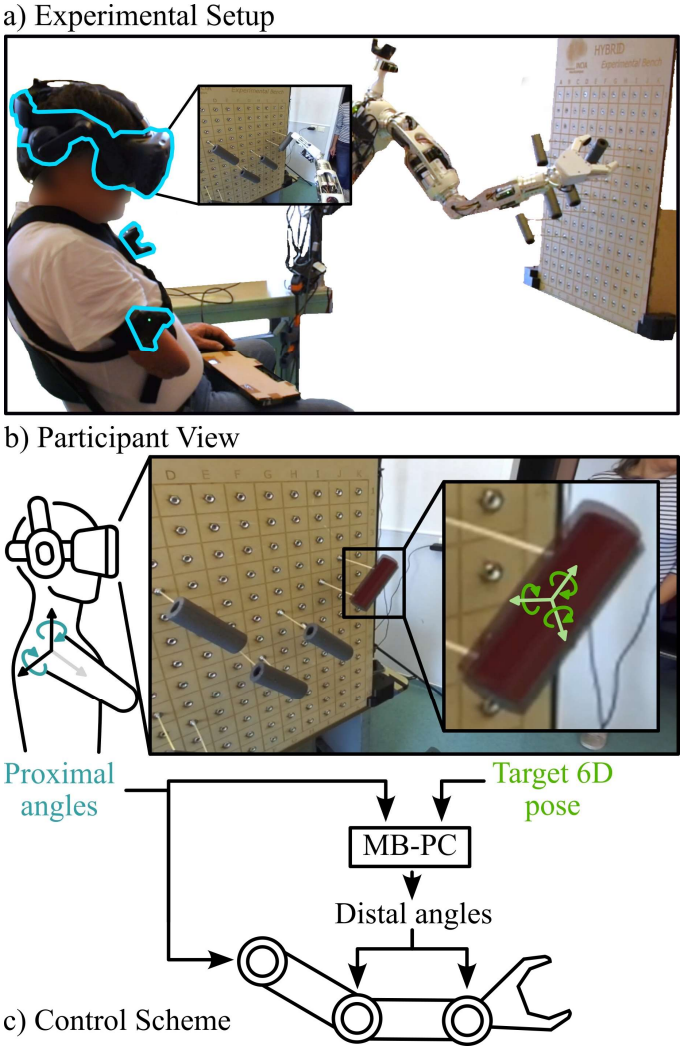


Figure 2. Téléopération en vision égocentré du robot Reachy avec contrôle de prothèse sur les articulations distales du bras. ***a****. Le patient amputé équipé du casque et des traqueurs de réalité virtuelle commande le bras du robot Reachy pour attraper des cylindres en mousse en face de lui.* ***b****. vue du participant, avec la position et l'orientation du cylindre qu'il doit attraper identifiées par vision artificielle aidée de son regard (cf référentiel vert).* ***c****. Le réseau de neurones appelé « Movement-Based Prosthesis Control » (MB-PC) prédit les mouvements distaux sur la base de la pose en 6 dimensions (3 positions et 3 orientations) cible et des mouvements d'épaule, et applique ces derniers sur le bras robotiques. D'après (4).*

Vidéo 3 *: https://www.youtube.com/watch?v=OUepbeQpBOg*
Principe de contrôle et illustration de l'étude sur patients amputés et valides téléopérant le robot Reachy.

Une transformation sensorimotrice bio-inspirée

Alors que l'intention de mouvement obtenue par vision artificielle aidé du regard peut être accessible dans le référentiel de la tête de l'opérateur, le contrôle de prothèse illustré précédemment opère dans un référentiel centré sur son épaule. Si le passage d'un référentiel à l'autre ne pose pas de difficulté en réalité virtuelle ni sur Reachy, où l'on connait précisément la position des éléments virtuels ainsi que la chaine cinématique, cette transformation est plus difficile à obtenir en situation réelle du fait de la biomécanique complexe de la nuque et de l'épaule. Inspirée des transformations sensorimotrices et du codage spatial multisensoriel par population de neurones (5), nous proposons une méthode effectuant la transformation tête-épaule nécessaire (6), et l'adaptant en temps réel aux mouvements et à la morphologie

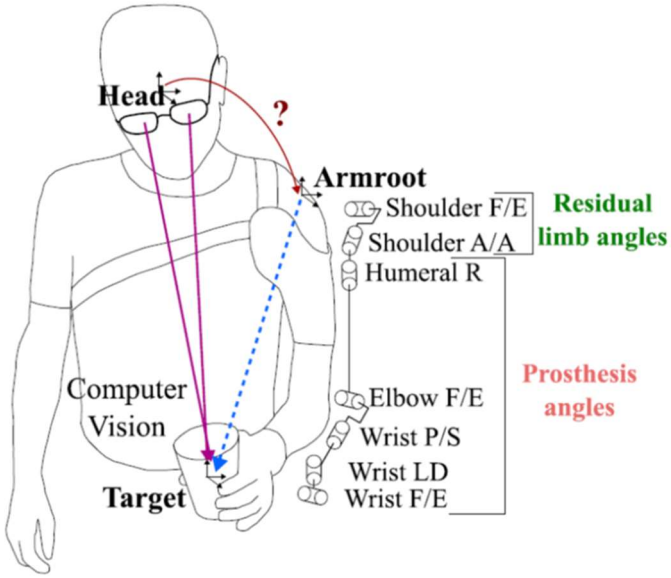


Figure 3 : Transformation de référentiel de la tête à l'épaule. *Partant de la position de la main disponible visuellement dans le référentiel de la tête, la méthode remonte au référentiel de l'épaule à partir de la cinématique de la prothèse, et code la relation tête-épaule dans une carte spatiale bio-inspirée. D'après (6).*

Vidéo 4 : *https://www.youtube.com/watch?v=SQNKJrkQBXE*
Adaptation de la carte spatiale bio-inspirée pour la transformation de référentiel de la tête à l'épaule, illustrée lors de la télé opération en vision égocentré du robot Reachy.

spécifique de chaque participant. Cette méthode part de la position de la main prothétique disponible visuellement dans le référentiel de la tête, remonte au référentiel de l'épaule à partir de la cinématique connue de la prothèse (proprioception artificielle), et code la relation tête-épaule ainsi obtenue dans une carte spatiale adaptative bio-inspirée (Figure 3 et Vidéo 4). Une expérience sur participants valides et amputés a ensuite permis de montrer l'efficacité de cette méthode pour le bon fonctionnement de notre contrôle de prothèse (6).

Vers l'intégration des régulations spinales dans le contrôle de bras artificiel

Si les développements illustrés précédemment montrent qu'il est en principe possible de contrôler une prothèse de bras pour attraper des objets, les interactions physiques avec le monde requièrent souplesse et robustesse vis-à-vis de perturbations. Dans ce contexte, on sait que les propriétés mécaniques du système musculo-squelettique combinés aux boucles sensorimotrices de bas niveau dans la moelle épinière sont essentielles à la production de mouvements coordonnés, et aux réponses fonctionnelles observées en cas de perturbations (7, 8). Nous avons montré par simulation qu'un réseau spinal couplé à un modèle musculo-squelettique réaliste du bras est capable de générer intrinsèquement les activités musculaires et neuronales complexes responsables de la production de mouvements coordonnés (Figure 4) (9). De simples commandes à palier, réglant l'activité de base des neurones spinaux (SET) et le déclenchement du mouvement (GO), suffisent en effet pour que le système produise seul les activités nécessaires, en vertu de la transformation en continu des afférences sensorielles qu'il génère et reçoit en commandes motrices adaptées. Un problème majeur reste cependant de trouver les commandes simples (SET et GO) permettant à ce système de générer des mouvements valides, l'espace des paramètres étant trop large pour une exploration systématique, et trop complexe pour les algorithmes d'optimisation à base de gradient. Ce verrou a été levé en utilisant une technique d'exploration à base de curiosité inspirée du babillage des enfants, proposée en robotique développementale et apprentissage automatique inspiré

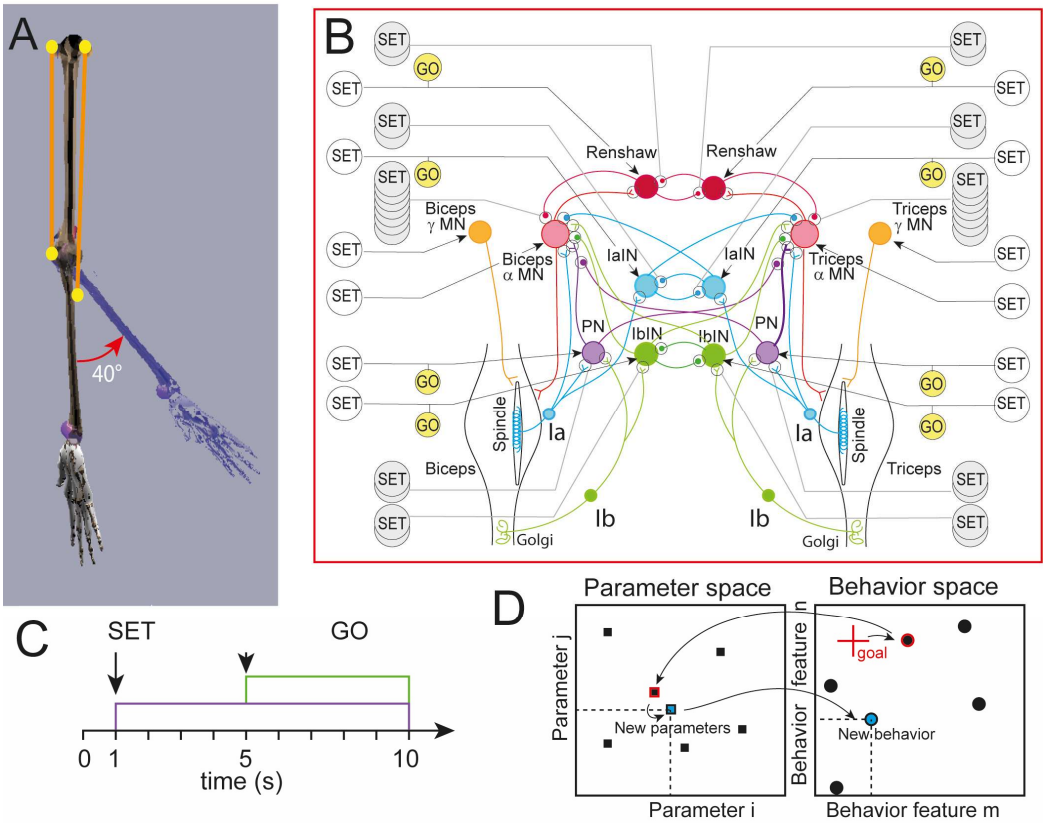


Figure 4. Modèle neuro-mécanique et principe de contrôle par régulation spinale. A. *Système musculo-squelettique.* ***B****. Système sensorimoteur, comprenant les principaux motoneurones, interneurones et leurs afférences sensorielles.* ***C****. Commandes simples SET et GO à palier utilisées pour contrôler le réseau sensorimoteur. Les commandes SET ajustent les niveaux d'activité des neurones pour amener le réseau dans un état prêt à déclencher le mouvement (sans produire le mouvement lui-même). Les commandes GO lancent le mouvement. Les commandes SET et GO sont les paramètres de l'espace des paramètres.* ***D****. Algorithme d'exploration « Goal Exploration Process (GEP). Pour un nouvel objectif demandé dans l'espace des comportements (croix rouge), l'algorithme recherche le comportement valide le plus proche, récupère les paramètres associés à ce comportement, et les rejoue en appliquant une petite variation aléatoire, de manière à produire un nouveau comportement. D'après (9).*

par la biologie et les neurosciences (10). A partir de petites variations aléatoires appliquées dans l'espace des paramètres, l'algorithme identifie de proche en proche une variété de comportements physiologiques, i.e., reproduisant une trajectoire de moindre secousse (*minimum jerk*) optimal d'un point de vue énergétique, avec un patron d'activité musculaire classiquement observé chez l'homme et l'animal (*triphasic muscle pattern*) (9). Branché en permanence sur le comportement qu'il produit, ce système est en mesure de répondre au plus vite à un imprévu comme le font nos réflexes spinaux (8), et devrait être capable de reproduire fidèlement les régulations spinales en réponse à des perturbations.

Bien que des articulations à raideur variable pilotées par des muscles artificiels intégrant des éléments élastiques existent en robotique bio-inspirée (11), leur contrôle avec les outils classiques de robotique reste un problème majeur. Les récents progrès en intelligence artificielle fournissent des outils permettant de composer avec les non-linéarités de ces systèmes, mais les réseaux de neurones artificiels profonds qu'ils impliquent sont très couteux en calcul et en énergie, représentent une approche « boite noire » difficilement interprétable et généralisable, et sont souvent couplés à des apprentissages par renforcement eux-mêmes extrêmement gourmands en expériences et données d'apprentissage. Nous pensons que le principe de contrôle bio-inspiré illustré ici offre un fort potentiel pour l'intégration des régulations spinales et des propriétés de l'appareil musculo-squelettique en robotique et pour les interfaces homme-robot.

Conclusion

Si les approches présentées ici illustrent l'intérêt des outils d'intelligence artificielle pour le contrôle de prothèses et de bras robotiques, elles montrent également que la compréhension fine des mécanismes sensorimoteurs naturels est indispensable à leur bonne intégration dans le contrôle d'un bras artificiel. Dans ce contexte, nous sommes persuadés que les étapes de modélisation et de test visant à cette bonne intégration sont de nature à améliorer de manière interactive notre compréhension des mécanismes de contrôle sensorimoteur naturels.

aymar.derugy@u-bordeaux.fr

Références

(1) E. Segas et al. (2023) *eLife,* 12: p. RP87317.

(2) E. Segas, et al. (2024) *Biomimetics,* 9 (9):9.

(3) I. González-Diaz et al. (2024) *IEEE J. Biomed. Health Inform,* p 1-17.

(4) V. Leconte et al. (2024) *TechRxiv* doi: 10.36227/techrxiv.172253934.48844352/v2

(5) A. Pouget et al. (2002) *Nat Rev Neurosci*, 3(9):741-747.

(6) B. Lento, V. (2024) *IEEE Robot Autom Lett*, 9(9):7875-7882.

(7) G. Raphael et al. (2010) *J. Neurosci*, 30(28):9431-9444.

(8) J. Weiler et al. (2019) *Nat Neurosci,* 22(4):529-533.

(9) D. Cattaert et al. (2023) *bioRxiv*, doi: 10.1101/2023.12.22.572982.

(10) J. Gottlieb & P.-Y. Oudeyer (2018) *Nat. Rev. Neurosci* 19(12):758-770.

(11) C. Richter et al. (2016) *IEEE Robot. Autom. Mag*. 23(4):128-137.

Dossier

Modélisation et exploration de graphes d'IRMs cérébrales sans *a priori* anatomiques

ANTOINE BOURLIER[1,2], MOHAMED SLIMANE[2], JEAN-YVES RAMEL[3], ELODIE CHAILLOU[1]

[1], INRAE, CNRS, Université de Tours, UMR PRC. [2], Université de Tours, LIFAT.
[3], Université Savoie-Mont-Blanc, LISTIC

1. Introduction

Un des objectifs de l'exploration des IRMs (imagerie par résonance magnétique) cérébrales est d'extraire divers critères morphologiques et structuraux de l'ensemble ou d'une partie de l'encéphale. Pour se faire, plusieurs étapes automatisées de prétraitements permettent d'améliorer la qualité et l'interprétabilité des images. Ces étapes incluent notamment le débruitage, le recalage et la normalisation en intensité. Souvent ces prétraitements précèdent la segmentation et la labellisation des régions d'intérêt (ROIs). La segmentation, basée sur le contraste des intensités des voxels de l'image et sur des connaissances expertes, peut être réalisée manuellement (1) ou automatiquement (2). Lorsque la segmentation est manuelle, elle est extrêmement chronophage et très dépendante des opérateurs, même s'ils sont experts en neuroanatomie (3). Pour la segmentation automatique, il faut disposer d'atlas anatomiques digitaux tels qu'il en existe pour l'humain (4), le macaque (5), la souris (6) et quelques espèces non-conventionnelles comme l'ovin (7). Qu'elle soit manuelle ou automatique, cette étape est dirigée par des *a priori* anatomiques, taille, forme ou position des ROIs (structures anatomiques, tissus etc.), et exige donc de connaître l'anatomie cérébrale de l'espèce étudiée (1).

A partir d'IRMs cérébrales segmentées et labellisées, il est possible d'extraire différentes mesures pour caractériser les ROIs et les comparer entre elles (6,8). La modélisation sous forme de graphes associée à l'intelligence artificielle est une autre option pour analyser et explorer les images segmentées. Les graphes permettent de représenter l'organisation d'objets en fonction des relations qui existent entre eux. Beaucoup utilisée pour représenter les réseaux sociaux, la modélisation sous forme de graphe est tout à fait adaptée à l'encéphale. En effet, il est aisé d'assimiler les différentes régions cérébrales à des nœuds et les connexions à des arêtes. Les nœuds et les arêtes qui composent le graphe sont caractérisés par des attributs. Pour un graphe d'encéphale, les attributs des nœuds peuvent être des volumes ou des intensités de signal, et les attributs des arêtes peuvent être des densités ou des longueurs de fibres blanches par exemple. Mais là encore, il est nécessaire de connaitre l'anatomie des encéphales étudiés. Afin d'apporter un autre point de vue, nous proposons d'explorer les IRMs cérébrales sans *a priori* anatomiques lors des étapes de segmentation et de modélisation sous forme de graphes, avant de déployer des approches d'intelligence artificielle pour analyser et comparer les graphes. Notre volonté de s'affranchir des connaissances en neuroanatomie peut paraitre surprenante. Néanmoins, celle-ci n'est pas si différente de celle prônée par les chercheurs utilisant des approches à base de réseaux de neurones convolutifs[1].

2. Algorithme de segmentation automatique sans *a priori* anatomiques

Pour s'affranchir de tous les prérequis neuroanatomiques, nous avons adapté un algorithme de segmentation automatique uniquement basée sur les informations contenues dans l'image telles que l'intensité de signal. Il existe de nombreux algorithmes de segmentation automatique avec des niveaux de complexité variables.

---

[1] Un réseau de neurones convolutif, ou Convolutional Neural Network (CNN) en anglais, est une architecture spécifique de réseaux de neurones artificiels utilisée en apprentissage automatique sur des images.

Dans notre démarche, nous avons opté pour un algorithme naïf qui ne requiert pas d'autres paramètres qu'un critère d'homogénéité et de taille minimale d'une région : le « split and merge ». Cet algorithme fonctionne en deux phases distinctes : d'abord une division (« split ») du volume en cubes homogènes, puis une fusion (« merge ») pour regrouper certains cubes homogènes voisins. Lors du « split », l'algorithme découpe récursivement l'image en cubes de plus en plus petits jusqu'à obtenir une intensité homogène. Pour le « merge », l'algorithme fusionne les cubes adjacents dès lors que leur intensité de signal satisfait un second critère d'homogénéité. Pour cet algorithme, l'opérateur a donc 3 critères à définir : la taille minimale d'un cube, l'homogénéité pour le « split », et l'homogénéité entre deux cubes adjacents pour le « merge ». Le nombre final de régions obtenues dépend directement de ces trois paramètres de segmentation (Figure 1). Les paramètres qui imposent des régions plus homogènes (image de gauche de la Figure 1) produisent des régions segmentées plus petites, comme constaté au niveau du cervelet. L'ensemble des régions ainsi segmentées peut alors être utilisé pour modéliser l'encéphale sous forme de graphe.

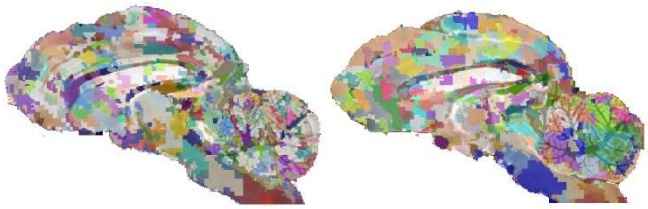


Figure 1 *: Résultats de deux segmentations automatiques "split and merge" d'une IRM cérébrale de mouton (vue sagittale). A gauche : paramètres [homogénéité split : 10, homogénéité merge : 40, taille minimale d'un cube : 1] ; A droite, paramètres [homogénéité split : 10, homogénéité merge : 60, taille minimale d'un cube : 1].*

3. Modélisation de l'organisation cérébrale sous forme de graphe

La modélisation de l'encéphale sous forme de graphe présente une forte flexibilité permettant d'explorer différentes représentations. Le choix de la construction des arêtes du graphe et de ses attributs (descripteurs) permet d'obtenir une diversité de représentations qui révèlent différents niveaux de relations entre les structures. De plus, la théorie des graphes et les techniques d'analyse de graphes peuvent être utilisées pour explorer et comparer des graphes entre eux. Des classes spécifiques d'encéphales, comme différents profils de pathologie (9), peuvent ainsi être identifiées. Par ailleurs, les approches d'intelligence artificielle accroissent encore l'intérêt d'utiliser des graphes pour explorer l'organisation cérébrale.

Pour créer les graphes sans *a priori* anatomique, chaque région obtenue après la segmentation automatique « split and merge » génère un nœud du graphe (Figure 2).

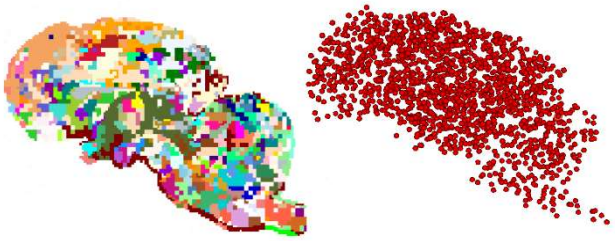


Figure 2 *: Création des nœuds du graphe. Chaque région segmentée à l'issue de la segmentation « split and merge » (à gauche) représente un nœud du graphe (à droite).*

Les arêtes[2] du graphe sont également indépendantes des relations anatomofonctionnelles connues entre les régions. Nous avons exploré plusieurs stratégies de création de graphes. Des graphes dans lesquels tous les nœuds sont reliés entre eux (graphe complet). Des graphes dans lesquels les nœuds sont reliés entre eux en fonction de leur similarité en termes d'intensité de signal, de quantité de surface de contact et/ou de la distance qui les sépare (Figure 3). A ce stade, selon les stratégies choisies pour construire les graphes, ils peuvent avoir un sens biologique. Le critère de seuil d'intensité peut rendre compte de l'organisation tissulaire, et celui de la distance entre deux nœuds peut rendre compte de la plasticité morphologique de l'encéphale. En fonction du seuil choisi pour la similarité d'intensité ou la distance entre les bords

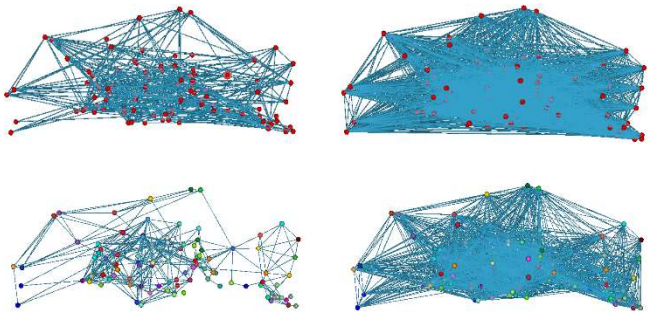


Figure 3 *: Création des arêtes du graphe en fixant un seuil sur l'intensité (en haut) et sur la distance (en bas) à partir d'une image segmentée manuellement (pour une meilleure visibilité). Pour le seuil d'intensité, une arête est créée entre deux nœuds si la différence d'intensité entre les deux régions, qu'ils représentent, est inférieure à 1 (en haut à gauche) ou à 10 (en haut à droite). Pour le seuil de distance, une arête est créée entre deux nœuds si les régions, qu'ils représentent, se touchent (seuil de distance = 0mm, en bas à gauche) ou si la distance entre elles est inférieure à 50mm, en bas à droite).*

[2] *Le terme « arêtes » est utilisé lorsque les graphes sont dits non dirigés (les connexions se font dans les deux sens). Pour un graphe dirigé, le terme « arcs » sera privilégié.*

externes de deux nœuds, les graphes obtenus sont très différents et peuvent mettre en évidence différentes facettes de l'encéphale.

Cette méthode ouvre un vaste panel de possibilités, et ce notamment grâce au choix des attributs à calculer et à stocker dans les nœuds et les arêtes des graphes générés. Pour les nœuds, la position calculée à partir de leur centre de gravité, le volume, la surface et la forme caractérisée par la sphéricité sont quelques possibilités. Les arêtes peuvent servir à stocker des informations de connexité et proximité entre régions. A ces attributs topologiques, peuvent être ajoutés des attributs d'intensité de signal et de descripteurs de texture. Enfin, il est encore possible d'enrichir le graphe en incluant des attributs issus de la théorie des graphes tels que le coefficient de clustering, la détection de communautés, la centralité [3] etc. En créant des graphes contenant le maximum de descripteurs, l'exploration de l'organisation cérébrale à partir d'IRMs anatomiques gagne en flexibilité, et n'est pilotée que par les données et/ou par des mesures et non par l'état des connaissances actuelles des opérateurs.

L'analyse et la comparaison de ces graphes peuvent ensuite être réalisées par des réseaux neuronaux sur graphes (GNN), qui permettent d'exploiter la structure des graphes pour des tâches de classification ou de prédiction. Des auto-encodeurs variationnels sur graphes (VGAE) peuvent également être utilisés pour apprendre des représentations latentes et générer de nouvelles structures de graphes.

4. Classification et comparaison de graphes

Dans notre étude, nous avons choisi d'utiliser les GNN ou les VGAE pour comprendre l'impact du mode d'élevage sur le développement cérébral des ovins. En effet, il a été montré que l'allaitement artificiel altère la maturation cérébrale des agneaux (8) et nous cherchons à évaluer l'impact de stratégies d'enrichissement (maternage, prébiotiques, musique) pour contrecarrer ces effets délétères sur le développement cérébral. Dans leur ensemble, ces connaissances participent à une meilleure compréhension de l'adaptation des agneaux aux contraintes d'élevage et, par conséquent, à améliorer leur bien-être.

Pour produire et exploiter des GNN ou des VGAE, il convient de disposer d'un jeu de graphes dédié à l'apprentissage et d'un jeu de graphes dédié aux tests. Il est alors possible de développer un GNN cherchant à résoudre une tâche de classification d'encéphales selon le mode d'élevage. Un VGAE permettra de générer des représentations vectorielles des graphes, représentatifs des encéphales, afin de pouvoir comparer indirectement leur structure.

4.1. Classification des graphes avec les GNN

Les GNN sont conçus de manière à identifier les relations complexes dispersées dans les nœuds et les arêtes des graphes. Ils peuvent détecter divers types de relations selon la stratégie choisie lors de la création des arêtes dans les graphes (cf. section 3). Pour cela, les GNN exploitent principalement 2 types de couches de réseaux de neurones dans les architectures : la couche de convolution et la couche de pooling. La couche de convolution a pour rôle de propager et d'agréger les informations contenues dans les nœuds du graphe. Cela permet à chaque nœud de mettre à jour sa représentation (appelé embedding) en combinant ses propres caractéristiques avec celles de ses voisins. L'intérêt est d'identifier les relations locales dans le graphe tout en respectant sa structure. La couche de pooling, quant à elle, permet de réduire la taille du graphe et de produire une représentation globale et résumée. Les GNN incluent généralement des couches « denses » chargées de produire un résultat de classification en synthétisant les embeddings produits par les autres couches du réseau de neurones artificiels. Par exemple, l'architecture GNN dédiée à la classification se compose de deux couches de convolution, deux couches de pooling et de couches « denses » constituant un « classifieur » (Figure 4). Il est courant d'alterner entre des couches de convolution et de pooling afin de produire progressivement des descripteurs de plus en plus complexes tout en s'assurant de rendre le modèle robuste aux variations.

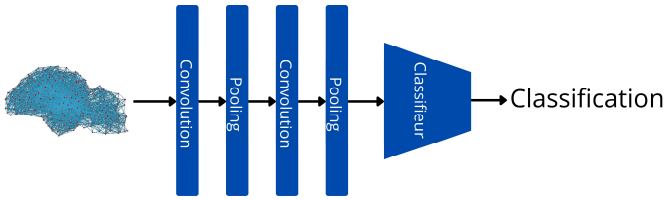


***Figure 4** : Exemple d'architecture de GNN pour résoudre une tâche de classification. Le classifieur se compose de plusieurs couches « denses ».*

En résumé, l'intérêt est de compresser des graphes de grande taille et de créer des représentations compactes. Dans une étude préliminaire, nous avons appliqué les GNN sur des graphes construits à partir d'images segmentées avec l'algorithme « split and merge », ce qui

[3] *Le coefficient de clustering mesure la probabilité que deux nœuds soient connectés sachant qu'ils ont un voisin en commun. La détection de communauté consiste à trouver des groupes de nœuds fortement liés entre eux. La centralité capture la notion d'importance dans un graphe.*

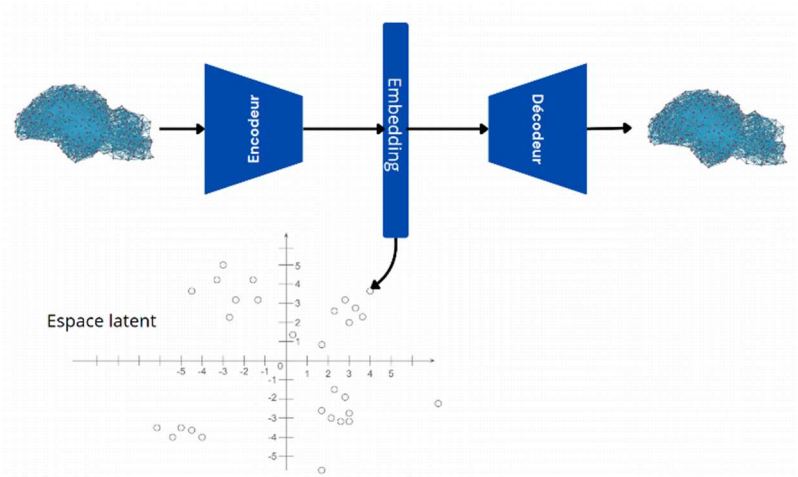


***Figure 5** : Architecture simplifiée d'un VGAE (encodeur + décodeur). Celui-ci apprend à représenter un graphe en un vecteur latent. Chaque graphe peut donc être visualisé comme un point dans un espace (espace latent). Deux points proches dans cet espace représentent deux graphes qui se ressemblent.*

a permis de résoudre une tâche de classification d'âge et de confirmer la pertinence de notre méthode (10).

4.2. Comparaison de graphes et/ou de nœuds avec les VGAE

Les VGAE sont utilisés pour créer des espaces de représentations vectoriels appelés espaces latents interprétables par des experts. L'objectif est de pouvoir transformer les graphes (ou les nœuds des graphes) en des points dans un espace euclidien de plus ou moins grande dimension pour faciliter leur comparaison. Par exemple, deux graphes similaires seront représentés par deux points proches dans l'espace. Les VGAE sont composés d'un encodeur et d'un décodeur (Figure 5). L'encodeur transforme le graphe en un vecteur. Le décodeur apprend ensuite à reconstruire le graphe initial à partir de cette représentation latente. Les VGAE permettent de compresser un graphe en un vecteur de dimension n tout en conservant l'information utile à la tâche, comme du clustering de graphes ou de nœuds (les graphes ou les nœuds sont regroupés par similarité).

5. Plateforme 3DBrainMiner

Pour créer et visualiser des graphes, nous avons développé la plateforme *3DBrainMiner,* disponible en accès libre (Figure 6, https://scm.univ-tours.fr/projetspublics/lifat/3dbrainminer).

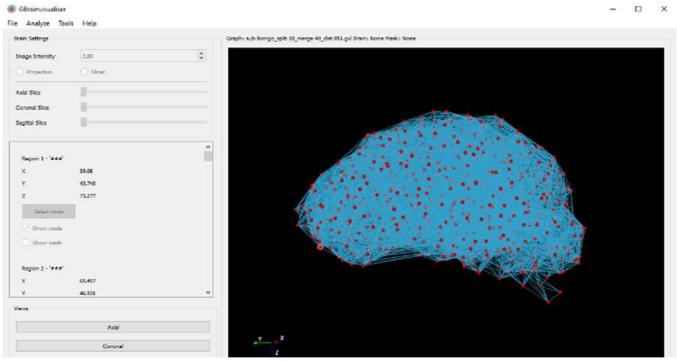


***Figure 6** : Interface de la plateforme 3DBrainMiner avec un graphe d'encéphale de cerf-souris.*

La plateforme 3DBrainMiner peut être utilisée avec tout type de résultats de segmentation manuelle ou automatique (obtenus avec split and merge ou d'autres algorithmes de segmentation). Cette segmentation, couplée à une IRM, permet de créer autant de graphes que nécessaires en utilisant un jeu de paramètres quasiment infinis. Les graphes peuvent être visualisés et sauvegardés dans divers formats (.gxl, .graphml, .pt) constituant des ensembles de données inédits pouvant être diffusés et exploités pour entraîner diverses architectures d'apprentissage profond sur graphes (GNN ou VGAE).

Comme nous l'avons vu précédemment, notre objectif est d'utiliser cette approche sans *a priori* pour identifier, si elles existent, des différences et des similarités entre les graphes d'encéphales d'agneaux élevés dans des conditions différentes, et ce de manière plus exhaustive (8). De plus, une telle approche nous semble pertinente pour comparer l'organisation cérébrale de différentes espèces animales et initier ainsi des approches comparatives sans *a priori*. En effet, si de plus en plus d'IRMs cérébrales d'espèces animales non conventionnelles sont disponibles (DigitalBrainBank https://open.win.ox.ac.uk/DigitalBrainBank/#/, braincatalogue.org), leur exploration reste limitée du fait du manque de connaissances anatomiques.

Comme exemple, nous avons appliqué notre méthode de segmentation « split & merge » sur des IRMs d'encéphale de bongo (*Tragelaphus eurycerus)* et de cerf-souris (*Tragulus javanicus)* et créé des modélisations sous forme de graphes (Figure 7). En appliquant des GNN et des VGAE sur ces graphes, nous espérons mettre en évidence des particularités d'organisation d'encéphale qui pourraient être examinées sous le pan des caractéristiques comportementales propres aux espèces (approche neuroéthologique). De telles relations cerveau-comportement ont déjà été observées à l'échelle de structures (substance grise périaqueducale (11)) ou de connexions cérébrales (cailles d'émotivités différentes (12)).

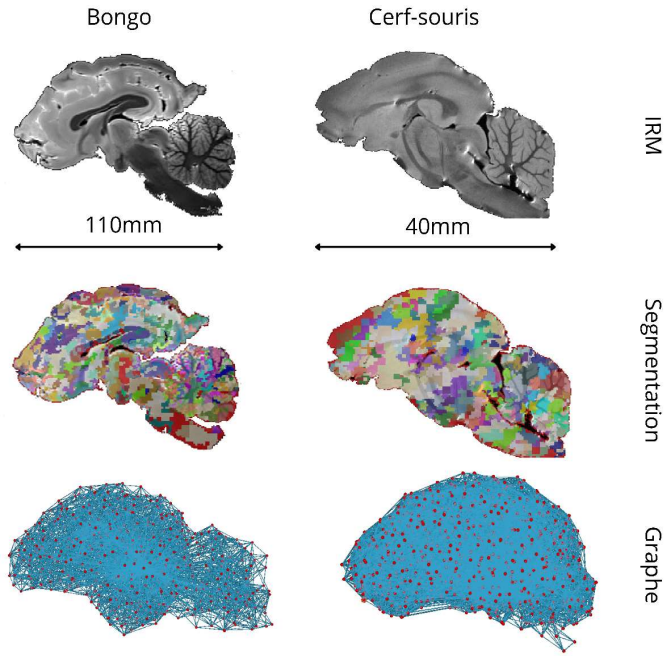


Figure 7 : *Segmentation "split and merge" puis modélisation sous forme de graphes d'un encéphale de bongo et de cerf-souris. Les représentations sont produites avec les mêmes paramètres (homogénéité du split : 10, homogénéité du merge : 40, taille minimale des cubes : 1) et les graphes sont construits en connectant deux nœuds si les régions sont adjacentes.*

6. Conclusions et perspectives.

L'exploration des IRMs cérébrales sans *a priori* anatomiques que nous proposons repose, au départ, uniquement sur les informations contenues dans l'image, permettant ainsi de s'affranchir de toute connaissance ou prérequis anatomiques. Dans cette démarche, nous avons développé une méthode pour segmenter automatiquement des images IRM, pour ensuite les modéliser sous forme de graphes. Nous développons ensuite des architectures de réseaux de neurones sur graphes pour analyser et comparer les graphes. Les avancées récentes faites sur les GNN offrent un large panel de possibilités d'analyses allant de la classification de graphes (encéphales) à la prédiction d'arêtes (relations) en passant par la détection de motifs (sous-graphes) particuliers. Quant à l'utilisation des VGAE, elle ouvre des opportunités d'interprétabilité sans nécessiter de grosses cohortes annotées manuellement du fait qu'elle fonctionne de manière non supervisée. Cette approche pourrait également être utilisée dans le cas où l'anatomie de l'encéphale est connue. Il serait alors possible d'utiliser ces connaissances *a posteriori* pour nommer et/ou comprendre les zones d'intérêts de nos segmentations et graphes.

antoine.bourlier@inrae.fr
elodie.chaillou@inrae.fr
jean-yves.ramel@univ-smb.fr
mohand.slimane@univ-tours.fr

Références

(1) Darrault F. et al. (2025) *J Anat.* 246(5):819-828. doi: 10.1111/joa.14167.

(2) Noorizadeh N. et al. (2024) *Sci Rep.* 14(1), 19114. https://doi.org/10.1038/s41598-024-69995-z

(3) Collier D. C. et al. (2003) *J Appl Clin Med Phys.* 4(1), 17–24. https://doi.org/10.1120/jacmp.v4i1.2538

(4) Shi L. et al. (2017) *Front Hum Neurosci.* 17;11:414. doi: 10.3389/fnhum.2017.00414.

(5) Seidlitz J. et al. (2018), *NeuroImage*, 170, 121–131. https://doi.org/10.1016/j.neuroimage.2017.04.063.

(6) Barrière D. A. et al. (2021) *NeuroImage*, 230, 117776. https://doi.org/10.1016/j.neuroimage.2021.117776

(7) Ella A. et al. *J Comp Neurol.* (2017) 525(3):676-692. doi: 10.1002/cne.24079.

(8) Love S. et al. (2022) *Dev Neurobiol.* 82(2), 214–232. https://doi.org/10.1002/dneu.22869

(9) Fornito, A., Zalesky, A., & Breakspear, M. (2015) *Nat Rev Neurosci.* 16(3), 159–172. https://doi.org/10.1038/nrn3901

(10) Bourlier A. et al. (2025). Comparison between CNN and GNN pipelines for analysing the brain in development. In *Proceedings of the 20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume* 3: VISAPP, ISSN 2184-4321, pages 475-482.

(11) Menant O. et al. (2018) *Brain Struct Funct.* 223(7), 3297–3316. https://doi.org/10.1007/s00429-018-1689-y

(12) Yebga Hot R. et al. (2022) *Brain Struct Funct.* 227(5), 1577–1597. https://doi.org/10.1007/s00429-022-02457-2

Dossier

Les architectes de l'ombre : comprendre et modéliser les processus de construction collective des nids d'insectes sociaux

GUY THERAULAZ

Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), F-31062 Toulouse Cedex 9, France.

Décrypter les mécanismes de coordination des activités chez les insectes bâtisseurs

Les insectes sociaux, tels que les fourmis et les termites sont des architectes hors pair, capables de construire des nids complexes dont la taille peut atteindre dans certains cas plusieurs milliers de fois celle des individus qui les construisent et ceci sans supervision centralisée (1, 2, 3). Ces architectures sont par ailleurs adaptées aux fonctions vitales des colonies, comme la protection contre les prédateurs et certaines d'entre elles permettent également la régulation thermique et l'optimisation des échanges gazeux (4).

Comment des milliers d'individus, parviennent-ils à coordonner leurs actions pour édifier des structures aussi sophistiquées ? Ce phénomène intrigue les scientifiques depuis des décennies mais ce n'est que récemment que l'analyse computationnelle et les concepts issus de la physique statistique ont permis d'en percer les secrets. Grâce aux avancées en vision artificielle et en modélisation mathématique, on est parvenu aujourd'hui à décoder les interactions complexes entre les insectes qui sont impliquées dans la coordination de la construction du nid et à comprendre comment elles façonnent le comportement des groupes.

L'introduction par Pierre-Paul Grassé du concept de *stigmergie* à la fin des années cinquante a permis d'apporter un premier élément de réponse à cette question. Pierre-Paul Grassé (1895-1985) est un célèbre zoologiste français, auteur d'un important Traité de zoologie et également d'un ouvrage en trois volumes sur les termites, *Termitologia*, qui fait encore référence dans le domaine (1). En étudiant la construction du nid chez les termites, Grassé a découvert que ces insectes coordonnaient leurs activités au moyen d'interactions indirectes en utilisant pour cela les traces de leurs activités passées. C'est ce mécanisme qu'il a désigné sous le terme *stigmergie*, un mot dérivé des termes grecs *stigma* (piqûre) et *ergon* (travail) (5).

La stigmergie est essentiellement un processus de type stimulus-réponse. Les traces laissées sur le sol par un insecte lorsqu'il se déplace comme des pistes chimiques ou les ébauches de construction qui résultent de son activité passée vont constituer des sources d'information qui, lorsqu'elles sont perçues par les insectes, vont déclencher des comportements spécifiques chez ces derniers. L'insecte va alors modifier le stimulus qui a déclenché son comportement, ce qui va conduire à la formation d'un nouveau stimulus qui, lui-même, déclenchera chez les insectes de nouveaux comportements spécifiques. Ce processus peut ainsi conduire sous certaines conditions à une parfaite coordination de l'activité des insectes, tout en donnant l'illusion que la colonie dans son ensemble suit un plan prédéfini (6). Par conséquent, pour comprendre comment les insectes d'une colonie construisent leur nid, nous devons identifier et caractériser toutes les boucles stimulus-réponse car elles vont jouer un rôle clé dans la coordination des comportements des individus ainsi que dans la croissance et la forme des nids.

La construction du nid chez les fourmis

Nous avons pu décrypter les interactions stigmergiques qui sont utilisées par les fourmis *Lasius niger* pour construire leur nid (7). Ces derniers sont construits à partir

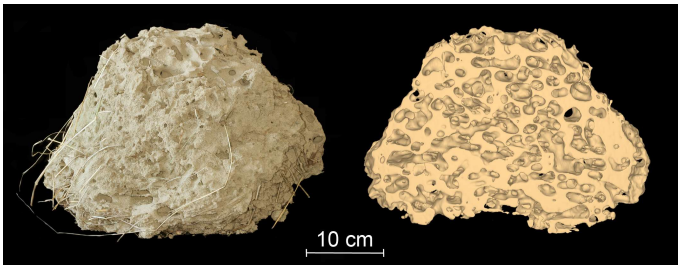


Figure 1. *Nid de la fourmi noire des jardins Lasius niger, (taille de la colonie : environ cinq à quinze mille individus). A gauche, partie épigée. A droite, une tomographie aux rayons X montrant la structure interne du nid*

de l'assemblage de boulettes de terre et leur structure interne ressemble à celle d'une éponge avec un grand nombre de chambres en forme de bulles, étroitement imbriquées les unes dans les autres (voir Figure 1). Pour étudier l'activité de construction chez cette espèce, on utilise un dispositif constitué d'une boite de pétri contenant une fine couche de sable humidifié dans laquelle on place ensuite un groupe de 500 fourmis (Figure 3A). Le dispositif étant éclairé en permanence, ces conditions stimulent très fortement l'activité de construction des fourmis qui essayent de construire un abri pour se protéger. Mais comme la quantité de sable disponible est limitée, le processus s'arrête juste après les toutes premières étapes de construction du nid qui conduisent à la formation de piliers et de murs régulièrement espacés.

Les comportements des fourmis de prise et de dépôts de boulettes de terre impliqués dans la construction des nids sont modulés par les traces de leurs actions passées dans l'environnement de la manière suivante. Après avoir prélevé des boulettes de terre sur le sol les fourmis les imprègnent d'une phéromone (*i.e.* une substance chimique qui agit comme un message entre les fourmis). Celle-ci, lorsqu'elle est détectée par les fourmis, stimule un comportement de dépôt de boulettes dans des zones riches en phéromone où d'autres boulettes ont déjà été récemment déposées ce qui crée un feed-back positif : plus le nombre de boulettes déjà déposées en un endroit est important, plus les fourmis ont tendance à venir y déposer d'autres boulettes. Mais dans le même temps, cette même phéromone va inhiber le comportement de prise de boulettes de terre au même endroit. Plus la zone détectée par une fourmi est riche en phéromone, plus le nombre de boulettes récemment déposées dans la zone est important, et plus la probabilité de prendre une boulette dans cette zone diminue. Enfin, lorsque les piliers qui résultent de l'amoncellement de boulettes atteignent une hauteur qui correspond approximativement à la longueur moyenne du corps des fourmis, celles-ci cessent de construire en hauteur et déposent des boulettes sur la paroi latérale des piliers. Les interactions des fourmis avec les structures qu'elles construisent vont ainsi leur permettre de coordonner leurs actions sans avoir besoin d'échanger directement de l'information. Chaque fourmi pouvant travailler dans son coin sans se soucier de ce que font les autres au même moment dans le nid.

Nous avons ensuite modélisé ces processus de construction. Dans le modèle nous avons représenté explicitement les fourmis et les boulettes d'argile qu'elles utilisent pour construire leur nid. Les fourmis sont représentées par des agents qui se déplacent dans un espace en trois dimensions et discret, c'est à dire qu'il est constitué de cellules cubiques. Le mouvement des agents est contraint par les structures qu'ils construisent, de sorte qu'au cours de leurs déplacements, ils restent en contact avec la surface de l'architecture, comme le font les fourmis. Le déplacement des fourmis et leurs comportements de prise et de dépôt de boulettes sont contrôlées par les trois fonctions qui ont été calibrées à

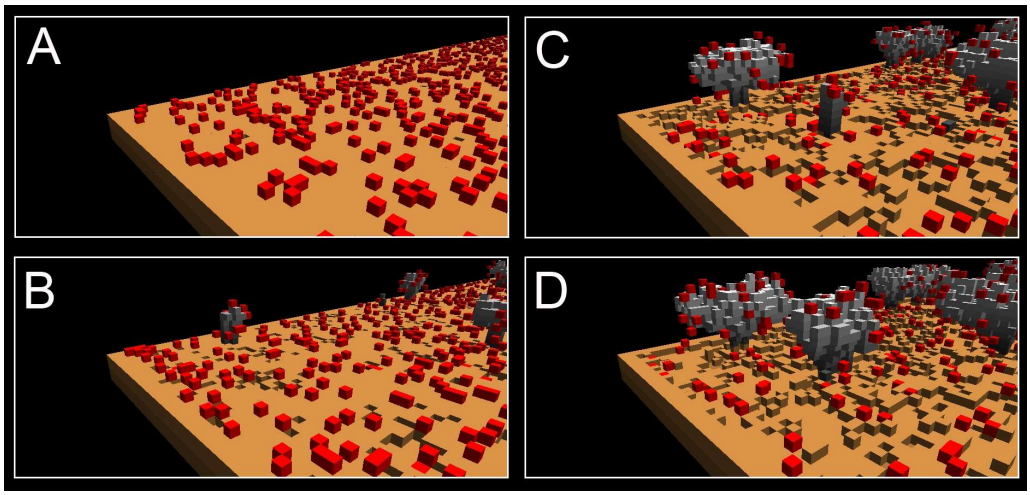


Figure 2*. Simulations d'un modèle de construction de nid de fourmis. Les « agents-fourmis » représentés par les cubes rouges prélèvent et déposent des boulettes de terre en suivant les règles de comportement caractérisées par l'expérience (A). Par hasard lorsque la densité de boulettes en un endroit est suffisamment élevée, ces zones deviennent des germes à partir desquels les fourmis construisent les piliers grâce au feed-back positif induit par la phéromone de construction qui stimule le dépôt (B). Par la suite, lorsque la taille des piliers est suffisamment grande, les fourmis réalisent des dépôts en hauteur ce qui conduit à la formation de chapiteaux de forme globulaire (C). Enfin lorsque deux chapiteaux sont suffisamment proches ils fusionnent pour former un passage voûté (D).*

partir de l'analyse des données expérimentales. Les simulations numériques du modèle montrent que la combinaison des interactions entre les fourmis et les structures qu'elles construisent conduisent à la formation de piliers régulièrement espacés, puis de chapiteaux au-dessus de ces piliers dont les formes reproduisent fidèlement celles réalisées par les fourmis dans les expériences (Figure 2A-D) (7).

La simulation du modèle dans lequel ont été transposés les processus de construction identifiés expérimentalement montre que lorsque la quantité de boulettes en un endroit est suffisamment élevée, celles-ci deviennent des germes à partir desquels les fourmis construisent des piliers (Figure 2B). Par la suite, lorsque la taille des piliers est suffisamment grande, les fourmis commencent à réaliser des dépôts en hauteur qui produisent des structures globulaires et massives en forme de chapiteaux surplombant les piliers (Figure 2C). Puis, lorsque deux chapiteaux sont suffisamment proches l'un de l'autre, ils peuvent fusionner, créer un passage voûté et puis progressivement une chambre (Figure 2D).

La comparaison des structures produites dans le modèle à celles obtenues dans les expériences permet de vérifier que les interactions et les comportements identifiés dans l'expérience et ensuite implémentés dans le modèle reproduisent fidèlement la dynamique de construction et la forme des nids construits par les fourmis (Figure 3).

Le modèle a permis de mettre en évidence le rôle clé de la phéromone ajoutée par les fourmis au matériel de construction dans l'adaptation de la forme du nid aux changements des conditions environnementales. En effet, selon les conditions climatiques, la phéromone (dont la demi-vie est de l'ordre de 20 minutes dans un environnement humide) se dégrade plus ou moins rapidement, ce qui conduit les fourmis à construire un nombre moins élevé de piliers dans un environnement sec. Les chambres construites sont alors beaucoup plus grandes, et les fourmis peuvent s'y agréger en grand nombre ce qui permet de conserver le peu d'humidité disponible (Figures 3B et 3F-H). Inversement, dans un environnement humide, la phéromone persiste plus longtemps, ce qui conduit à la construction d'un nombre plus élevé de piliers et à des chambres plus petites (Figures 3A et 3C-E).

Ainsi la modulation de la durée de vie de la phéromone par les conditions environnementales permet d'adapter automatiquement la forme du nid, sans que les fourmis aient besoin de changer leurs règles de construction. Les processus de construction permettent donc à une colonie de produire des structures dont la forme va changer selon les conditions de température et d'humidité. Ces processus permettent ainsi une économie de codage des mécanismes comportementaux qui à l'échelle individuelle permettent aux insectes de coordonner leurs activités et de construire des structures qui, chez certaines espèces, peuvent atteindre un très haut degré de complexité.

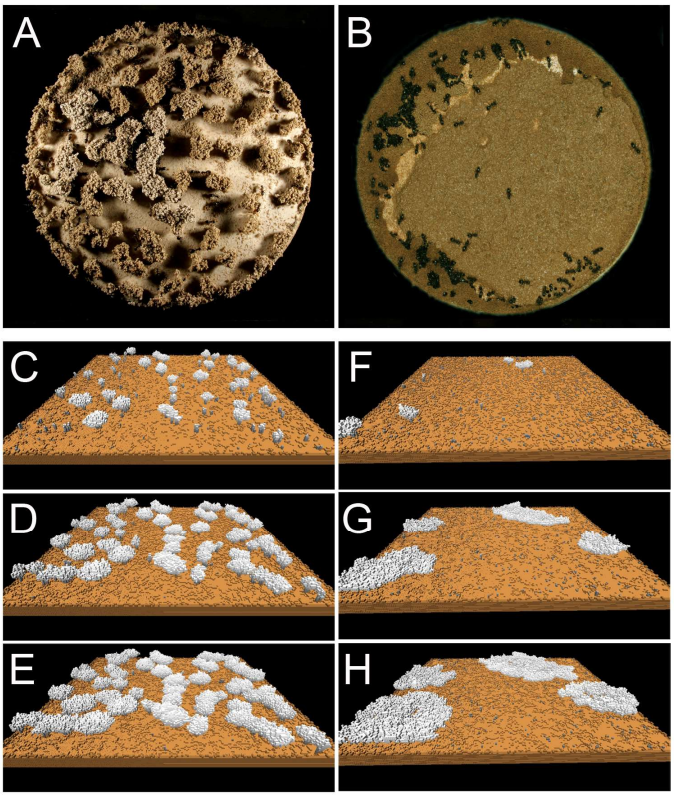


***Figure 3**. Impact des conditions environnementales sur la durée de vie de la phéromone ajoutée par les fourmis au matériel de construction et sur la forme du nid. Les photographies A et B montrent les structures construites par des groupes de 500 fourmis en 36 heures dans deux milieux différents. Le diamètre des boites de Petri utilisées dans ces expériences est de 10 cm. Dans un milieu humide (A), la durée de vie de la phéromone est élevée et les fourmis construisent des piliers et des murs régulièrement espacés. Dans un milieu sec (B), la phéromone a une durée de vie beaucoup plus courte et les fourmis construisent de très larges plateaux. Les simulations informatiques du modèle de construction de nid reproduisent l'effet de la durée de vie de la phéromone sur la forme des structures construites par les fourmis : trois étapes successives (12h, 24h et 36h) d'une simulation lorsque la durée de vie de la phéromone est élevée (C,D,E) et lorsque la durée de vie de la phéromone est courte (F,G,H).*

L'origine de la complexité des nids de termites

Chez les insectes sociaux, les termites du genre *Apicotermes* ont développé l'art de la construction des nids à un niveau inégalé dans le règne animal. La complexité, la taille mais également la solidité des nids sont en cela bien supérieures aux structures construites par les fourmis. Pour édifier le nid, les ouvriers termites utilisent essentiellement du mortier stercoral constitué d'excréments et exsudations anales. Ces sociétés de termites vivent exclusivement en Afrique dans la forêt et la savane et construisent leurs nids sous terre entre vingt et soixante centimètres de profondeur. De forme ovoïde

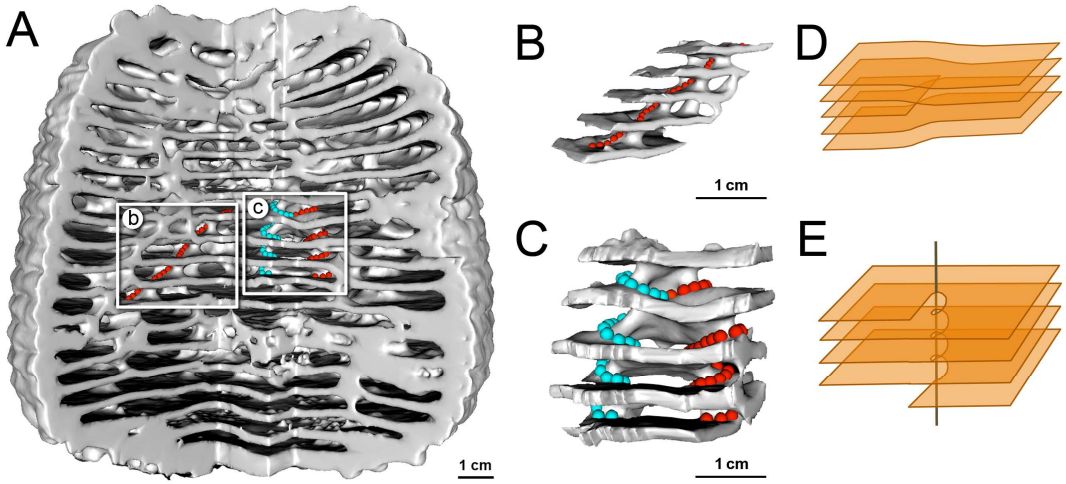


Figure 4 : *Reconstruction tridimensionnelle d'un nid d'Apicotermes lamani à partir d'une tomographie aux rayons X (A) originaire de Pointe Noire (République du Congo) au sein duquel sont représentées les positions d'une rampe linéaire (B) et de deux rampes hélicoïdales (C) reliant plusieurs étages du nid. Ces rampes, matérialisées par les sphères rouges et bleues résultent de la création de défauts topologiques ou dislocations produites au cours du remodelage du nid par les termites et qui conduisent au désalignement des planchers délimitant les différentes chambres. Les rampes linéaires apparaissent au niveau de dislocations coin qui résultent de l'introduction d'un demi plancher (sur la gauche de la figure) dans la succession régulière des planchers du nid (D). Les rampes hélicoïdales apparaissent au niveau de « dislocations vis » qui résultent d'un désalignement des planchers autour d'un plan de glissement perpendiculaire aux planchers (E).*

et pouvant atteindre une hauteur de quelques dizaines de centimètres, la surface des nids est couverte de petits orifices en forme de gargouilles disposés en lignes très régulièrement espacées et qui parcourent toute la circonférence externe de la paroi. L'intérieur est constitué d'une succession de vastes chambres superposées les unes aux autres et qui sont délimitées par des planchers parallèles et régulièrement espacés ; par ailleurs des rampes linéaires et hélicoïdales assurent la communication entre les différents étages du nid (Figure 4).

Pour comprendre les processus intervenant dans la construction de ces structures particulièrement complexes nous avons tout d'abord caractérisé et quantifié l'architecture des nids de ces termites au moyen de techniques d'imagerie 3D (tomographie aux rayons X). Comme chez les fourmis, le comportement de construction des termites s'organise en réponse à des signaux locaux tels que la forme des structures déjà construites et la présence de phéromone dans ces structures et dont les composés chimiques sont encore peu connus. L'architecture du nid dicte à chaque instant les espaces accessibles aux termites ; la densité de termites dans les différentes zones du nid détermine à son tour la concentration de phéromone de construction qui y est présente, et cette dernière sert enfin de gabarit pour le remodelage de l'architecture du nid par les termites. Les ouvriers termites détruisent certaines parties du nid et elles reconstruisent de nouvelles structures de sorte que le remodelage du nid est permanent. Nous avons ensuite construit un modèle mathématique simple décrivant les interactions physiques et biologiques entre les termites, la phéromone de construction sécrétée par ces derniers qui guide leur comportement bâtisseur et enfin le matériel utilisé pour construire le nid (8).

Les simulations numériques du modèle reproduisent fidèlement la structure des nids d'*Apicotermes* avec des planchers régulièrement espacés qui sont reliés en certains endroits par des rampes linéaires et hélicoïdales (Figure 5). L'analyse du modèle montre également que ces rampes résultent de la création de défauts topologiques ou dislocations. Au cours du remodelage du nid il arrive en effet que des planchers contigus ne soient plus alignés. Lorsque cela se produit, le désalignement des planchers crée soit des « dislocations coin » au niveau desquelles apparaissent des rampes linéaires reliant les étages adjacents ou des « dislocations vis » (un désalignement des planchers autour d'un plan de glissement) qui fournissent l'axe de rotation autour duquel apparaissent des rampes hélicoïdales (Figure 4D-E). De tels défauts topologiques ne se retrouvent pas dans les nids de fourmis probablement en raison des propriétés des matériaux utilisés dans la construction de leurs nids qui rendent les structures beaucoup plus friables.

Ces résultats montrent que la complexité des nids construits par les termites *Apicotermes* résulte d'un décalage spatial et temporel dans la croissance des différents éléments qui les composent. Ces résultats illustrent également une des principales propriétés des processus d'auto-organisation chez les insectes sociaux qui est de permettre l'émergence de propriétés nouvelles à l'échelle collective à partir de comportements individuels et d'interactions physiques et biologiques extrêmement simples.

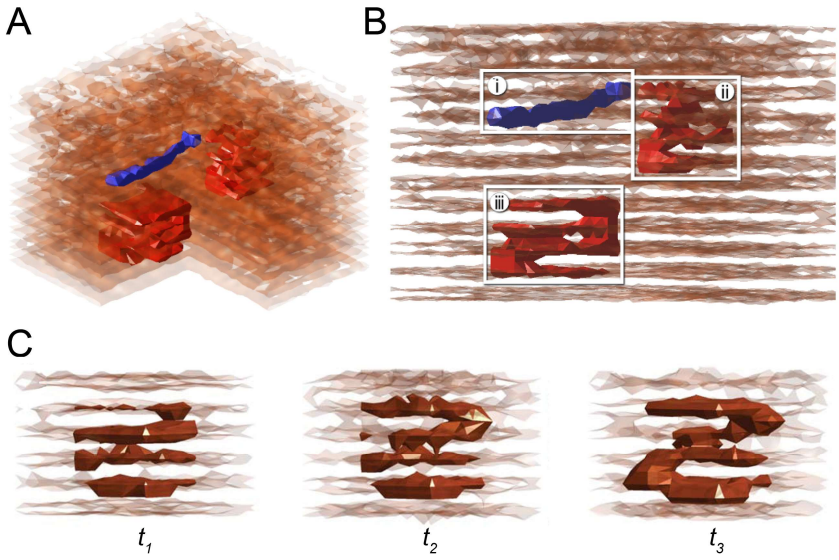


Figure 5 : *Simulations numériques du modèle de construction de nid (A, B). Les structures simulées reproduisent la structure des nids d'Apicotermes lamani avec une rampe linéaire (i, bleu) et deux rampes hélicoïdales (ii-iii, rouge) reliant plusieurs étages du nid. Trois étapes successives (t_1, t_2 et t_3) de la formation d'une rampe hélicoïdale dans une simulation du modèle de construction de nids (C).*

Conclusions

L'étude des processus de construction collective chez les insectes sociaux a permis d'identifier des mécanismes robustes et efficaces d'auto-organisation (9). La formation des nids repose sur des interactions simples entre les individus et leur environnement conduisant à des boucles de rétroaction positive et négative. L'architecture des nids n'est pas figée mais elle évolue au cours du temps en fonction des besoins de la colonie et des contraintes environnementales. Enfin l'adaptabilité des nids repose sur un équilibre subtil entre les actions individuelles de construction et de remodelage et les contraintes physiques qu'imposent l'architecture sur les déplacements des insectes. Ces recherches offrent non seulement des perspectives fascinantes pour la biologie de l'évolution et la cognition collective, mais elles inspirent également des applications dans des domaines tels que la robotique en essaim et l'architecture biomimétique (10, 11).

guy.theraulaz@univ-tlse3.fr

Références

(1) Grassé P.P. (1984) *Termitologia, Vol. II: Fondation des sociétés – Construction.* Masson, Paris.

(2) Hansell M. (2005) *Animal Architecture*. Oxford, Oxford University Press.

(3) Perna A. & Theraulaz G. (2017) *J Exp Biol*, 220: 83-91.

(4) Singh K. et al. (2019) *Sci Adv*, 5: eaat8520.

(5) Grassé P.P. (1959) *Insectes Sociaux* 6 : 41–83.

(6) Theraulaz G. & Bonabeau E. (1999) *Artif Life*. 5: 97-116.

(7) Khuong A. et al. (2016) *PNAS* 113: 1303–1308.

(8) Heyde A. et al. (2021) *PNAS* 118 (5): e2006985118.

(9) Camazine S. et al. (2001) *Self-Organization in Biological Systems.* Princeton, Princeton University Press.

(10) Yang G. et al. (2022) *Buildings*, 12, 2225.

(11) Pinter-Wollman N. et al. (2018) *Philos Trans R Soc Lond B Biol Sci.* 373: 20170232.

Tribune Libre – Questions d'actualité

Comment mener des recherches expérimentales en milieu scolaire ?

SEBASTIEN GOUDEAU[1], MATTHEW J. EASTERBROOK[2], MARIE-PIERRE FAYANT[3]

[1]Centre de Recherches sur la Cognition et l'Apprentissage, Université de Poitiers - CNRS, France
[2]University of Sussex, Falmer, United Kingdom
[3]Laboratoire de Psychologie Sociale, Université Paris Cité, Boulogne-Billancourt, France

Les recherches en milieu scolaire peuvent éclairer les enseignants et les décideurs sur les pratiques qui sont (in)efficaces pour améliorer les résultats des élèves (par exemple, augmentation des acquis, amélioration du bien-être, réduction des inégalités ou du harcèlement). Elles apportent également des éclairages précieux sur des aspects souvent inaccessibles en dehors du cadre scolaire, comme le climat de classe, les relations avec les enseignants ou encore les interactions entre élèves.

Ces dernières années, les avancées en neurosciences ont suscité un intérêt croissant pour leur contribution potentielle à la compréhension des mécanismes de l'apprentissage. En identifiant les processus cognitifs impliqués dans l'attention, la mémoire, la motivation ou encore la régulation des émotions, les neurosciences permettent d'enrichir la réflexion sur les conditions optimales d'enseignement. Cela dit, le propos de cette tribune n'est pas de discuter de l'intérêt des recherches en neurosciences pour l'école – sur ce point, nous renvoyons les lecteurs intéressés au livre Les neurosciences à l'école : leur véritable apport[1] – mais de s'interroger sur les précautions méthodologiques nécessaires à la mise en œuvre de recherches expérimentales dans le cadre scolaire. Avec l'intérêt croissant des institutions publiques et des enseignants pour les essais randomisés contrôlés à l'école, la question de leur faisabilité et de leur évaluation à grande échelle devient un enjeu majeur. Les recherches en classe sont ainsi essentielles pour évaluer l'efficacité des pratiques dans leur contexte réel et garantir leur applicabilité à large échelle.

Dans un article publié dans Nature Human Behaviour (1), nous avons utilisé notre expérience en tant que chercheurs pour mettre en lumière les enjeux à chaque étape du processus de recherche, tout en fournissant des recommandations pratiques sur la manière de mener des expérimentations en éducation qui soient fiables,

[1] *Les neurosciences à l'école : leur véritable apport, par Edouard Gentaz, Ed O Jacob, 2022.*

permettent de statuer sur la nature causale des liens, génèrent du savoir applicable aux élèves et situations scolaires qui dépassent le contexte de l'étude, et soient éthiques. Ces recommandations visent également à favoriser des collaborations positives avec les écoles, les enseignants et autorités académiques. Bien que nous nous concentrions principalement sur les recherches expérimentales et quasi-expérimentales, certains arguments sont également pertinents pour d'autres types de méthodologies (par exemple, les enquêtes par questionnaire ou par observations).

Validité et reproductibilité des recherches en éducation

Un des enjeux fondamentaux concerne la validité interne, qui assure que les effets observés sont bien dus aux variables manipulées et non à des facteurs externes (2).

En contexte scolaire, la présence des enseignants, les interactions entre élèves, ou encore les attentes implicites des participants peuvent biaiser les résultats. Pour limiter ces biais, les chercheurs doivent veiller à l'élaboration d'un protocole rigoureux, incluant des mécanismes de contrôle. De plus, l'implication des enseignants et leur compréhension des principes expérimentaux peuvent jouer un rôle déterminant dans la fiabilité des résultats.

La validité externe, qui permet de généraliser les conclusions à d'autres contextes éducatifs, peut être remise en question si l'expérimentation s'éloigne trop des conditions réelles d'apprentissage (3). Ainsi, il est recommandé d'effectuer des études dans des environnements scolaires variés et d'inclure des échantillons diversifiés pour accroître la généralisabilité des résultats. La reproductibilité est également un défi majeur. Les contraintes logistiques des écoles (emplois du temps, diversité des environnements pédagogiques) rendent souvent difficile la réplication exacte des études. L'adoption de méthodologies rigoureuses et la transparence dans le partage des données et protocoles sont essentielles pour assurer la robustesse des conclusions (4). Documenter chaque étape du processus expérimental et publier les jeux de données anonymisés peut également favoriser la reproductibilité.

Défis et solutions à chaque étape du processus

Avant l'étude : anticiper les obstacles

Recrutement et engagement des enseignants : L'implication des enseignants est essentielle mais leur connaissance des hypothèses de l'étude peut dans certains cas biaiser les résultats si elle influence leur comportement en classe. Une formation préalable sur le protocole expérimental, sans dévoiler l'hypothèse centrale de l'étude, permet de limiter ces biais (5). Cette formation peut inclure des séances de sensibilisation aux méthodologies de recherche, ainsi qu'une discussion sur les enjeux liés à la validité des résultats. Mettre les informations clés dans un document (par exemple, comment présenter l'étude, les termes à éviter et comment répondre aux questions courantes comme « Ce test sera-t-il noté ? ») peut augmenter la validité.

Consentement parental et éthique : La réalisation de recherches en milieu scolaire implique souvent l'obtention d'un consentement parental explicite (opt-in), ce qui suppose que les écoles doivent envoyer des formulaires aux parents, collecter et enregistrer leurs réponses, puis suivre quels élèves peuvent ou ne peuvent pas participer. Cette exigence administrative peut être particulièrement difficile à mettre en œuvre lorsque le protocole expérimental requiert la participation de classes entières, comme dans certaines recherches sociométriques. De plus, cette procédure peut entraîner une sous-représentation de certains groupes d'élèves, notamment ceux issus de familles à faibles revenus. Lorsque l'étude se déroule dans des contextes de classes ordinaires et que les élèves réalisent des tâches proches de leurs apprentissages du quotidien, le consentement opt-out est une alternative qui permet aux élèves de participer automatiquement à l'étude sauf si leurs parents expriment explicitement leur refus. Ce modèle, lorsqu'il est accepté par les comités d'éthique, facilite le recrutement et réduit les biais de sélection tout en garantissant les principes éthiques.

Protection des données : L'appariement des données issues de l'étude avec d'autres variables (socio-démographiques, scolaires, de classe et/ou des enseignants) doit respecter les réglementations en matière de protection des données, nécessitant souvent une anonymisation ou une pseudonymisation. L'anonymisation supprime ou modifie les informations personnelles de manière irréversible afin que les individus ne puissent pas être identifiés, même en croisant plusieurs sources de données. La pseudonymisation remplace les identifiants personnels par des pseudonymes, réduisant le risque d'identification, mais permettant une ré-identification si la clé de pseudonymisation est accessible. Pour assurer un appariement efficace et conforme aux exigences éthiques et réglementaires, il est essentiel de fournir aux établissements scolaires un protocole détaillé expliquant les données requises, le processus d'appariement et un calendrier réaliste. Une rencontre avec l'administrateur des données peut faciliter la mise en œuvre. Enfin, lorsque cela est applicable, fournir aux élèves des codes d'anonymat en amont de la collecte garantit une

organisation fluide et conforme aux exigences de confidentialité.

Assignation du traitement expérimental : Un problème éthique peut survenir lorsqu'un traitement visant à améliorer des résultats positifs (ex. réussite scolaire) ou à réduire des effets négatifs (ex. harcèlement) est testé. Il est essentiel d'éviter de pénaliser certains élèves en leur refusant un traitement potentiellement bénéfique. Cette préoccupation, légitime pour les enseignants, peut les inciter à influencer l'assignation (ex. en veillant à ce que les élèves en difficulté bénéficient du traitement) ou à refuser de participer si tous les élèves n'y ont pas accès, ce qui compromet la validité interne de l'étude. Il est donc important d'expliquer aux enseignants l'intérêt des études randomisées pour évaluer l'efficacité d'un traitement et montrer que l'assignation aléatoire constitue une méthode juste et impartiale pour répartir des ressources limitées. Trois stratégies peuvent être envisagées : 1) mettre en place un système de liste d'attente permettant aux participants de la condition contrôle de recevoir le traitement plus tard ; 2) utiliser une assignation basée sur une variable quantitative, comme la moyenne scolaire, pour garantir que ceux qui en ont le plus besoin bénéficient du traitement ; 3) inclure tous les participants dans le groupe expérimental via des modèles intra-sujets où les élèves sont exposés alternativement aux différentes conditions ou via des séries temporelles interrompues où le traitement est introduit au milieu de la période d'observation. Enfin, il est fortement recommandé que les chercheurs contrôlent l'assignation et ne révèlent jamais le critère de répartition des élèves entre les conditions, sauf en cas de randomisation.

Pendant l'étude : assurer la rigueur méthodologique

Stabilité de l'assignation aux conditions : Un problème fréquent dans les études expérimentales concerne la stabilité de l'assignation des participants aux conditions (6). Celle-ci peut être compromise lorsque les élèves ayant déjà complété l'étude en discutent avec ceux qui ne l'ont pas encore fait, ou lorsqu'ils remarquent des différences de matériel dans une même classe. Pour limiter ces risques, il est essentiel de collaborer avec l'école afin de restreindre les interactions entre élèves de différentes classes, par exemple en ajustant les horaires de récréation ou de repas. Le moment du débriefing doit être soigneusement planifié. Si l'assignation est intra-classe, il faut s'assurer que les élèves ne puissent pas voir le matériel d'une autre condition, soit en utilisant des écrans de séparation, soit en rendant les supports visuellement identiques. Une explication crédible peut être donnée pour justifier d'éventuelles différences, par exemple : « Pour éviter la triche, les exercices ne sont pas identiques. »

Mise en œuvre des conditions expérimentales : L'application rigoureuse du traitement et de la condition contrôle est essentielle pour examiner les relations causales. Toutefois, les salles de classe présentent des contraintes difficiles à maîtriser, comme la disposition des lieux, le stress des élèves lié à d'autres évaluations, ou encore les commentaires des enseignants sur les tâches expérimentales. Ces facteurs externes, souvent non contrôlés, augmentent le bruit statistique et réduisent la validité interne. Par exemple, un enseignant peut rassurer les élèves dans une condition évaluative ou, au contraire, leur signaler qu'un exercice comptera pour leur note, influençant ainsi les résultats. Il est donc crucial d'anticiper et de documenter ces variables. Pour limiter ces biais, il faut signaler dans les publications tous les facteurs externes identifiés, éviter toute variation systématique entre conditions (ex. classes différentes pour les groupes contrôle et traitement) et standardiser autant que possible les éléments contextuels (ex. utiliser les mêmes salles pour toutes les conditions). La présence des enseignants doit être limitée lors des tests, car leur influence est difficile à contrôler ; lorsqu'ils sont impliqués dans l'intervention, un bref questionnaire post-session peut aider à mesurer leur impact. L'instruction aux élèves doit être claire et uniforme, en précisant chaque détail (ex. type de stylo à utiliser, consignes de correction) et en lisant les consignes à haute voix aux jeunes enfants pour éviter toute disparité de compréhension. La recherche en classe peut aussi générer du stress lié à l'évaluation ou à la comparaison sociale, ce qui peut interférer avec les comportements des élèves ; il est donc recommandé d'insister sur le fait que la tâche n'évalue pas leurs compétences individuelles et de limiter l'accès aux performances des autres. Enfin, il est indispensable de documenter toutes les sources de variation, même celles qui semblent insignifiantes, en réalisant des tests pilotes et en enregistrant le protocole par vidéo ou photos pour garantir la reproductibilité des résultats.

Après l'étude : garantir un impact positif et transparent

Débriefing : Le débriefing est essentiel, en particulier dans les recherches menées en milieu scolaire, afin d'éviter que les participants et enseignants ne se sentent trompés. Révéler le protocole expérimental peut susciter de la méfiance ou du ressentiment chez les enseignants, et certaines manipulations peuvent avoir des effets négatifs durables (7), notamment en contexte évaluatif (ex. menace pour l'image de soi après un retour négatif). Il est donc important d'annoncer dès le départ que tous les détails ne peuvent pas être communiqués immédiatement, tout en garantissant un débriefing approfondi après l'étude. Une discussion détaillée, expliquant les objectifs et corrigeant d'éventuels malentendus, permet d'atténuer

les effets négatifs et de maintenir des relations de travail respectueuses. Lorsque cela est pertinent, une induction positive (ex. mise en avant de l'inclusion) en fin de session peut également restaurer l'état psychologique des participants (8). Le moment du débriefing doit être soigneusement choisi : si aucune communication entre élèves n'est possible, il peut être fait immédiatement après l'étude ; sinon, il est recommandé d'organiser le débriefing à un moment précis en collaboration avec l'administration scolaire, afin d'éviter toute diffusion du traitement entre les groupes.

Protection des élèves : Les établissements scolaires ont la responsabilité de garantir la sécurité de leurs élèves. Certaines recherches en classe peuvent impliquer la production de documents nécessitant une vérification pour identifier d'éventuels signaux d'alerte, comme des intentions d'automutilation dans des rédactions. Il est donc crucial de prévoir des ressources suffisantes pour assurer ces procédures et d'établir un protocole clair définissant les mesures à prendre en cas de problème, y compris les personnes à contacter et la manière dont les informations seront partagées, en tenant compte des engagements de confidentialité faits aux élèves. Ces aspects doivent être anticipés lors du processus d'approbation éthique, mais nécessitent souvent une action spécifique après la fin de l'étude.

Conclusion

Les recherches expérimentales en éducation offrent des perspectives précieuses pour améliorer les pratiques pédagogiques, à condition d'être conduites avec rigueur. En intégrant dès la conception de l'étude des précautions méthodologiques adaptées, il est possible de concilier exigences scientifiques et respect des spécificités du milieu scolaire. Ces recherches, lorsqu'elles sont bien menées, constituent un levier essentiel pour identifier les pratiques les plus efficaces, comprendre leurs mécanismes d'action, et en évaluer l'impact dans des contextes réels. Elles permettent ainsi d'éclairer les choix pédagogiques et les décisions politiques en s'appuyant sur des données solides. C'est précisément parce que leur potentiel est si important qu'il est crucial d'en soigner la mise en œuvre.

sebastien.goudeau@univ-poitiers.fr
m.j.easterbrook@sussex.ac.uk
marie-pierre.fayant@u-paris.fr

Références

(1) Goudeau, S., Easterbrook. M., & Fayant, M. P. (2025). How to do research in classroom settings . Nature Human Behaviour. doi: 10.1038/s41562-024-02027-y

(2) Chester, D. S., & Lasko, E. N. (2021). Construct validation of experimental manipulations in social psychology: Current practices and recommendations for the future. Perspectives on Psychological Science, 16(2), 377-395.

(3) Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. Perspectives on Psychological Science, 7(2), 109-117.

(4) Nosek BA, Hardwicke TE, Moshontz H, Allard A, Corker KS, Dreber A, Fidler F, Hilgard J, Kline Struhl M, Nuijten MB, Rohrer JM, Romero F, Scheel AM, Scherer LD, Schönbrodt FD, Vazire S. (2022). Replicability, robustness, and reproducibility in psychological science. Annual review of psychology, 73(1), 719-748.

(5) Uluğ, Ö. M., Zoodsma, M., Sandbakken, E. M., Figueiredo, A., Rocha, C., Sagherian-Dickey, T. Acar Y G, Moss S M, Saab R & Woo, Y. T. (2023). How can social psychologists become more participatory in their research? A reflection on working 'with'communities and participants rather than 'on' them. Social Psychological Review, 25(1), 9-14.

(6) West, S. G., Cham, H., & Liu, Y. (2014). Causal inference and generalization in field settings: Experimental and quasi-experimental designs. In H. T. Reis & C. M. Judd (Eds.), Handbook of research methods in social and personality psychology (2nd ed., pp. 49–80). Cambridge University Press.

(7) Miketta, S., & Friese, M. (2019). Debriefed but still troubled? About the (in) effectiveness of postexperimental debriefings after ego threat. Journal of Personality and Social Psychology, 117(2), 282.

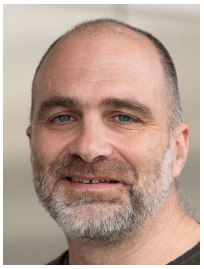
(8) Zwolinski, J. (2014). Does inclusion after ostracism influence the persistence of affective distress? Group Dynamics: Theory, Research, and Practice, 18(4), 282.

Nouveautés en Neurosciences

Le thalamus cognitif

MATHIEU WOLFF

INCIA, UMR5287, Bordeaux Neurocampus



Les fonctions thalamiques font actuellement l'objet de nouvelles découvertes. Trop longtemps considéré comme un simple relai, le thalamus apparaît désormais comme un acteur majeur dont les fonctions intégratives opèrent à l'échelle du cerveau entier. Retour sur un véritable changement de paradigme.

Le thalamus, une structure plus hétérogène que complexe

Le thalamus est souvent considéré comme une structure profonde dont la complexité serait intimidante en raison notamment d'une riche connectivité avec d'autres régions cérébrales et de l'existence de multiples noyaux thalamiques définissant autant de cas particuliers. Pourtant, les principes d'organisation élémentaires du thalamus sont assez simples. Il est en effet pour l'essentiel constitué de neurones de projection glutamatergiques qui contactent principalement des zones précises du cortex cérébral. A l'inverse du cortex, il y a peu d'interneurones locaux et les principales sources d'inhibition sont donc extra-thalamiques (1). Cette relative simplicité a conduit à largement privilégier une analyse des voies anatomiques d'entrée et de sortie de cette structure pour envisager ses fonctions. Au plan histologique, il est relativement aisé d'identifier au sein du thalamus différents regroupements cellulaires qui forment autant de noyaux thalamiques. C'est en fait de leur nombre relativement élevé que résulte la supposée complexité thalamique. En réalité il s'agit donc simplement d'une structure hétérogène qui se singularise par une connectivité forte avec le cortex cérébral.

Thalamus de premier ordre et thalamus d'ordre élevé

Cette hétérogénéité du thalamus a très vite posé la question de la nomenclature à adopter pour mieux cataloguer, et donc mieux comprendre, son schéma d'organisation. L'une des premières propositions conséquentes propose un focus sur la nature des voies de sorties du thalamus vers le cortex qui peuvent être relativement peu nombreuses mais avec de larges terminaisons bien définies (Core), ou alors plus nombreuses, mais de façon plus clairsemée et diffuse (Matrix). C'est la dichotomie Core/Matrix proposée par Jones (2). Une alternative influente issue de la neurophysiologie des systèmes sensori-moteurs propose une approche plus fonctionnelle qui pose la question de la nature des afférences qui arrivent au thalamus. Ces afférences peuvent en effet être classifiées comme *driver,* c'est-à-dire capable de déclencher de l'activité neuronale dans une cible cellulaire jusqu'alors silencieuse, ou alors de type *modulator,* qui sont simplement capables de moduler de l'activité neuronale existante sans pour autant la déclencher. Les afférences de type *driver* sont relativement peu nombreuses mais se constituent typiquement en larges terminaisons qui vont directement innerver les corps cellulaires. A l'inverse les afférences de type *modulator* sont nombreuses, mais de taille plus modeste et innervent les parties plus distales de l'arbre dendritique du neurone cible. La proposition majeure de Ray Guillery et Murray Sherman consiste à classifier les

noyaux thalamiques en fonction de l'origine des afférences de type *driver* (3). Si cette origine est périphérique, il s'agit de noyaux thalamiques qui sont considérés comme de premier ordre (*first-order*, FO). C'est typiquement le cas des relais thalamiques sensoriels. Mais l'origine de ces afférences peut aussi être corticale, comme c'est le cas pour les noyaux thalamiques dits d'ordre élevé (*higher-order*, HO) qui dès lors ne peuvent se concevoir comme de simples relais sensoriels. Dans tous les cas, les noyaux FO et HO reçoivent des afférences modulatrices issues de la couche la plus profonde du cortex (Figure 1).

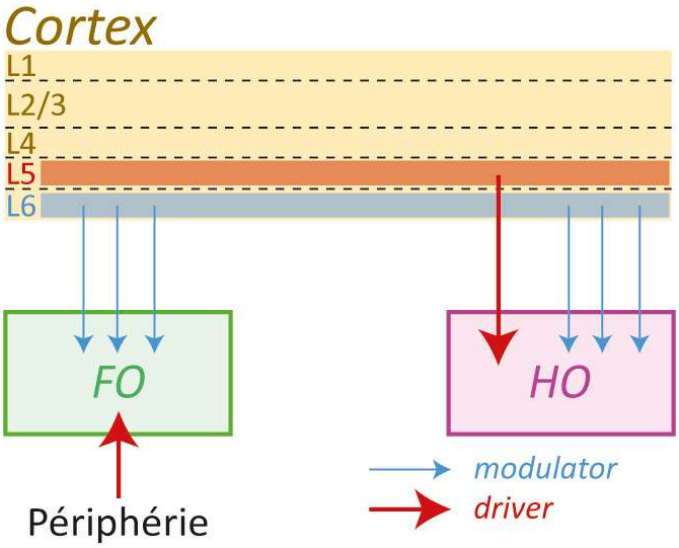


***Figure 1**. Tous les noyaux thalamiques reçoivent une innervation corticale de type modulator issue de la couche 6 mais seuls les noyaux thalamiques de type Higher-Order (HO) reçoivent également une innervation corticale de type driver de la couche 5. Cette dernière provient de la périphérie pour les noyaux de type First- Order (FO).*

Cette nomenclature est toujours influente mais elle est issue d'une recherche portant sur les circuits sensori-moteurs et peut donc perdre en pertinence lorsque l'on s'intéresse aux fonctions intégratives et cognitives qui sont par définition multi- voire a-modales. Surtout, on voit que les différentes nomenclatures proposées se focalisent sur une analyse des entrées et de sorties thalamiques sans considérer la contribution fonctionnelle propre au thalamus, ce qui a certainement contribué à généraliser l'idée classique que cette structure joue un simple rôle de relai.

Du circuit de Papez à une conception moderne des circuits de la mémoire

De façon paradoxale, le rôle important joué par le thalamus dans les fonctions intégratives et même cognitives est apparu de façon assez précoce, avec un effort de recherche qui s'est finalement développé en parallèle du travail de neurophysiologie des circuits sensori-moteurs. On doit en effet au médecin Korsakoff des observations absolument fondamentales sur des patients atteints de démence résultant le plus souvent de consommation d'alcool chronique et de malnutrition. Il rapporte en effet dès la fin du 19ème siècle que ces patients manifestent souvent des troubles mnésiques sévères qu'il rapporte notamment dans les termes suivants : « *the patient constantly asks the same questions and repeats the same stories … may read the same page over and again sometimes for hours…* » (4).

Cette description correspond sans équivoque à une amnésie dite antérograde (une incapacité à mémoriser de nouvelles informations) mais Korsakoff et bien d'autres après lui prennent aussi bonne note du fait qu'elle est souvent accompagnée d'une amnésie rétrograde avec gradient (les souvenirs récents sont perdus mais des souvenirs plus anciens peuvent être conservés). Ce tableau clinique en ce qui concerne l'amnésie est en fait en tout point conforme à celui qui sera établi quelques décennies plus tard par Brenda Milner pour le patient HM qui deviendra l'un des plus fameux patients de l'histoire de neuroscience cognitives (5). Mais là où le cas HM était très clair en ce qui concerne les déterminants biologiques de ses troubles (une opération chirurgicale durant laquelle la formation hippocampique lui fut retiré), les patients atteints du syndrome de Korsakoff ont un lobe temporal intact ! En réalité ce sont des lésions du diencéphale qui caractérisent ces patients. Parmi les structures candidates, ce sont surtout les corps mamillaires et les noyaux thalamiques qui semblent émerger. Il sera établi au début des années 2000 que ce sont les dommages au niveau du thalamus antérieur qui sont le plus systématiquement associés aux troubles mnésiques. Il en résultera un focus sur un circuit connectant l'hippocampe à l'ensemble de ces régions diencéphaliques qui correspond finalement au fameux circuit de Papez, initialement conçu comme le support anatomique hippocampo-mamillo-thalamo-cingulaire de l'expérience émotionnelle. Sous l'impulsion notamment de John Aggleton, ce circuit anatomique est progressivement reconsidéré au plan fonctionnel, d'abord en reconnaissant clairement sa contribution aux processus mnésiques, pour lequel le thalamus antérieur semble crucial, et ensuite en remettant en question son principe d'organisation purement sériel qui ne prend pas suffisamment en compte l'importance de nombreuses projections réciproques. Au final, les bases neurales de la mémoire sont désormais conceptualisées comme un ensemble tripartite avec des flux diencéphaliques et hippocampiques partiellement indépendants mais qui peuvent interagir au niveau des sites corticaux vers lesquels ils convergent (6) (Figure 2).

Cette recherche au long cours a donc finalement permis de mettre en évidence des éléments clés pour

comprendre les fonctions thalamiques en mettant en avant son intégration stratégique au sein de circuits largement distribués et l'importance des interactions avec les régions corticales.

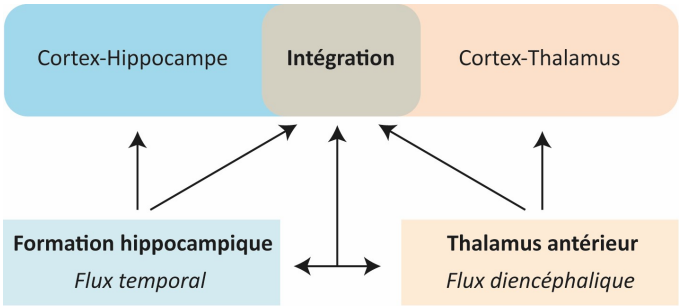


***Figure 2**. Modèle tripartite de la consolidation mnésique. L'hippocampe et le thalamus antérieur contribuent de façon partiellement indépendante à ce processus mais il y a convergence et intégration au niveau cortical en plus des interactions possibles au niveau sous-cortical. Adapté d'après (6).*

Des fonctions thalamiques intégratives

Bien entendu, au-delà du thalamus antérieur, de nombreux noyaux thalamiques contribuent à des fonctions intégratives complexes. Pour reprendre l'exemple de la mémoire, le rôle important d'une région médiane et profonde du thalamus, le complexe *reuniens*/rhomboïde, a émergé ces dernières années pour ce qui concerne tout particulièrement la capacité à former des souvenirs sur le long terme, ce qui implique vraisemblablement un transfert au niveau cortical des traces mnésiques. Il semblerait que cette région thalamique puisse contribuer de façon importante à ce processus de consolidation mnésique en connectant la région hippocampique à différents modules corticaux et notamment préfrontaux (7).

Cette connectivité corticale spécifique est un élément important pour mieux aborder les fonctions associées aux différents noyaux thalamiques. Les régions préfrontales sont particulièrement connectées avec le thalamus, ce qui est en soi un indicateur probant de la nature intégrative des fonctions thalamiques. C'est tout particulièrement le cas pour le thalamus médiodorsal, dont les projections vers l'ensemble du cortex préfrontal furent considérées comme un critère constitutif de ce cortex préfrontal : sont considérées comme préfrontales les aires corticales qui reçoivent une innervation thalamique en provenance du thalamus médiodorsal. La prise en compte de cette particularité anatomique a naturellement abouti à suspecter un rôle tout particulier pour le thalamus médiodorsal (mais d'autres noyaux thalamiques sont également concernés) dans les fonctions intégratives de haut niveau qui sont typiquement associées au cortex préfrontal : contrôle cognitif, flexibilité comportementale, attention focalisée, prise de décision… Pour l'essentiel, un effort de recherche internationale qui a pris de l'ampleur dans les dix dernières années a en effet permis de confirmer que le thalamus médiodorsal semble contribuer de façon spécifique à l'ensemble de ces fonctions dites exécutives (8). De façon intéressante, les données sont globalement en convergence entre ce qui a été accompli chez le rongeur et chez le primate, ce qui accrédite l'idée que les principes fonctionnels à l'œuvre dans les circuits thalamocorticaux sont plutôt bien conservés dans le phylum.

Un exemple qui illustre bien ce point concerne les comportements dits « orientés vers un but » qui s'inscrivent dans le registre comportemental flexible et adaptatif des organismes. L'idée principale est qu'il s'agit d'un mode de contrôle de l'action qui est fondamentalement réflexif et qui dépend notamment de la valeur courante du but poursuivi. Des manipulations comportementales qui peuvent s'effectuer aussi bien chez le rongeur que le primate humain ou non humain, permettent de manipuler la valeur du but pour étudier la prise de décision adaptative des sujets (Figure 3).

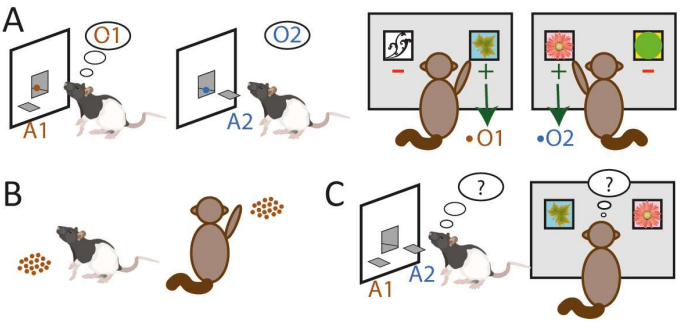


***Figure 3**. Prise de décision adaptative chez le rongeur et le primate. A. Dans une première phase d'apprentissage, les sujets forment des associations spécifiques : l'appui entre un levier spécifique (A) et une récompense alimentaire (O) qui lui est spécifiquement associée pour un rongeur, ou différents items visuels qui peuvent être associés à des récompenses chez le primate. B. On peut modifier la valeur du but en donnant l'une des deux récompenses à volonté aux sujets. Son appétence relative va donc diminuer si on propose au sujet de choisir ensuite entre ces deux récompenses. C. Lorsque le sujet est ensuite en situation de choix avec deux actions (ou items) préalablement associées aux deux récompenses, le sujet va donc manifester une réponse adaptative en choisissant l'option dont la valeur n'a pas été dévaluée. Adapté d'après (9).*

De nombreuses données lésionnelles ou des manipulations plus subtiles (plus sélectives et transitoires) ont ainsi permis de montrer non seulement que l'intégrité fonctionnelle du thalamus médiodorsal est nécessaire pour permettre au sujet d'exprimer un tel comportement adaptif, mais aussi que les interactions fonctionnelles avec le cortex préfrontal sont alors déterminantes (9).

Représentations mentales et synergie thalamique

D'une façon générale, le nouveau tableau qui se dessine est donc celui d'un thalamus largement dévolu à la

cognition avec des fonctions intégratives pour plusieurs noyaux thalamiques qui semblent opérer de façon complémentaire entre fonctions mnésiques (apprendre, se rappeler) et fonctions exécutives (planifier, décider, contrôler). Nous avons respectivement considéré le thalamus antérieur et le thalamus médiodorsal pour aborder ces fonctions mais il est clair que d'autres noyaux thalamiques contribuent. Un point essentiel et qui ressort de façon assez systématique est l'importance des interactions fonctionnelles entre noyaux thalamiques et les régions corticales qui leur sont associées. Ces interactions, de par leur complexité, font actuellement l'objet d'une recherche soutenue pour en décrypter les principes fonctionnels et les mécanismes cellulaires. Des données indiquent que les projections thalamiques pourraient notamment exercer un contrôle fin sur des microcircuits corticaux, permettant ainsi au thalamus de jouer un rôle important dans la capacité à former et à façonner nos représentations mentales. Ces représentations mentales, ou cartes cognitives, sont conçues comme un espace abstrait dans lequel l'inférence peut se développer. Au plan computationnel, c'est une condition nécessaire pour trouver des solutions rationnelles à un problème, s'adapter à des conditions changeantes ou tout simplement faire des choix. Une hypothèse fonctionnelle émergente est donc que le thalamus d'ordre élevé (HO) pourrait jouer un rôle fondamental pour façonner les représentations corticales (10). Différents noyaux thalamiques seraient appelés à concourir de façon complémentaire pour implémenter une telle fonction. Il y a donc vraisemblablement à l'œuvre une véritable synergie thalamique qui opère à l'échelle de multiples circuits thalamocorticaux.

On peut maintenant revenir à l'un des points évoqués en première partie de cet article, à savoir le rôle fonctionnel du thalamus dit HO qui a donc pour caractéristique de recevoir des afférences de type *driver* du cortex. Une idée influente est que les noyaux HO pourraient constituer un lien « transthalamique » pour permettre d'intégrer différents éléments représentés au niveau cortical. Construire des représentations mentales riches et élaborées nécessite en effet de mettre en relation un grand nombre d'éléments qui sont susceptibles d'être représentés dans différentes régions corticales. Les représentations corticales étant par ailleurs souvent complexes et multidimensionnelles, l'un des rôles du thalamus pourrait être de compresser ces représentations corticales pour faciliter une prise de décision rapide et adaptée (8). Ces concepts restent à démontrer formellement mais forment dès à présent un socle fécond qui guide un effort de recherche en ce sens.

Dysfonctions thalamiques et troubles mentaux

Les pathologies affectant le thalamus sont fréquemment associées à des troubles mentaux. Nous avons déjà survolé l'impact profond des lésions du thalamus antérieur sur les fonctions mnésiques et on peut aussi indiquer ici que les affections vasculaires sont également une grande cause de dysfonctionnement du thalamus (11). Ces dernières années ont donc vu émerger de nombreuses données indiquant qu'un trait transversal à certaines pathologies mentales réside dans un syndrome de « dysconnectivité » fonctionnelle au sein de circuits cérébraux largement distribués (12). Dès lors, le thalamus apparait comme un élément critique, au vu de son positionnement stratégique central. Plusieurs de ces pathologies sont donc susceptibles de dépendre, à des degrés divers (soyons prudents et reconnaissons des causes vraisemblablement multifactorielles aux pathologies mentales complexes) de troubles de la connectivité thalamique. L'une des pathologies qui revient le plus fréquemment en exemple est assurément la schizophrénie, avec des éléments probants montrant que l'architecture thalamocorticale est impactée pendant le développement (13). Cela étant, une réduction de la connectivité entre noyaux thalamiques et aires préfrontales est un élément transversal qui l'on retrouve aussi dans les troubles bipolaires ou la dépression. Ces liens qui existent entre troubles cognitifs et troubles de l'humeur nous rappellent que l'on a longtemps qualifié le thalamus de structure « limbique » (14). Au final, les troubles de connectivité au sein des circuits thalamocorticaux apparaissent comme trait commun à nombre de pathologies mentales, soulignant ainsi les enjeux de la recherche en cours visant à mieux conceptualiser les fonctions thalamiques.

Conclusions

Ce survol des fonctions thalamiques met donc clairement en évidence un changement de paradigme. Initialement envisagé comme une structure-relai sans réelle contribution fonctionnelle spécifique, le thalamus apparaît désormais comme un point nodal stratégique au sein de circuits largement distribués. Les neurosciences actuelles ont certainement pour travers un focus un peu trop important sur l'analyse des entrées et des sorties de chaque structure, au détriment d'une réflexion sur leur fonction spécifique. Il en découle une conception pauvre du « relai » dont on peine à saisir l'utilité sans computation spécifique.

Le rôle des noyaux thalamiques a été largement appréhendé au niveau cortical dans cet article mais il est clair qu'une connectivité sous-corticale tout aussi importante est à considérer avec notamment des liens privilégiés avec le système des ganglions de la base et le cervelet. Appréhender les fonctions thalamiques au niveau du cerveau entier est donc un enjeu majeur d'une recherche qui se décline désormais aussi bien chez la personne humaine qu'avec de nombreux modèles animaux, avec pour perspective importante l'apport des approches computationnelles. La période est stimulante et la réunion d'un grand nombre de disciplines qui œuvraient auparavant de façon relativement indépendante ouvre de nouvelles opportunités. En ce sens, conduire une recherche résolument interdisciplinaire associant neurophysiologistes et spécialistes des systèmes sensori-moteurs avec les neurosciences comportementales, cognitives et computationnelles semble la direction la plus prometteuse pour mieux comprendre ce qu'il convient bien d'appeler le thalamus cognitif.

mathieu.wolff@u-bordeaux.fr

Références

(1) Halassa, M. M.; Acsady, L. (2016) Trends Neurosci 39, 680–693.

(2) Jones, E. G. (1998) Neuroscience 85, 331–345.

(3) Sherman, S. M.; Guillery, R. W. J (1996) Neurophysiol 76, 1367–1395.

(4) Kopelman, M. D. (2015) Neurosci Biobehav Rev 54, 46–56.

(5) Scoville, W. B.; Milner, B. (1957) The Journal of neuropsychiatry and clinical neurosciences 12, 103–113.

(6) Aggleton, J. P.; O'Mara, S. M. (2022) Nat Rev Neurosci. 23, 505-516

(7) Ferraris, M.; Cassel, J.-C.; Pereira de Vasconcelos, A.; Stephan, A.; Quilichini, P. P. (2021) Neurosci Biobehav Rev 125, 339–354.

(8) Wolff, M.; Halassa, M. M. (2024) Neuron 112, 893–908.

(9) Wolff, M. (2023) In The Cerebral Cortex and Thalamus; Usrey, M., Sherman, M. S., Eds.; Oxford.

(10) Wolff, M.; Vann, S. D. (2019) J Neurosci 39, 3–14.

(11) Carlesimo, G. A., Lombardi, M. G., Caltagirone, C. (2011) Neuropsychologia 49, 777-789.

(12) van den Heuvel, M. P.; Sporns, O. (2019) Nat. Rev. Neurosci. 20, 435–446.

(13) Benoit, L. J.; Canetta, S.; Kellendonk, C. (2022) Biol Psychiatry 92, 491–500.

(14) Vogt, B. A., et M. Gabriel. *Neurobiology of Cingulate Cortex and Limbic Thalamus*. Édité par B. A. Vogt et M. Gabriel. Boston: Birkhäuser, 1993.

Hommage

Hommage au Professeur Marie-Jo Besson

JOCELYNE CABOCHE, JEAN-ANTOINE GIRAULT, PETER VANHOUTTE, EMMANUEL VALJENT, SYLVIE RETAUX, KAREN BRAMI-CHERRIER & MARIE-LOU KEMEL

MJ Besson, Paris 2003

C'est avec une profonde tristesse que nous avons appris le décès de Marie-Jo Besson, survenu le dimanche 30 mars 2025, à l'âge de 85 ans.

Agrégée en sciences de la vie de l'Université de Nancy, Marie-Jo Besson a rejoint en 1967 le laboratoire de Neuropharmacologie dirigé par le professeur Jacques Glowinski au collège de France. Elle y a débuté une brillante carrière d'enseignante-chercheuse et y a animé une équipe qui a fait des contributions importantes à l'étude de la libération des neurotransmetteurs dans les ganglions de la base (dopamine, acétylcholine, GABA, glutamate). Ses étudiants et ses collègues la reconnaissent unanimement comme une figure majeure de ce laboratoire, dans une période de grande effervescence scientifique.

En 1986, Marie-Jo Besson entreprend une année sabbatique au sein du laboratoire d'Ann Graybiel, au *Massachusset Institute of Technology* à Boston, où elle effectue des travaux pionniers sur la distribution des récepteurs de la dopamine et des neuropeptides dans les compartiments du striatum et leurs structures cibles. Elle rejoint ensuite l'Université Pierre et Marie Curie à Jussieu, en tant que Professeure de Neuropharmacologie, et y monte son propre laboratoire de Neurochimie-Anatomie, jusqu'à sa retraite en 2004.

L'engagement scientifique et la persévérance de Marie-Jo Besson ont permis des découvertes majeures qui ont inspiré de nombreuses recherches tant en France qu'à l'étranger. Son ouverture d'esprit et sa curiosité intellectuelle ont facilité le développement de nouvelles approches dans son laboratoire comme l'étude de la signalisation intracellulaire et de l'expression des gènes. Elle a participé à et coordonné des programmes collaboratifs européens à leurs débuts. Ses travaux ont ainsi contribué de manière décisive à l'avancée de nos connaissances sur la physiologie des ganglions de la base, dans une vision intégrée et fonctionnelle et dans une perspective incluant les pathologies, comme les addictions, la maladie de Parkinson, ou la maladie de Huntington.

Au-delà de ces contributions à la recherche, Marie-Jo Besson a œuvré très activement pour l'enseignement et la formation des jeunes chercheurs en Neurosciences. Elle a dirigé le DEA de Neurosciences, puis l'École Doctorale « Cerveau, Cognition et Comportement » qu'elle a créée, jouant un rôle déterminant dans le développement de la formation en Neurosciences à Paris. Elle a accompagné les étudiants pendant des années, toujours disponible pour les conseiller et les orienter, toujours curieuse de suivre leur parcours.

Nombreux sont les étudiants qui, grâce à son talent pédagogique, se sont passionnés pour l'anatomie et les fonctions des systèmes monoaminergiques, les voies de signalisation associées, et leur rôle dans les fonctions motrices et cognitives des ganglions de la base.

Tout au long de sa carrière, Marie-Jo Besson a su transmettre à de nombreux jeunes chercheurs et chercheuses sa rigueur intellectuelle et scientifique, sa curiosité insatiable et sa bienveillance.

Femme scientifique dans un monde encore très masculin, elle a tracé sa route avec une spontanéité, un dynamisme et une force de caractère dont tous se souviennent, marquant ainsi plusieurs générations de chercheurs.

Nous garderons le souvenir d'une grande dame, avide de partager sa passion et ses connaissances, d'un modèle de femme scientifique au grand talent pédagogique qui a influencé et accompagné des générations de jeunes chercheurs.

Nous lui sommes infiniment reconnaissants.

jocelyne.caboche@sorbonne-universite.fr