

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

CONCEVOIR LES PROCESSUS COGNITIFS COMME DES SUCCESSIONS DE
TALOPASS (ASSEMBLÉES LOCALES ET TRANSITOIRES DE SOUS-SYSTÈMES
PHYSIQUES ET SOCIAUX): UNE RÉPONSE AU PROBLÈME DU
LOCALISATIONNISME DANS LES NEUROSCIENCES COGNITIVES

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN PHILOSOPHIE

PAR

GABRIEL BOUCHARD-JOLY

AVRIL 2023

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.04-2020). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Ce mémoire n'est pas le produit d'un cerveau isolé, déconnecté de son environnement, mais le produit d'une combinaison fructueuse de conditions facilitantes et « permettantes ». Parmi ces conditions, je me dois absolument de mentionner en premier lieu l'aide précieuse de mon directeur de recherche, Pierre Poirier, qui a su, tout au long du processus de rédaction, m'aiguiller avec adresse dans les méandres de la recherche et de la réflexion philosophique. Nombre de ses recommandations m'ont permis de me surpasser épistémiquement et de perfectionner cet ouvrage. Plus que son avis éclairé, je retiens surtout ses qualités humaines qui ont fait de ce processus laborieux une expérience pédagogique agréable de dépassement intellectuel.

Je tiens également à remercier les membres de mon comité d'évaluation, Luc Faucher et Serge Robert, pour leurs commentaires constructifs à l'occasion de la correction de mon chapitre 3, ainsi que certains membres du corps professoral de l'UQAM (ex. Mauro Rossi, Dario Perinetti, France Simard, etc.) avec qui j'ai eu la chance d'entretenir des conversations cordiales et enrichissantes.

Évidemment, je me dois aussi de remercier grandement ma fiancée, Stéphanie Moreau, qui m'a épaulé tout au long de la rédaction de ce mémoire. Comme le veut l'expression « pour le meilleur et pour le pire », Stéphanie m'a accompagné autant dans les moments d'extases intellectuels que dans les moments d'angoisses. Je peux affirmer avec conviction qu'elle a échafaudé ma régulation émotionnelle à travers cette épreuve.

Enfin, ces remerciements ne seraient pas complets si je négligeais de mentionner mes amis Valdez Hilaire, Luke Greau, Camille Corriveau-Bienvenue, Marie-Loup Faubert, Bradley Rolland, et les autres que je ne mentionnerai pas par manque d'espace (ils se reconnaîtront). Bien qu'ils ne soient pas directement impliqués dans le processus de rédaction, ils ont su me soutenir moralement et intellectuellement tout au long des dernières années. Un merci particulier à William Fortin-Auger pour m'avoir aidé à corriger ma posture à la suite de ces longues heures de rédaction. Ma colonne vertébrale souffrante lui sera éternellement reconnaissante.

TABLE DES MATIÈRES

Remerciements	ii
Liste des figures	v
Liste des abréviations, des sigles et des acronymes	vi
Résumé	vii
INTRODUCTION	1
CHAPITRE 1	
Les neurosciences cognitives et la question du localisationnisme	11
1.1 La question de la localisation en neurosciences cognitives : comment vient-on à localiser des facultés mentales ?	15
1.1.1 Les preuves basées sur l'étude des lésions	16
1.1.2 Les preuves basées sur les techniques de neuroimagerie	20
1.1.3 Les preuves basées sur les études d'enregistrement neuronal	26
1.2 Les problèmes avec le localisationnisme en sciences cognitives	28
1.2.1 Le problème définitionnel : qu'est-ce qui porte la marque du cognitif	29
1.2.2 Le problème des frontières de la cognition et de la dispersion spatiale	33
1.2.3 Le problème de la taxonomie du mentale et du niveau de spécificité	39
1.2.4 Le problème de la réalisabilité multiple et du caractère dynamique de la cognition	41
CHAPITRE 2	
Au-delà de la boîte noire, les échafauds extraneuronaux de la cognition	46
2.1 Le cadre théorique de l'échafaudage actif de l'esprit	46
2.1.1 Préambule : qu'est-ce que l'échafaudage actif ?	46
2.1.2 Pourquoi des échafauds ?	52
2.2 Échafaud intrasomatique: comment le corps contribue à la cognition ?	55
2.3 Échafaud intersomatique : comment l'univers social contribue à la cognition ?	59
2.4 Échafaud extrasomatique: comment le monde matériel contribue à la cognition ?	64
2.5 Objections possibles à l'échafaudage de l'esprit	68

CHAPITRE 3	
La cognition comme réseau complexe distribué et dynamique	72
3.1 Une introduction à la science des réseaux	76
3.1.1 Théories des graphes et sciences des réseaux	79
3.1.2 Quelques notions de base sur les types de réseaux	81
3.2 Complexité et problèmes potentiels à modéliser la cognition via une architecture réseautique	87
3.2.1 Un mot sur les systèmes complexes	87
3.2.2 Techniques de réduction de la complexité	90
3.3 Intégration : le cas de la récupération de l'information en mémoire épisodique	98
CONCLUSION	109
LISTE DES RÉFÉRENCES	113

LISTE DES FIGURES

1.1 Modèle d'explication mécaniste en neurosciences cognitives	14
2.1 Saccades oculaires dans l'expérience de Ballard et al. (1995)	56
3.1 Illustration du cycle cognitif dans Sutton et al. (2006)	78
3.2 Représentation visuelle du problème du pont de Königsberg	80
3.3 Variété des réseaux	83
3.4 Procédures de <i>black boxing</i> et de <i>point cutting</i> dans les réseaux	94
3.5 Réseau cérébral de la récupération de l'information en mémoire	99
3.6 Réseau cognitif complet intra- inter- et extrasomatique	101
3.7 TALoPASS incluant l'échafaud extrasomatique	103
3.8 TALoPASS incluant l'échafaud intrasomatique	104
3.9 TALoPASS incluant l'échafaud intersomatique	106

LISTES DES ABRÉVIATIONS, DES SIGNES ET DES ACRONYMES

ASC	Échafaudage actif de la cognition
BOLD	<i>Blood oxygen level dependent</i>
DTI	<i>Diffusion tensor imaging</i>
ECOG	Électrocorticogramme
EEG	Électroencéphalogramme
EOG	Électro-oculogramme
FMRI	Imagerie par résonance magnétique fonctionnelle
ICM	Modèle inter-causal
IRM	Imagerie par résonance magnétique
MEG	Magnétoencéphaographie
MM	Manipulabilité mutuelle
ROI	<i>Region of interest</i>
SMRI	Imagerie par résonance magnétique structurelle
TALONS	<i>Transiently assembled localized neural subsystems</i>
TALOPASS	<i>Transiently assembled localized physical and social subsystems</i>
TECS	<i>Transient extended cognitive system</i>
TSPT	Trouble du stress post-traumatique

RÉSUMÉ

Dans le cadre de ce mémoire, nous examinerons la relation de réduction ontologique et épistémologique des états mentaux aux états cérébraux présente dans les neurosciences cognitives. Cette réduction trouve son point culminant dans le localisationnisme qui consiste, dans ce domaine, à circonscrire les facultés cognitives à l'intérieur de régions cérébrales spécifiques. Nous examinerons critiquement, dans le chapitre 1, les différentes preuves qui soutiennent le localisationnisme et les problèmes philosophiques auxquels cette stratégie explicative fait face. Nous nous intéresserons particulièrement aux problèmes qui concernent la variété ontologique des composantes participant au mécanisme qui réalise les processus cognitifs, la réalisabilité multiple et le caractère dynamique de la cognition. Pour traiter les deux premiers problèmes, nous présenterons, dans le chapitre 2, le cadre théorique de l'échafaudage actif de l'esprit développé par Varga (2018). Nous nous efforcerons, dans les marges de ce cadre, de montrer que différentes composantes extraneuronales contribuent constitutivement à la réalisation et la régulation des fonctions cognitives à partir du critère de démarcation de la manipulabilité mutuelle. Enfin, nous proposerons, au terme de ce mémoire, d'abandonner le localisationnisme comme stratégie explicative en sciences cognitives au profit de stratégies qui mettent de l'avant l'identification de composantes clés à la cognition, leurs propriétés, leurs interactions et la manière dont leurs activités évoluent à travers le temps. Pour rendre compte de ces aspects, nous proposerons de modéliser l'évolution des fonctions cognitives dans le temps comme une série de configurations d'activités transitoires entre les composantes qui participent à leurs réalisations ; ce que nous nommerons des TALoPASS pour « *transiently assembled local physical and social subsystems* ».

Mots clés : Réductionnisme psychoneural, localisationnisme, explication mécaniste, échafaudage de l'esprit, systèmes dynamiques, science des réseaux, TALoPASS.

ABSTRACT

As part of this dissertation, we will examine the ontological and epistemological reduction relationship of mental states to brain states present in cognitive neuroscience. This reduction culminates in localizationism, which consists in circumscribing cognitive faculties within specific brain regions. We will critically examine, in Chapter 1, the various evidence that support localizationism and the philosophical problems that this explanatory strategy faces. We will focus our attention on problems that concern the ontological variety of components involved in the mechanism that realizes cognitive processes, multiple realizability and the dynamic character of cognition. To address the first two problems, we will present, in Chapter 2, the theoretical framework of active mind scaffolding developed by Varga (2018). We will try, within the margins of this framework, to show that different extra-neuronal components contribute constitutively to the realization and regulation of cognitive functions with the use of the demarcation criterion of mutual manipulability. Finally, we will suggest, at the end of this dissertation, to abandon localizationism as an explanatory strategy in cognitive sciences in favor of strategies that put forward the identification of key components of cognition, their properties, their interactions, and the way that their activities evolve over time. To account for these aspects, we will propose to model the evolution of cognitive functions over time as a series of transient activity configurations between the components that participate in their realization; what we will call TALoPASS for "transiently assembled local physical and social subsystems".

Keywords : Psychoneural reductionism, localizationism, mechanistic explanation, scaffolding of the mind, dynamic systems, network science, TALoPASS.

Introduction

Dans son livre de 2008, *Supersizing the mind*, Andy Clark assemble, via une multitude d'exemples qui couvrent les domaines de la robotique, des neurosciences, de la psychopathologie et bien d'autres, une vision globale de la cognition selon laquelle celle-ci serait étendue sur l'organisme et l'environnement. Bien que le présent mémoire n'ait pas autant d'envergure que l'œuvre susmentionnée, il défendra un point commun avec celui-ci : le rejet de ce que le philosophe nomme habilement le *modèle brainbound*¹ de la cognition (qu'on peut traduire par « modèle de la cognition *confinée* au cerveau »). Ce modèle, tel qu'il le définit, correspond à la tendance prévalente que nous avons culturellement et scientifiquement de rattacher, de confiner spatialement, la rationalité et la cognition au cerveau. Loin d'être anodine, cette tendance révèle un problème philosophique important dans les sciences cognitives : la réduction ontologique et épistémologique des états mentaux aux états cérébraux, qui trouve son point culminant dans le localisationnisme. Dans cette perspective, non seulement les états mentaux correspondent aux états cérébraux (relation type-type), mais ces derniers peuvent être rattachés individuellement à des régions spécifiques du cerveau (relation occurrence-occurrence). Ainsi, dans un modèle de la cognition confinée au cerveau, c'est l'activité neuronale seule qui réalise les processus cognitifs humains. Le corps, par contraste, ne serait que le véhicule sensorimoteur liant l'organe cérébral à l'environnement.

Rejeter le modèle *brainbound* de la cognition ne signifie pas que nous endossons une vision où le cerveau ne contribuerait pas ou peu à l'émergence des facultés mentales. Il est, aujourd'hui, indéniable que l'encéphale joue un rôle important dans la réalisation des différents processus cognitifs – nous ne prétendons pas le contraire. Un long historique de recherches cliniques, datant des travaux de Willis (pour plus de détails, voir Molnár, 2004) – et même avant –, nous

¹ Question de ne pas créer d'ambiguïté, nous emploierons, à travers ce mémoire, les expressions « *brainbound* », « cérébrocentrisme » et « réductionnisme cérébral » de manière interchangeable pour désigner à tour de rôle cette attention exclusive qui est portée sur le rôle de l'appareil cérébral dans la cognition (même s'il existe quelques petites différences subtiles entre celles-ci).

rappelle que la cognition est intimement liée à l'organisation du cerveau. Les études classiques de Broca et de Wernicke (Geschwind, 1970 tiré de Tremblay & Dick, 2016) nous montrent, par exemple, qu'il existe un corrélat anatomo-fonctionnel entre certaines régions spécifiques du cortex et différentes composantes du langage, soit respectivement sa production et sa réception, entre autres choses². Ce corrélat est d'autant plus convainquant qu'il repose sur une double association anatomique et fonctionnelle, c'est-à-dire que des lésions à une aire cérébrale x correspond à une perte d'une fonction cognitive y , et, inversement, une diminution de la fonction y équivaut assez fiablement à une atrophie ou à une lésion de la région x en charge de celle-ci. Loin d'être limitées à l'étude des lésions, les neurosciences cognitives cumulent les preuves en faveur de la réduction ontologique et épistémologique des états mentaux à travers les études en neuroimagerie (voir Ganis & Kosslyn, 2002; Poldrack et al. 2011) et en enregistrement neuronal (Bunge & Kahn, 2009).

Il est difficile de dire alors que le cerveau n'est pas impliqué dans certaines fonctions cognitives. Toutefois, qu'il soit, en partie ou dans sa totalité, lié à un processus cognitif ne fait pas de lui l'unique responsable de ce processus ; chose qui semble parfois être endossée implicitement – et même explicitement – par les neurosciences cognitives (voir Mudrik & Maoz, 2014; Frith, 2007; Cacioppo & Decety, 2009; LeDoux, 2002 pour des exemples). Il pourrait n'être, comme le suggère Fuchs (2011), qu'un médiateur entre l'environnement, le corps et l'esprit, qu'un nœud important, parmi d'autres, dans le réseau complexe dynamique et distribué qui donne naissance à la cognition. C'est ce que nous tenterons de démontrer au terme de ce mémoire.

En effet, la grande majorité des opérations mentales et des comportements que nous accomplissons sont liées à un besoin organique et/ou contextuel. Si nous faisons usage de la parole, autant sous forme orale qu'écrite, c'est en premier lieu parce qu'il s'agit d'un outil indispensable pour atteindre nos objectifs: signaler nos déséquilibres homéostatiques (ex. faim, sommeil, etc.), demander de l'information pour éviter des désagréments qui pourraient

² Question d'éviter de dévier du point principal, je me restreins à parler des régions de Broca et de Wernicke de la manière la plus stéréotypée et réductionniste qu'il soit en les ramenant à une division « production/réception ». En réalité, les relations anatomo-fonctionnelles qu'entretiennent ces régions sont beaucoup plus complexes et suggèrent que la contribution fonctionnelle de ces aires se chevauchent au moins partiellement. En outre, l'aire de Broca serait également impliquée dans la perception du langage (Hickok & al. 2012); ce qui contraste bien évidemment avec la vision classique selon laquelle cette région serait dédiée exclusivement à la production du langage.

éventuellement mettre en péril notre équilibre psychophysique (comme demander des indications d'un passant pour éviter de se perdre pendant une longue marche) ou, encore, communiquer une compréhension des neurosciences cognitives dans un mémoire pour obtenir le titre de maître en philosophie et décrocher l'emploi qui assurera, en harmonie avec les valeurs personnelles et les passions de l'auteur de ce travail, une régulation stable de ses états psychophysiques sur le long terme.

Même chose pour l'utilisation de nos capacités logico-mathématiques, mémorielles, visuo-spatiales et les autres; elles prennent naissance dans des contextes spécifiques - rarement dans une salle blanche -, et elles répondent aux besoins de régulation de l'organisme et l'influencent en retour. Ce n'est que dans des conditions ou pratiques humaines très précises et exceptionnelles qu'on peut détacher ces capacités de ces besoins organiques ou contextuels (ex. le cas d'un individu reclus qui résout un théorème mathématique difficile ou celui de deux philosophes qui discutent de questions métaphysiques), et, encore, nous pouvons croire que l'exercice de ces facultés, dans ses situations particulières, pourraient tout de même contribuer à stabiliser la perception qu'un individu pourrait avoir de son identité (comme étant un bon mathématicien ou, simplement, comme étant une personne intelligente); ce qui en retour pourrait améliorer sa capacité de régulation homéostatique via des choix d'actions appropriées (voir par exemple Joffily & Coricelli, 2013).

Pris dans une telle interaction avec l'organisme et l'environnement, il semble mal avisé de limiter *a priori* notre étude des fonctions cognitives au cerveau. La littérature scientifique nous montre que les processus cognitifs sont soutenus et modulés par une myriade de structures extérieures à lui. Que ce soit des exemples comme l'hypothèse des marqueurs somatiques de Damasio (1996; 1998; 2018), où les viscères contribuent à la prise de décision, les systèmes de mémoire transactionnelle (Wagner, 1986; Theiner, 2013), soit les capacités mnésiques qui peuvent émerger de la collaboration d'individus dans un groupe, ou, encore, les nombreux cas d'implémentation d'intention (Orbell et al. 1997; Gollwitzer, 1999; Prestwich et al. 2015), où les objets de l'environnement permettent le rappel ultérieur d'intentions passées, nous voyons comment différents types de structures (ex. notre organisme, notre univers social et notre univers matériel) permettent, facilitent et façonnent la forme que prennent nos processus

cognitifs. Nous tenterons de montrer, dans le chapitre 2, de quelle manière précisément ces structures contribuent à la réalisation de certaines facultés cognitives³.

Pourtant, même si le réductionnisme psychoneural stricte semble être une position indéfendable, il est très largement endossé, autant implicitement qu'explicitement, par les neurosciences cognitives traditionnelles (voir par exemple Mudrik & Maoz, 2014; Frith, 2007; Cacioppo & Decety, 2009; LeDoux, 2002), et contribue à freiner, voir littéralement détourner, les progrès épistémiques et pratiques dans les domaines concernés par celles-ci (ex. sciences cognitives, neuropsychologie, psychiatrie, etc.). Dans la recherche, le cérébrocentrisme – qu'il soit implicite ou non – réduit la lentille d'analyse et exclut au passage des variables d'intérêt pour comprendre la cognition humaine et la psychopathologie. Pire encore, celui-ci contribue dans certains domaines à générer et à préserver des préjugés néfastes, comme c'est notamment le cas avec la dichotomie populaire entre « cerveau masculin » et « cerveau féminin » propulsée par Baron-Cohen (2003; 2004), Brizendine (2007), Gurian (2004), et d'autres, où un profil neurocognitif particulier, *hardwired* et différencié est rattaché à chacun des sexes⁴. Le cerveau des hommes, selon cette perspective, serait organisé de telle sorte que ceux-ci seraient *essentiellement* meilleurs que les femmes en mathématiques et dans les tâches d'orientation visuo-spatiale. À l'inverse, ces dernières seraient fondamentalement meilleures que les hommes en ce qui concerne les facultés linguistiques et relationnelles. Même des traits de personnalité, tels que l'agressivité, la compétitivité et l'émotionnalité, pour ne prendre que quelques exemples stéréotypés, sont attribués aux différences neurobiochimiques et neuroendocriniens observées entre les sexes (Hoffman, 2012; Fine, 2012).

³ Il ne faut pas se méprendre, cependant. En parlant de ces différentes structures, on n'évoque pas ici que de simples relations de possibilité (ex. sans x , y n'existe pas) ou de relations facilitantes (ex. x facilite y , sans pour autant contribuer causalement à sa réalisation), mais de relations de constitution, c'est-à-dire que les éléments qu'on postule comme étant impliqués dans la réalisation des fonctions cognitives appartiennent au mécanisme responsable de l'émergence de ces fonctions. Nous en parlerons plus en détail dans les chapitres 1 et 2.

⁴ Comme le but n'est pas de traiter spécifiquement de la relation entre sexe et genre, je m'en tiendrai ici à une distinction simple entre les deux, question d'être clair et de ne pas alourdir inutilement mon propos. Ainsi, j'utiliserai le mot « sexe » pour désigner l'identité biologique d'un individu, déterminer par les caractéristiques sexuelles primaires, et « genre » pour parler de l'identité psychologique (e.g. traits de personnalité, fonctions cognitives, etc.) de cet individu, même si les deux sont, en réalité, intimement liées à un point tel qu'il est considéré comme insensé de les détacher l'un de l'autre (Kaiser, 2012).

Cette vision réductionniste du genre au cerveau est hautement problématique parce qu'elle laisse à croire en un fatalisme biologique: une personne de sexe féminin ne sera, selon cette optique, jamais aussi compétente qu'une de sexe masculin dans le domaine des mathématiques en raison de l'organisation intrinsèque de son cerveau (ou du moins cela lui pendra beaucoup plus d'effort qu'à un homme, car la maîtrise des mathématiques irait contre ses dispositions naturelles). Cela crée et renforce une injustice profonde quant à l'égalité des chances entre les sexes dans différents domaines associés aux facultés cognitives et aux traits de personnalité considérés comme genrés (on peut penser à la femme qui aspire à une carrière d'ingénieure ou l'homme qui désire faire carrière dans l'enseignement primaire). Alors qu'en réalité, ces présumées dissimilitudes pourraient s'expliquer en bonne partie par des facteurs extraneuronaux, qui sont, hélas, masqués par la place prépondérante des explications neurocentrées sur la scène explicative, telles que la position sociale de l'individu (Maibom & Bluhm, 2014) et, plus significativement, parmi d'autres, la manière dont les institutions en place dans nos sociétés contraignent et régulent différenciellement les opportunités de production et la forme de la cognition chez l'homme et la femme à titre de constituants du système cognitif (pour un bon exemple, voir Merritt, 2013 qui applique avec doigté la thèse des institutions mentales défendue par Gallagher & Crisafi, 2009)⁵. En réduisant le profilage cognitif des sexes à l'organisation anatomo-fonctionnelle du cerveau on passe donc sous silence toutes ces variables pertinentes à la compréhension des différences entre les hommes et les femmes, et on transforme au passage ce qui ne serait fort probablement qu'une différence d'organisation cérébrale individuelle en une différence intergroupe entre les sexes.

Dans la pratique clinique, les préjugés cérébrocentriques guident les cibles de traitement et les priorités thérapeutiques de manière significative. Prenons par exemple la schizophrénie, qui est considérée comme l'un des « symboles sacrés » de la psychiatrie (Sass et al. 2018). L'une des

⁵ En effet, Gallagher et Crisafi, dans leur article de 2009, démontrent que des institutions, comme le système juridico-légal, participent à la cognition de manière analogue aux exemples classiques donnés par Clark et Chalmers (1998). Les institutions sont accessibles fiablement lorsque nécessaire et les informations qu'elles diffusent peuvent être acceptées d'une façon similaire à la façon qu'elles le seraient si elles étaient présentes dans le cerveau. Par ailleurs, il est possible de les modifier conformément à des besoins épistémiques et cognitifs sociétaux (ex. en changeant les pratiques départementales d'un Cégep ou d'une Université), modifications qui en retour offriront des opportunités cognitives aux individus et qui pourraient mener à d'autres révisions de ces institutions. Dans cette optique, on peut dire que ces institutions font partie de la cognition. Merritt reprend d'ailleurs cette thèse pour montrer comment l'étiquette du genre se comporte comme une institution et comment il façonne et est façonné réciproquement par les comportements et les idées véhiculés sous son égide.

explications majeures de cette psychopathologie est l'hypothèse neurodéveloppementale proposée conjointement par Weinberger, et Murray et Lewis (Dans Jablensky et al. 2017). Selon cette hypothèse, l'installation de ce trouble au début de l'âge adulte serait la résultante de l'interaction entre l'effet des gènes et de lésions cérébrales prénatales. Les effets de ces lésions, comme Jablensky et son équipe le notent, seraient latents et apparaîtraient lorsque les structures endommagées, particulièrement le cortex préfrontal dorsolatéral, seraient recrutées dans l'activité cognitive plus tardivement dans la vie de l'individu. Plusieurs études remarquent d'ailleurs une perturbation glutamatergique sous-jacente qui serait en cause dans le dysfonctionnement de cette région cérébrale et d'autres impliquées dans le portrait clinique de la schizophrénie (Takahata & Moghaddam, 2003; Goldman, 2011; Li et al. 2020, pour ne prendre que quelques exemples). Ainsi, il n'est pas étonnant qu'une des cibles majeures de traitement pour ce trouble soit le glutamate, ou plus précisément les enzymes GCPII qui auraient le potentiel d'influencer la concentration extracellulaire de ce neurotransmetteur (Zhou et al. 2005).

Cette préférence pour cette cible de traitement, en l'occurrence le glutamate, influence, en retour, les priorités thérapeutiques adoptées à l'égard de la schizophrénie. Plutôt que de songer, en premier lieu, à une thérapie cognitive-comportementale ou à évaluer le contexte familial et matériel dans lequel se trouve l'individu, le psychiatre peut être amené à privilégier une intervention médicamenteuse qui traiterait spécifiquement la dysrégulation de ce neurotransmetteur, sans s'intéresser aux autres avenues thérapeutiques. Cela n'est évidemment pas de dire qu'il s'agit d'une intervention mal intentionnée ou mal informée de la part du psychiatre. Vu la nature et l'intensité des symptômes présents dans certains cas, la prise d'un médicament se présente comme un outil crucial pour atténuer le risque que l'individu schizophrène se blesse ou ne blesse un autre.

Cependant, il ne faut pas réduire l'essence de cette psychopathologie au rôle seul du glutamate – et des autres molécules impliquées (ex. dopamine, sérotonine, etc.). La schizophrénie est, certes, un trouble qui présente une signature neuronale et moléculaire particulière, mais elle n'est pas ontologiquement que cela. Elle est également un agrégat de symptômes interreliés qui prennent racines dans un couplage unique et complexe entre un individu, son historique personnel de développement (Andreasen, 2010) et son environnement social et matériel (Leask, 2018). La schizophrénie n'affecte pas que le cerveau, mais l'ipséité de l'individu touché, soit son rapport avec lui-même, son soi, et sa capacité à s'insérer activement dans le monde autour

de lui. Comme Sass et al. (2018) le font remarquer dans leur modèle bio-phéno-social, la perturbation du soi, dans la schizophrénie, repose sur trois aspects principaux en interaction: un sentiment de présence de soi diminué, une hyper-réflexivité et une saisie appauvrie sur le monde. Ces aspects incluent une composante neuronale, Sass et ses collaborateurs parlent d'ailleurs d'un problème au niveau du réseau de saillance et un qui concerne les copies d'efférences, mais ils ne se réduisent pas qu'à cela.

L'individu schizophrène adopte une posture particulière par rapport à ses activités internes qui l'amène, en relation avec ces variables neurologiques, à les considérer comme étrangère à lui. Cette hyper-réflexivité l'empêche ainsi de saisir, pour reprendre les propos de Krueger (2018), la gamme variée d'*affordances*, de possibilités interactives, que lui offriraient en temps normal les différents objets et individus qui peuplent son univers, de sorte que, les différentes rencontres avec ces choses apparaissent à la personne schizophrène comme dépourvues de sens. Se concentrer uniquement sur les neurotransmetteurs et les régions cérébrales touchées par cette psychopathologie nous priverait d'une compréhension plus complète des mécanismes à l'œuvre dans celle-ci, mais aussi elle nous amènerait à négliger des avenues thérapeutiques possibles. On pourrait imaginer, par exemple, parmi ces avenues, une thérapie axée sur la rééducation motrice, sociale et émotionnelle autour des différents types d'*affordances*, mais de telles considérations vont bien au-delà des objectifs du présent travail.

C'est précisément ce genre de problèmes fondamentaux (ex. la réduction d'un sexe à un profil neurologique particulier) et pratiques (ex. le traitement clinique de la schizophrénie) qui motivent la critique du localisationnisme que nous tenterons de dresser dans ce mémoire. Le modèle *brainbound* de la cognition tel qu'il est présent dans les neurosciences cognitives et les disciplines subsumées par celles-ci met en péril une investigation profonde et complète des variables déterminantes dans le fonctionnement cognitif, autant normal qu'anormal, contribue à stigmatiser des groupes sociaux sur la base de données incomplètes (e.g. cerveau masculin et cerveau féminin) et compromet potentiellement l'efficacité des avenues thérapeutiques envisagées pour le traitement de certaines psychopathologies.

Avec cette idée de critique en tête, ce mémoire a trois objectifs principaux: soulever les problèmes liés au localisationnisme dans les sciences cognitives, identifier quelques-unes des composantes extraneuronales qui participeraient constitutivement au système cognitif et proposer une modélisation pour capturer et intégrer certains des mécanismes responsables de

l'émergence et de la régulation des facultés cognitives qui ne repose pas sur la localisation des composantes et des processus. Pour atteindre ces objectifs, nous diviserons cet ouvrage en trois chapitres.

Dans le premier chapitre « Les neurosciences cognitives et la question du localisationnisme », il sera question de la place particulière qu'occupe le localisationnisme sur la scène explicative des sciences cognitives; plus précisément des preuves que les neurosciences cognitives fournissent en faveur de la réduction des processus cognitifs à des états neurologiques du cerveau. L'étude sélective des différentes méthodes de collecte de données (ex. étude des lésions cérébrales, neuroimagerie et enregistrement de l'activité cérébrale) employées dans ce domaine permettra de dégager certains manquements conceptuels et méthodologiques auxquels nous tenterons une réponse dans les chapitres suivants. En particulier, nous verrons dans ce chapitre que la localisation est non seulement inadéquate comme méthode d'accroissement des connaissances en sciences cognitives, mais qu'elle est carrément impraticable en raison de la variété ontologique des composantes appartenants aux systèmes cognitifs et du caractère dynamique de la cognition.

Dans le deuxième chapitre, « Au-delà de la boîte noire, les échafauds extraneuronaux de la cognition », nous examinerons la possibilité que certaines structures extraneurales contribuent constitutivement à l'émergence et la régulation de certains processus cognitifs. Pour ce faire, nous présenterons le cadre théorique de l'échafaudage actif de la cognition (ASC) introduit par Varga (2018). Ce cadre nous permettra d'éviter quelques débats stériles sur la nature de la cognition (ex. définition de la cognition, marque du cognitif, etc.) et nous permettra également d'explorer le rôle de certaines structures – les échafauds – dans la production de la cognition, comme l'organisme, l'univers social et l'univers matériel de l'individu. Le point focal de ce chapitre sera que les échafauds ne contribuent pas tous de la même façon et aux mêmes moments à la cognition, que certaines de ces structures ne participent que distalement ou très indirectement à l'émergence et la régulation des processus cognitifs, et qu'il faut, dans cette perspective, se prémunir d'un critère efficace pour distinguer celles qui sont pertinentes explicativement de celles qui le sont moins ou pas du tout. Ce critère sera celui de la manipulabilité multiple introduit dans le chapitre 1.

Enfin, dans le troisième et dernier chapitre « La cognition comme réseau complexe et distribué », nous soutiendrons, à la lumière des éléments discutés dans les chapitres précédents, que la question de déterminer le lieu d'émergence de la cognition est inadéquate ontologiquement et méthodologiquement. Vu la variété ontologique et la dispersion des éléments qui participent à l'émergence et la régulation des processus cognitifs, la question de sa localisation, le « où », devrait être remplacée par celle de sa composition, le « quoi »; de *quoi* cette cognition est composée pour prendre la forme particulière qu'elle a, et, par extension, quelle est la pondération relative de chacune des composantes identifiées dans l'émergence moment-par-moment de celle-ci? Pour répondre à cette nouvelle question (le quoi), nous suggérerons de modéliser la cognition comme une succession de configurations d'activité transitoires entre différentes composantes neuronales et extraneuronales, ce que nous appellerons les *transiently assembled local physical and social subsystems* (TALoPASS). Cette suggestion est inspirée de concepts tels que les TALoNS et les TECS développés par Anderson (2014) et Clark (2010) respectivement.

Cette modélisation devrait, ultimement, nous permettre de capturer – si ce n'est qu'un instantané – certaines des multiples variables et opérations qui participent à l'émergence de la cognition, autant normale « qu'anormale », tout en contournant certains des problèmes qu'engendrent l'entreprise localisationniste, particulièrement celui qui consiste à réduire ontologiquement la cognition à un processus statique qui ne se réaliserait qu'à l'intérieur du cerveau. Elle devrait également enrichir certains débats existants présentement en sciences cognitives, comme celui portant sur le niveau adéquat de spécificité à adopter (*fine-grained* contre *coarse-grained*) à l'égard de certaines facultés cognitives et nous informer sur les relations fonctionnelles entre celles-ci, par l'étude des transitions dynamiques entre les états du système cognitif – les configurations d'activités que sont les TALoPASS.

Enfin, il est bien de noter, avant de débiter, que nous utiliserons tout au long de cet ouvrage le processus de récupération de l'information en mémoire comme exemple de travail. Nous jugeons qu'il révèle bien, par son enchevêtrement avec d'autres fonctions cognitives, la complexité inhérente à l'étude des fonctions cognitives (ex. relations fonctionnelles avec d'autres processus cognitifs, niveau de spécificité à adopter, classification taxonomique, etc.). Plus que cela, la récupération de l'information en mémoire met en relief la contribution respective des différents échafauds et les interactions dynamiques qui existent entre ceux-ci à travers le temps. Mais ne sautons pas d'étapes. Commençons par les prolégomènes de ce récit:

comment en sommes-nous arrivés à une équivalence identitaire entre le cerveau et la cognition?
C'est ce que nous allons voir dans le chapitre à suivre.

Chapitre 1

Les neurosciences cognitives et la question du localisationnisme

Nous avons mentionné à quelques reprises depuis le début de ce texte l'expression « neurosciences cognitives », sans vraiment définir en quoi consiste celle-ci. L'intuition voudrait qu'il s'agisse simplement d'une intersection disciplinaire entre les sciences neurologiques et les sciences cognitives, mais le portrait véritable de cette intersection est beaucoup plus complexe que cela. L'expression « neurosciences cognitives » reflète en fait un changement paradigmatique important dans la sélection des types d'explications prédominantes en sciences cognitives, que Boone et Piccinini (2015) soulèvent avec adresse. Ces derniers soulignent que cette fusion des sciences cognitives avec les neurosciences prend l'allure d'une « sorte d'oxymore ». C'est qu'il semble en effet que les neurosciences cognitives constituent un mariage très particulier entre deux perspectives disjointes: une perspective computationnaliste qui trouve sa voix dans les approches symboliques-représentationalistes des sciences cognitives traditionnelles et une perspective physicaliste qui trouve écho dans les sciences neurologiques.

En effet, avant d'être intégrées aux neurosciences, les sciences cognitives étaient dominées par les approches symboliques-représentationalistes (McClelland, 2014). La cognition, dans cette perspective, est considérée comme un système de manipulation de représentations internes, analogue à un ordinateur avec des règles qui régissent le genre de manipulations qui doivent être effectuées sur les symboles présents dans le système. Ce sont les règles de manipulation symbolique qui sont au cœur de l'explication dans ces approches. Par exemple, Braine et O'Brien (1998) soutiennent que le raisonnement humain est le produit d'une logique mentale, soit l'application de règles d'inférence sur des propositions formelles (ex. P implique Q, et P, alors Q). Ici, il n'est donc pas question de lier des processus cognitifs à des états physiques dans le monde, mais de les expliquer via les règles de la logique et des représentations propositionnelles de ce monde.

Il est pertinent de noter que l'analogie avec l'ordinateur est loin d'être insignifiante, dans ce contexte, car elle a motivé une vision particulière du fonctionnement cognitif. En effet, si on peut prouver que l'ordinateur fait de la cognition au même titre qu'un agent humain, alors qu'il est composé de matériaux différents de ce dernier, il est alors tentant de dire que

l'implémentation physique exacte (ex. le cerveau) de la cognition n'a pas vraiment d'importance pour comprendre celle-ci. Elle peut être réalisée de plusieurs façons – c'est d'ailleurs la thèse de la réalisabilité multiple (pour plus d'information, voir Barrett, 2013) –, et il faudrait donc davantage s'intéresser aux règles de manipulation symbolique et à la relation existante entre les fonctions cognitives⁶ qu'à la matière qui produirait cette cognition.

À l'opposé de cette perspective, les neurosciences soutiendront que les facultés mentales prennent leur origine dans un type de matériel très spécifique: les cellules neuronales, et qu'il est possible, en vertu de cette implémentation physique, de comprendre la cognition en étudiant les propriétés et l'organisation de ces cellules. Contrairement aux approches symboliques, les explications en neurosciences ne portent pas sur des règles de manipulation ou d'inférence logique qui s'appliqueraient sur des représentations internes, mais sur des mécanismes. Craver (2008) définit les mécanismes comme étant « des collections d'entités et d'activités intégrées au sein de la production de changements réguliers allant des conditions de départ ou de mise en place aux conditions terminales ou d'achèvement ». Pour le dire autrement, un modèle mécaniste expliquerait la manière qu'un système (ex. celui de la cognition) change d'un état initial à un autre état en faisant appel aux parties de celui-ci et aux interactions qui existent entre celles-ci.

Toujours selon Craver (2007), les explications mécanistes possèdent trois caractéristiques fondamentales: 1) elles décrivent des mécanismes, cela, 2) sur plusieurs niveaux hiérarchiques verticaux (moléculaire, neuronal, intégratif), et 3) intègrent les différentes données issues de plusieurs domaines scientifiques (physique, chimie, biologie, etc.). L'interaction de ces caractéristiques culmine en un effet d'*émergence* où les phénomènes observés à un niveau E (e.g. lobe frontal) sont directement émergents des propriétés et des interactions entre les composantes de E du niveau D (e.g. neurones), qui, à leur tour, émergent des propriétés et des interactions entre les composantes de D au niveau C⁷ (e.g. molécules), et cela itérativement jusqu'au niveau d'explication le plus bas concevable (ex. les éléments subatomiques). Il existe

⁶ Loin l'idée de confondre computationnalisme et fonctionnalisme, ces deux aspects sont regroupés simplement à des fins pratiques pour montrer l'attitude qui régnait à l'égard des explications dans les sciences cognitives des années 50-70.

⁷ Ici, le choix des lettres est arbitraire et ne reflète pas le nombre réel de niveaux d'explication qu'on pourrait rencontrer en essayant d'expliquer un phénomène en neurosciences. Les lettres A, B et C auraient pu aussi bien être sélectionnées pour l'exemple, mais l'utilisation des lettres C, D et E a été privilégiée puisqu'elles ouvrent la porte à la possibilité qu'il y ait des niveaux d'explications inférieurs à C et supérieurs à E.

de ce fait des explications qui tentent de retracer un phénomène aussi complexe que la conscience aux mouvements hasardeux d'atomes (voir, par exemple, Hameroff & Penrose, 2014).

Plus important encore, le phénomène émergent (ex. la fonction cognitive étudiée) est lié à des modifications dans les composantes et les activités du niveau explicatif inférieur à celui-ci (ex. des régions cérébrales) – sans quoi il ne pourrait être expliqué exclusivement par ces éléments. Pour que ces éléments aient une pertinence explicative relativement au phénomène d'intérêt, il doit y avoir des relations de co-dépendance entre ceux-ci, de sorte qu'il ne peut pas y avoir de modifications dans le phénomène observé sans qu'il n'y ait de modifications dans les éléments constituant le système identifié pour le dit phénomène (le système serait dans ce cas-ci un modèle inter-niveaux des éléments et activités qui contribueraient à l'émergence de la cognition). C'est, dans ses grandes lignes, la thèse de la *survenance* (pour plus d'information, voir Mandik, 2011; Harbecke, 2014).

Émergence et survenance sont deux facettes d'une même médaille. L'émergence saisit la manière qu'un phénomène particulier naît des propriétés et des interactions entre les différentes composantes présentes à un niveau inférieur d'explication dans le système conçu pour étudier ledit phénomène, et la survenance saisit la relation de codépendance entre ces composantes et le phénomène d'intérêt. La **figure 1** ci-dessous montre à quoi ressemble un modèle d'explication mécaniste typique dans les neurosciences cognitives qui inclut les relations d'émergence et – plus implicitement – de survenance.

Nous allons voir dans les sections qui suivent – et plus précisément avec l'exemple de la mémoire – comment les explications en termes d'émergence-survenance occupent une place importante dans les neurosciences et comment ces explications sont influencées par et influencent à leur tour la localisation de mécanismes dans le cerveau. Cette relation émergence-survenance reviendra également dans le chapitre 2 lorsqu'on parlera de la question de la manipulabilité mutuelle (Craver, 2007; Varga, 2018) et dans le chapitre 3, lorsqu'on parlera des *transiently assembled local physical and social subsystems* (TALoPASS). Bien que la direction explicative ne sera pas la même, beaucoup des éléments discutés ici reviendront dans notre discussion des systèmes réseautiques complexes et dans la question de l'émergence des facultés cognitives à partir de l'organisation de ceux-ci.

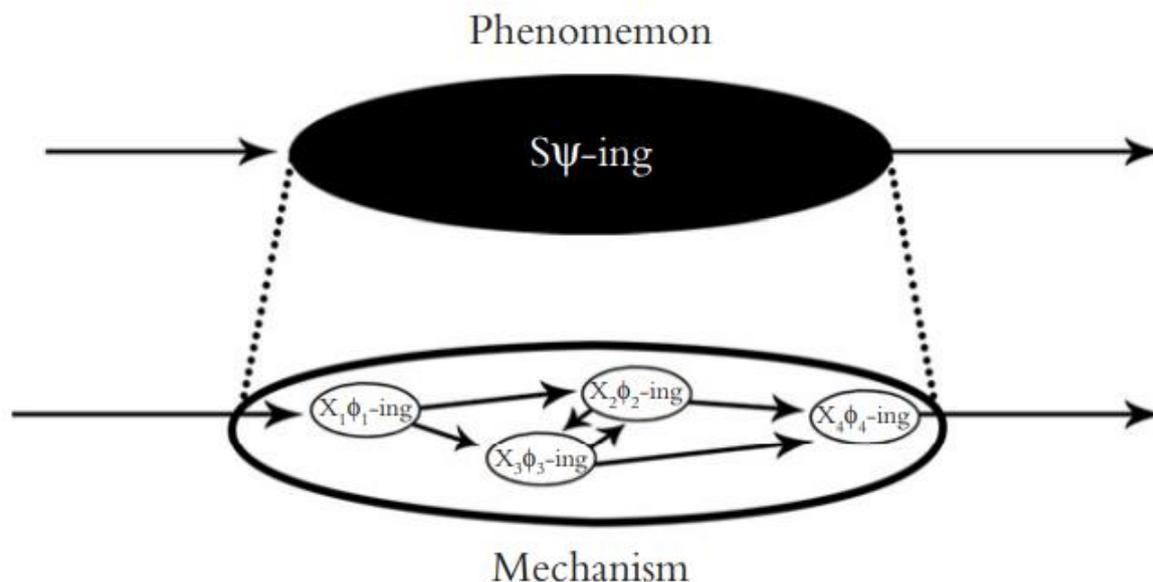


Figure 1.1 - Représentation visuelle d'une explication mécaniste en neurosciences cognitives. La pastille noire correspond à la fonction cognitive (Ψ) émergente qu'on tente d'expliquer - le phénomène - alors que la pastille blanche, au-dessous de celle-ci, correspond au mécanisme explicatif. Nous voyons que celui-ci est constitué de différentes parties en interaction. Bien qu'il ne soit pas indiqué clairement dans le modèle, on pourrait visualiser des relations de survenance entre les composantes et le phénomène Ψ à l'aide de flèches centrales qui, à partir des composantes du mécanisme, pointeraient vers le phénomène étudié. Source: Craver (2007).

Enfin, un mot général sur le déroulement de ce chapitre. Celui-ci sera divisé en deux sections principales. Dans la première, « La question de la localisation en neurosciences cognitives », nous examinerons sommairement, mais de manière critique, les méthodes et les preuves amassées par les neurosciences en faveur de la réduction onto-épistémologique des facultés mentales à des états cérébraux, et comment la démarche localisationniste est enchevêtrée dans cette réduction. Nous procéderons, ensuite, dans la deuxième section, à la présentation de différents problèmes associés au localisationnisme; problèmes qui s'appliquent autant, à notre avis, aux versions réductionnistes qu'émergentistes. Nous verrons comment des *a priori* conceptuels, comme la définition qu'on accorde à la cognition et l'étendue conceptuelle des fonctions cognitives, peuvent impacter l'intention même derrière la recherche d'un *locus* pour la cognition. Plus important encore, nous verrons comment des problèmes, comme celui de la dispersion spatiale, de la réalisabilité multiple et du caractère dynamique de la cognition, rendent la localisation de la cognition impossible.

1.1. La question de la localisation en neurosciences cognitives: comment vient-on à localiser des facultés mentales?

La première question que nous pouvons nous poser à l'égard de la localisation des facultés mentales en neurosciences cognitives, et qui fait l'objet du titre de cette section, est comment en vient-on justement à localiser ces facultés, et, plus précisément, les localiser dans le cerveau? Comment passe-t-on de facultés intangibles comme la mémoire ou l'imagination à des systèmes cérébraux spécifiques? La réponse à cette question passe par une réduction épistémologique et/ou ontologique des états mentaux à des états cérébraux (pour plus d'information, voir Bickle, 2003, Kaplan, 2009; Bechtel, 2010). Dans le premier cas, on considère simplement que les deux types d'états sont mutuellement informatifs à l'égard de l'un et de l'autre. On peut, par exemple, soutenir que les fonctions cognitives dérivent ou peuvent être prédites à partir de certains états cérébraux sans nécessairement se réduire à ceux-ci. Toutefois, comme ces deux types d'états seraient distincts, dans cette perspective, on peut aussi penser que la portée des connaissances acquises serait limitée (ex. est-ce que c'est réellement le changement d'activité cérébrale ou la présence d'un phénomène médiateur qui conduirait au changement observé dans l'activité cognitive?). Nous sommes d'accord avec une réduction épistémologique partielle où l'activité cérébrale pourrait nous renseigner sur quelques-uns des mécanismes participant aux états mentaux, sans, néanmoins, être la source absolue des informations à leur sujet.

Dans le deuxième cas, qui est plus lourd d'implications, nous remédions à cette limitation en postulant que facultés cognitives et états cérébraux sont deux manifestations d'une seule et même chose. La réduction ontologique comprend la réduction épistémologique. En effet, comme les deux entités sont considérées comme étant une seule et même chose sur le plan ontologique, les informations obtenus sur les composantes et les opérations de l'un devraient trouver leur équivalence dans les composantes et les opérations de l'autre – elles seraient en principe la même chose. Les subdivisions de la cognition devraient, donc, dans cette optique, correspondre à des subdivisions du cerveau et vice-versa. Nous allons revenir sur la question de ces subdivisions un peu plus tard dans cette section lorsque nous parlerons des problèmes de la taxonomie du mental et du niveau de spécificité et de leurs implications pour l'établissement de corrélations anatomo-fonctionnelles entre le cerveau et la cognition.

La deuxième question que nous pouvons nous poser relativement à la localisation des facultés cognitives, et qui découle directement de la réponse à la première, est la suivante: « comment en vient-on à accepter les réductions épistémologiques et ontologiques de la cognition au cerveau? ». Autrement dit, qu'est-ce qui nous motive à accepter ce genre de réductions? Il est de notre avis que l'acceptation de ces réductions passe par la présence de différents types de « preuves », ou données, en neurosciences cognitives (Bechtel dans Poirier & Faucher, 2008). On peut regrouper ces types de preuve, selon une suggestion de McClelland (2001), en trois catégories: les preuves basées sur les études lésionnelles et cognitivo-comportementale, les techniques de neuroimagerie et les études d'enregistrement neuronal. Il faut préciser toutefois que ce dernier ne parle pas de preuves, mais de méthodes.

On peut à nouveau distinguer ces types de preuves selon qu'elles reposent sur l'objectif de mesurer ou de manipuler l'activité cérébrale (Ruff & Huettel, 2014), question d'être plus précis. Inutile de dire que les méthodes qui manipulent l'activité cérébrale possèdent une meilleure validité en ce qui concerne la localisation d'une fonction que celles qui ne font que la mesurer, comme elles permettent d'établir une relation entre une cause, une manipulation de la part de l'expérimentateur, et un effet, un changement dans la performance cognitive (c'est d'ailleurs l'idée centrale de l'approche interventionniste; voir le chapitre qu'y consacre Varga, 2018 et les travaux de Woodward, 2010, 2015, entre autres). Il n'est pas impertinent de noter que ces trois catégories de preuves reposent d'une manière ou d'une autre sur l'observation comportementale ou cognitive; c'est après tout les éléments que les neuroscientifiques tentent de localiser dans le cerveau (Uttal, 2001). Il est donc important d'avoir à notre disposition une bonne définition de ce en quoi consiste ces éléments si nous désirons les étudier convenablement. Mais ne sautons pas d'étapes, prenons quelques instants pour discuter individuellement de ces catégories de preuves afin d'évaluer leurs forces et leurs faiblesses.

1.1.1. Les preuves basées sur l'étude des lésions

D'abord, nous pouvons commencer par parler des preuves issues de l'étude des lésions cérébrales et de leurs effets sur le comportement et la cognition humaine. De toutes les catégories de preuves susmentionnées, il s'agit sans doute de la plus vieille. On peut retracer ses balbutiements à l'antiquité avec les travaux de médecins anatomistes tels qu'Alcméon de Crotona (Celesia, 2012), Hippocrate (Breitenfeld et al. 2014), Erasistrate (Stefanou, 2018) et

Galien (Rocca, 1997) qui ont été parmi les premiers à identifier des relations anatomofonctionnelles entre certaines régions du cerveau et des fonctions sensorielles, motrices et même cognitives.

Le principe derrière les preuves lésionnelles est relativement simple. Dans un cas, on remarque l'altération chez un individu d'une faculté cognitive particulière (ex. la compréhension du langage) qui correspond à une lésion cérébrale spécifique (ex. le gyrus de Heschl) observable *in vivo* ou à l'autopsie^{8 9}. Dans le cas inverse, une lésion cérébrale peut être remarquée en premier - que ce soit à la suite d'un examen neurologique, d'un accident ou d'une lésion expérimentale provoquée sur un animal (voir Jiang et al. 2000; Fijalkowski et al. 2006 pour des exemples)¹⁰ – et mener à l'identification de dysfonctions cognitives particulières.

L'intérêt de parler de ces deux trajectoires inverses, du cognitif au neurologique et du neurologique au cognitif, est celui d'un rappel; que les conclusions anatomofonctionnelles établies dans les sciences neurologiques, comme la neuropsychologie et la neurologie, reposent sur un dialogue en constante évolution entre un modèle des effets d'une lésion (voir par exemple Alstott et al. 2009; Yuan et al. 2017; Zhang et al. 2021) – quelles conséquences fonctionnelles devrait entraîner une lésion, disons, du lobe préfrontal gauche – et les effets observés suivant une lésion de cette région dans un cas clinique particulier, l'aphasie, l'agrammatisme, etc. S'il y a adéquation entre les deux, le modèle en vigueur se trouvera à être fortifié (tout va bien). En revanche, dans le cas où il y a inadéquation, une attention particulière devra être portée sur les

⁸ Évidemment, la relation entre fonctions cognitives affectées et lésions est plus complexe que ce que laisse croire ce portrait sommaire. En réalité, il arrive souvent de retrouver des cas où une lésion focale entraînerait des perturbations cognitives sur plusieurs domaines (Arbula et al. 2020). Il arrive également, à l'inverse, qu'une perturbation cognitive soit la résultante de plusieurs lésions diffuses. C'est le cas notamment des dysfonctions qu'on remarque après des traumatismes craniocérébraux (Arciniegas et al. 2002) ou dans le cas des maladies neurodégénératives (voir par exemple Deloire et al. 2005).

⁹ C'est, pour prendre un exemple classique, le cas du patient de Broca « Tan » qui, étant affligé par une paralysie des membres droits, n'a eu d'autres choix que d'être pris en charge par ce dernier (voir Broca, 1861). Or, il s'avérait que « Tan » souffrait déjà depuis quelques années d'un trouble non-diagnostiqué au niveau de la production du langage. Il était incapable de produire des phrases complètes et ne s'exprimait que par le biais du son « tan » - d'où le nom qui lui a été donné - et, occasionnellement, d'un seul juron. À l'autopsie, Broca a constaté différentes lésions, dont une primaire au niveau du lobe préfrontal gauche, ce qui l'a amené à défendre le rôle essentiel de cette région dans la production du langage.

¹⁰ Il est également possible de pratiquer l'inactivation temporaire de régions cérébrales spécifiques chez certaines espèces animales pour reproduire à courte échelle l'effet d'une lésion cérébrale permanente, bien que, comme le soulignent Vaidya et al. (2019), l'efficacité de cette procédure demeure restreinte. L'inactivation temporaire ne permet pas, en effet, de saisir les changements d'organisation cérébrale (ex. récupération fonctionnelle liée à la plasticité cérébrale) suivant la phase aiguë de la lésion.

caractéristiques particulières du cas qui enfreint le modèle (est-ce que le patient a d'autres lésions diffuses? a-t-il un profil neurologique atypique par rapport aux cas précédents?) ou est-ce plutôt le modèle lui-même qui est fautif (peut-être manque-t-il quelques variables d'intérêt ou qu'il a un champ d'application trop restreint ou trop vaste)?

Ces déviations par rapport au modèle aident à raffiner et à créer de nouveaux modèles qui, en retour, aideront à catégoriser et à guider le traitement de nouveaux cas; cas qui pourront, à leur tour, mettre au défi le ou les modèles existants. C'est, comme l'indique Medina & Fischer-Baum (2019), l'un des avantages issus de l'étude de cas unique de lésions cérébrales par rapport aux preuves basées sur la neuroimagerie (ce qu'ils appellent le « *big data* »): la possibilité de confronter les modèles en vigueur sur les effets d'une lésion donnée et de produire de nouveaux modèles neurocognitifs compatibles avec les écarts observés avec le modèle initial. Il faut préciser néanmoins que l'articulation de ce dialogue s'échafaude sur une taxonomie de la vie mentale, c'est-à-dire sur une collection de fonctions cognitives prédéfinies (par exemple la mémoire épisodique, l'attention sélective, la prise de décision, etc.) qu'on cherche à confiner dans des régions cérébrales, et que la précision avec laquelle on réussit effectivement à délimiter ces fonctions cognitives dans le cerveau dépend de notre capacité à bien les définir conceptuellement. Ce point fera l'objet de la section **1.2.3**.

Les preuves basées sur les études lésionnelles possèdent également d'autres avantages. Elles permettent de différencier les composantes nécessaires à la réalisation d'une fonction cognitive ou d'un comportement (ex. celles qui lorsque désactivées ou éliminées perturbent l'habileté observé) des composantes qui ne font que participer à celle-ci; ce que, encore une fois, les techniques de neuroimagerie ne peuvent pas faire (Vaidya et al. 2019). Autrement dit, elles permettent de mettre en relief la nécessité causale de certaines structures (Adolphs, 2016). Dans la même veine, les études lésionnelles contribuent à nous donner un aperçu de la manière que les réseaux neuronaux s'organisent (ex. spécialisation fonctionnelle de certaines régions, assemblées neuronales, connectivité structurelle, etc.) et nous ouvre la voie sur l'étude de certains processus, comme celui la plasticité neuronale (Lakes-Harlan, 2013; Feinstein, 2013).

En revanche, si l'étude des lésions permet d'identifier des aires cérébrales nécessaires, elle ne permet pas de recenser les régions suffisantes à l'émergence d'un processus cognitif particulier. Même si nous prouvons qu'une lésion du lobe préfrontal gauche mène systématiquement à un déficit de la production orale du langage, cela ne signifie pas pour autant que cette structure est

le siège unique ou le réseau complet pour cette fonction. De ce fait, nous savons qu'une panoplie d'autres structures participent à cette fonction: faisceau arqué, aire de Wernicke (Shuren et al. 1995; Binder, 2015; Salvan et al. 2017), etc. En gros, l'étude des lésions ne peut jamais garantir que nous connaissons réellement le réseau complet des composantes qui participent à l'émergence de la fonction cognitive étudiée. Ce type de preuve ne peut, au mieux, que révéler des composantes individuelles nécessaires, mais ne peut pas – ou que très difficilement - fournir un aperçu d'ensemble des éléments impliqués dans la production de la fonction cognitive à l'étude.

Pire encore, il se peut que les régions neuronales identifiées dans l'étude d'un cas clinique ne soient même pas responsables computationnellement de la fonction cognitive affectée (Kwok, 2013). En effet, les lésions cérébrales focales n'affectent pas que les neurones où elles se trouvent, mais aussi ce que D'Esposito (2000) nomme les fibres de passage qui relient certaines régions cérébrales entre elles. Ainsi, il est possible qu'une région A soit lésée dans le profil neurologique d'un patient, mais qu'elle ne soit pas en cause dans les perturbations cognitives observées en raison de ses connexions avec une région B ou C qui, elle, jouerait un rôle dans la régulation de ces fonctions cognitives. Ça serait alors non pas les neurones locaux lésés de la région A qui entraîneraient les dysfonctions spécifiques observées dans le profil cognitif du patient, mais les fibres lésées des régions B ou C qui traversent la région A.

Enfin, aucun cas clinique n'est strictement identique à un autre (Adolphs, 2016). Nous pouvons emprunter des heuristiques et comparer les cas entre eux sur la base de certains critères, comme l'étendue, la forme ou, encore, la distribution des lésions (Kwok, 2013), mais on ne peut jamais retrouver exactement un même profil neurologique. Cette constatation va de pair avec un autre problème que soulève Adolphs, soit que la quantité de données qu'on peut obtenir à partir de l'étude des lésions est très limitée. Non seulement est-il difficile d'obtenir des données similaires en raison de la nature idiosyncrasique des lésions, mais il est impossible éthiquement de contrôler expérimentalement l'expression de ces lésions chez l'être humain. Un défi, pour les études lésionnelles, est donc d'assembler un corpus suffisamment large de données pour tirer des inférences valides sur les relations anatomofonctionnelles entre régions cérébrales et fonctions cognitives (Adolphs, 2016).

1.1.2. Les preuves basées sur les techniques de neuroimagerie

Historiquement plus récentes que l'étude des lésions cérébrales, les techniques de neuroimagerie consistent en une série de méthodes qui ont comme objectif de visualiser, de mesurer et de cartographier approximativement les régions cérébrales impliquées – entre autres choses – dans les facultés cognitives. Couvrir adéquatement l'ensemble de ces méthodes serait une tâche colossale qui nécessiterait la rédaction de plusieurs chapitres (au minimum), c'est pourquoi nous allons nous contenter de présenter quelques-unes d'entre elles sommairement afin de donner une idée aux lecteurs de leurs implications potentielles pour l'étude de la cognition, avec une attention particulière sur l'imagerie par résonance magnétique fonctionnelle (fMRI). La fMRI occupera une place de prédilection dans cette analyse, car, non seulement est-elle actuellement la méthode de cueillette de données la plus populaire en neurosciences (Ganis & Kosslyn, 2002; Poldrack et al. 2011), mais elle permet également d'identifier des activités dynamiques transitoires dans le cerveau (Huettel, 2012), un aspect qui deviendra important dans notre discussion de la cognition comme système dynamique dans le chapitre 3 de ce mémoire.

L'imagerie par résonance magnétique fonctionnelle (fMRI) repose sur le principe que l'activation d'une région cérébrale donnée est accompagnée par une augmentation locale de la circulation sanguine dans cette région pour alimenter les neurones en oxygène (Glover, 2011). Cet approvisionnement en oxygène est, cependant, nettement supérieur aux besoins métaboliques et énergétiques des neurones, créant un surplus d'oxygène dans le sang. Un programme associé à la fMRI codifiera ensuite les variations locales en oxygène (le signal BOLD) à l'aide d'un spectre de couleur, habituellement du bleu jusqu'au rouge (Russell et al. 2011), reflétant le niveau d'activation neuronale dans certaines régions du cerveau selon la logique que plus une région est sollicitée, plus il y aura d'oxygène présent dans le sang. Lorsque le signal, dans une région, est codé dans la palette du bleu-violet, cela signifie qu'il n'y a peu ou presque pas d'activité à l'intérieur de celle-ci. Une couleur se situant dans le spectre du orange-rouge signifie, en revanche, que la région est particulièrement sollicitée pour la tâche en cours.

La fMRI est réputée pour offrir une très bonne résolution spatiale du cerveau (Glover, 2011). Lorsqu'elle est réalisée avec précaution, elle peut capturer avec précision la plupart des

subdivisions communes aux grands atlas neuroanatomiques (l'atlas de Talairach, l'atlas de Duvernoy, l'atlas probabiliste LONI, etc.) faisant d'elle une méthode de prédilection pour cartographier les relations structure-fonctions entre le cerveau et la cognition (Russell et al. 2011; Turner, 2016). Qui plus est, elle a également l'avantage de révéler des activités distribuées dans le cerveau, contrairement à l'étude des lésions qui s'attarde principalement aux conséquences de l'inactivité de certaines régions précises. Cette méthode nous permet donc plus aisément de postuler l'existence de mécanismes communs ou distincts à certaines activités cognitives, comme lorsque des régions similaires sont systématiquement activées – qu'elles covarient – lors de différentes tâches (Mather et al. 2013). Non seulement la fMRI peut-elle identifier des mécanismes redondants, mais elle peut aussi être utilisée, comme il a été dit, pour étudier les transitions dynamiques au sein de l'activité cognitive.

La fMRI possède cependant son lot de problèmes. D'abord, parce qu'elle nécessite justement beaucoup de précautions afin de fournir des données fiables. Une grande quantité de correctifs doivent être apportés aux données brutes qui sont obtenues via le signal BOLD (ex. correction pour les mouvements de la tête, pour les *drifts*¹¹ et pour les autres bruits parasites de toutes sortes qui peuvent fausser le signal), et il s'agit d'autant d'occasions où il est possible d'ajouter, de perdre ou de modifier de l'information (Russell et al. 2011; De Blasi et al. 2020). Par conséquent, il ne faut pas croire que les données obtenues via la fMRI reflètent l'activité *réelle* du cerveau, telle qu'elle se produit. Elles représentent plutôt une approximation de celle-ci qui est sujette à différentes erreurs de manipulation (ex. les correctifs) et d'interprétation.

Ensuite, la fMRI n'identifie pas les régions strictement nécessaires à certaines fonctions cognitives (Bunge & Kahn, 2009; Medina & Fischer-Baum, 2019). Il ne faut pas se méprendre, ce que la fMRI capture n'est pas spécifiquement le réseau cérébral à l'origine de l'activité cognitive étudiée, mais un portrait d'ensemble de l'activité cérébrale « totale » au moment du scan, duquel il faut déduire les activations propres à ladite activité cognitive. Cette tâche n'est évidemment pas facile, surtout lorsqu'on mêle à cela les différentes sources de variations dans le signal BOLD (celles mentionnées ci-dessus). Il existe, néanmoins, des méthodes pour

¹¹ Les *drifts*, ou dérapages, sont un phénomène mystérieux en fMRI où des variations du signal se produisent spontanément à très faible intensité. Certains croient que ces dérapages seraient imputables à une instabilité du scanner ou à des facteurs endogènes propres aux patients, mais les raisons exactes derrière ce phénomène demeurent très largement inconnues (Smith et al., 1999; Zahra et al., 1997; Aguirre et al., 1997 dans Russell et al. 2011).

éliminer ces variations, dont celles qui consistent à soustraire les activations neuronales propres à l'état de repos (Chen et al. 2012), mais elles ne sont pas infaillibles.

Un autre problème est que, bien que la fMRI puisse capturer des transitions dynamiques dans l'activité cérébrale, celle-ci possède une très faible résolution temporelle (Kim et al. 1997; Glover, 2011), ce qui va, en un sens, à l'encontre de cet objectif. Comment savoir, en effet, si la carte de l'activité cérébrale prise au temps T2 est une transition directe de celle obtenue au temps T1 ou si elle n'est pas plutôt la résultante d'une série de modifications perdues dans l'intervalle temporelle entre T1 et T2? Cette question est importante, car si des modifications ont lieu durant cet intervalle, il faudrait réviser les relations de causalité établies entre certains états cérébraux via la fMRI. Peut-être qu'en réalité l'état cérébral x observé au T2 n'est pas causé – ou précédé – par une configuration d'activité cérébrale y , comme le laisse entendre la carte prise au T1, mais par une ou plusieurs activités cérébrales spatialement et fonctionnellement hétérogènes.

D'autres méthodes de neuroimagerie incluent la tomographie par émission de positrons (le PET-scan), l'imagerie par résonance magnétique structurale (sMRI), l'IRM de diffusion (DTI), la tomodensitométrie, telle que la TACO (pour plus d'informations, voir Bunge & Kahn, 2009; Baars & Cage, 2010; Krishnan, 2012), etc. Pour les besoins de notre projet, nous allons nous contenter de discuter des trois premières seulement. Nous jugeons que cela sera suffisant pour illustrer la complémentarité de certaines méthodes en neuroimagerie, et comment il est possible de combiner celles-ci pour obtenir un portrait plus global des relations structure-fonction entre cerveau et cognition.

Commençons par le PET-scan. Méthode moins récente que la fMRI, le PET-scan utilise des marqueurs radioactifs qui se rattachent à des molécules spécifiques pour mesurer différents aspects du cerveau comme le métabolisme du glucose, la circulation sanguine, etc (Berger, 2003; Kropotov, 2016). Le fait que ces marqueurs radioactifs puissent être liés de la sorte à des molécules comme le glucose ou les molécules d'eau présentes dans le sang confère au PET-scan une certaine flexibilité. Il est possible, en effet, d'avoir différentes mesures de l'activité cérébrale selon les besoins, et même de visualiser parallèlement l'activité de ces différentes molécules pour avoir un profil plus détaillé des opérations qui ont lieu dans le cerveau. Toutefois, il présente des risques non-négligeables pour la santé des individus (Fernández et al. 2019), et, donc, il ne peut pas être utilisé dans des contextes expérimentaux usuels (ex. avec des

sujets sains issus de populations normales). Aussi, il présente une très mauvaise résolution spatiale (Moses, 2011 fait le tour de la question), ce qui fait qu'il est difficile d'établir des corrélats anatomo-fonctionnels précis entre certaines régions cérébrales et des fonctions cognitives via cette technique de neuroimagerie.

L'imagerie par résonance magnétique structurelle (sMRI), de son côté, utilise l'eau (spécifiquement les atomes d'hydrogène) contenue dans les tissus cérébraux pour créer des images à haute résolution du cerveau (Di Prospero et al. 2022). Plus précisément, cette technique permet, via les différences de quantité d'eau, de quantifier le volume de matière grise et de matière blanche présents pour certaines régions du cerveau et permet également de visualiser les variations de gyrification cérébrale, c'est-à-dire les circonvolutions sur la surface du cortex (Gifford et al. 2020). Vu les aspects qu'elle capture, la sMRI est utilisée dans des contextes variés pour étudier les différences de volumes cérébraux entre les sexes (Bhargava et al. 2021), les effets des lésions traumatiques (Yuh, 2017) ou encore pour évaluer la densité de la matière grise chez des personnes autistes (Stigler & McDougle, 2013), entre autres choses.

Bien qu'elle soit utile pour localiser des lésions ou des anomalies cérébrales à l'échelle macroscopique, la sMRI possède son lot de problèmes. L'un d'entre eux est, on pourrait dire, de nature philosophique: quel genre de conclusions peut-on tirer sur la cognition à partir de caractéristiques structurelles, comme le volume et la gyrification cérébrale? Un présupposé qui est souvent endossé dans les études qui emploient la sMRI est que le volume d'une région cérébrale est corrélé à sa performance (on le retrouve dans une myriade d'études, dont Ruben et al. 1999; Mortimer et al. 2012; Menary et al. 2013; Ellfolk et al. 2014; Aljondi et al. 2019). Or, plusieurs chercheurs s'entendent pour dire que ce n'est pas tant le volume cérébral qui compte pour comprendre les différences de performance cognitive, mais plutôt le degré de connectivité intra- et inter-régional à l'intérieur du cerveau (Wendelken et al. 2011; Hofman, 2014¹²; Sotiras et al. 2017), bien que les deux soient corrélés¹³. Par conséquent, on peut se

¹² Le titre de l'article d'Hofman (2014) semble suggérer une prise de position en faveur de l'importance du volume cérébral, position inverse à celle défendue ici. Toutefois, le propos d'Hofman est que, certes, un plus gros cerveau serait mieux, mais seulement en vertu du fait qu'une plus grande quantité de matière grise permettrait théoriquement un plus grand nombre de connexions - vu qu'il y aurait plus de neurones.

¹³ En effet, il est possible qu'un cerveau plus volumineux soit corrélé à une connectivité neuronale plus grande (Hänggi et al. 2014), selon la logique qu'une plus grande quantité de corps cellulaire permettrait un plus grand nombre de connexions. Toutefois, il arrive bien souvent qu'une région cérébrale s'atrophie en réponse à une spécialisation fonctionnelle (Mürner-Lavanchy et al. 2014; Hoekzema et al. 2016; Sotiras et al. 2017). Par ailleurs, il n'est pas rare, comme le soulève Hoffman (2012, p.51), que d'autres régions du cerveau compensent pour une activité en cours, de sorte que deux individus ayant une performance cognitive qualitativement similaire pourraient

demander si les informations obtenues via la sMRI sont réellement pertinentes en ce qui concerne l'étude de la cognition.

Pour finir, l'IRM de diffusion (DTI) utilise le taux de diffusion des molécules d'eau – le mouvement Brownien – dans les tissus pour créer une image tridimensionnelle colorée des projections d'une région du cerveau à d'autres (Razenberger & Snyder, 2021). L'aspect intéressant de cette méthode, par rapport à celles nommées jusqu'à présent, est sa capacité à capturer la directionnalité et l'anisotropie des faisceaux de neurones, ce qui permet concrètement de voir la progression de l'activité neuronale à partir d'un voxel et d'identifier des trajectoires neuronales dans le cerveau (Luo et al. 2019; Kieronska & Sloniewski, 2019). Des deux méthodes mentionnées précédemment, il est de notre avis que la DTI est celle qui apporte la contribution la plus pertinente à l'étude de la cognition, parce qu'elle capture la *connectivité* cérébrale (voir Skudlarski et al. 2008; Tymofiyeva et al. 2017), la façon dont certaines régions sont liées structurellement et susceptibles de s'influencer fonctionnellement. C'est un point qui prendra de l'importance lorsqu'on proposera de modéliser la cognition, dans le chapitre 3, comme un système réseautique dynamique.

La DTI est, néanmoins, sujette à d'importants problèmes en ce qui concerne l'étude de la cognition. Le plus gros de ces problèmes est d'évaluer la pertinence fonctionnelle des faisceaux identifiés (Kieronska & Sloniewski, 2019). En effet, la DTI ne nous informe pas sur la contribution respective des structures neuronales recensées aux processus cognitifs; elle ne communique de l'information que sur les éléments structurels des réseaux cérébraux, comme la directionnalité des fibres, leur anisotropie ou, encore, la myélinisation des axones (Lebel & Deoni, 2018). Une autre limitation importante de la DTI est que les voxels ne peuvent pas représenter plus d'une direction à la fois pour les fibres de grosses tailles (O'Donnell & Westin, 2011), ce qui signifie qu'il est possible, théoriquement, lorsque des faisceaux se croisent, que ce que nous présumons être la continuité du faisceau A soit plutôt la suite du faisceau B et vice-versa. La DTI est également sujette, similairement à la fMRI, à des artefacts importants lors de la capture des données. Elle est, par exemple, très sensible au mouvement, assez pour dire que les données peuvent être affectées par la respiration et les battements cardiaques (Soares et al. 2013).

avoir des volumes cérébraux différents pour les régions associées à cette performance. Compte tenu de ces phénomènes, des caractéristiques structurelles comme le volume du cortex ne devraient pas être utilisées comme marqueur de la capacité computationnelle du cerveau.

Malgré leurs défauts, il est de notre avis qu'une combinaison judicieuse de ces méthodes – et d'autres – pourrait répondre aux questions importantes des sciences cognitives à savoir les constituants – le quoi – et les processus – le comment – responsables de l'émergence et la régulation de l'activité cognitive. Comme le dit Mather et al. (2013), la fMRI n'a pas à s'arrêter à des questions triviales sur le localisationnisme - sur le où -, mais peut être utilisée, comme il a été mentionné, pour capturer des mécanismes redondants; ce avec quoi nous sommes d'accord. La fMRI peut, par ailleurs, être utilisée pour modéliser la connectivité entre les régions cérébrales (Russell et al. 2011), nous offrant un aperçu des patrons d'interactions transitoires entre certaines portions du cerveau. Elle peut donc être utilisée, via des techniques de modélisation, pour capturer des exemples de réseaux fonctionnels comme les *transiently assembled local neural subsystems* (TALoNS) dont parle Anderson (2014). Cet élément reviendra au chapitre 3.

Toujours dans cette optique, nous croyons que la fMRI devrait être utilisée en complémentarité avec d'autres approches pour pallier ses défauts. L'étude des lésions peut, par exemple, fournir un corpus de référence quant aux régions strictement nécessaires à une activité cognitive duquel il serait possible, ensuite, de déduire les activités plus superficielles qu'on observe sur les images du fMRI. Le PET-scan peut, de son côté, donner un aperçu varié des activités métaboliques qui ont lieu durant une activité cognitive, et potentiellement nous informer sur leurs contributions respectives. La sMRI, elle, peut compléter la capture donnée par la fMRI en relevant certaines caractéristiques macro-structurelles du cerveau; à condition de bien doser les interprétations tirées à partir du volume et de la gyrification cérébrale. Enfin, la DTI peut montrer concrètement les projections de certaines régions vers d'autres, nous permettant ainsi de recréer plus finement les réseaux fonctionnels du cerveau via la fMRI. Ces deux dernières méthodes sont d'ailleurs souvent combinées, comme l'une capture la dimension structurelle du cerveau et l'autre la dimension fonctionnelle (Sporns et al. 2005; Staempfli et al. 2008; Vassal et al. 2016). D'autres méthodes, comme l'électroencéphalogramme, dont nous allons parler dans les prochaines lignes, pourraient également être couplées à la fMRI pour résoudre le problème de résolution temporelle (Menon et al. 1997).

Ainsi, il est de notre avis qu'en combinant les avantages respectifs des différentes méthodes de neuroimagerie susmentionnées – et des autres méthodes existantes – il serait possible de produire des cartes plus complètes et moins biaisées des structures et des processus qui ont lieu dans le cerveau; cartes qui pourraient être plus informatives à l'égard du ou des mécanismes

qui permettent l'émergence et la régulation de l'activité cognitive. Jumelé à différentes mesures physiologiques, comportementales, interpersonnelles et écologiques, nous croyons que cette combinaison judicieuse des techniques de neuroimagerie est la première étape vers l'obtention d'un portrait d'ensemble – le réseau – des composantes et processus qui mènent à l'activité cognitive, un portrait qui ne serait pas confiné au cerveau.

1.1.3. Les preuves basées sur les études d'enregistrement neuronal

Enfin, nous pouvons discuter brièvement de la dernière catégorie de preuves, les preuves issues de l'enregistrement neuronal. Plutôt que de fournir une image de l'activité cérébrale, comme le feraient les techniques de neuroimagerie, les techniques d'enregistrement neuronal fournissent un signal sous la forme d'un tracé qui rend compte des oscillations - du rythme - dans l'activité neuronale à l'intérieur de différentes régions du cerveau (Tatum, 2014)¹⁴. Il existe plusieurs techniques d'enregistrement neuronal, dont le maintenant classique électroencéphalogramme (EEG), le magnétoencéphalogramme (MEG), électrocorticogramme (ECoG), l'électrooculogramme (EOG), etc. Nous jugeons que, bien que ces méthodes soient très utiles à l'étude du cerveau, il n'est pas pertinent de présenter chacune d'entre elles, un à un, comme l'EEG et le MEG, en particulier, partagent plusieurs similarités méthodologiques (da Silva, 2013). Nous allons plutôt nous contenter de survoler la contribution de ce type de preuve à l'étude de la cognition et des problématiques qui l'accompagnent.

Il est pertinent de parler des preuves basées sur l'enregistrement neuronal, car elles fournissent un regard particulier sur des événements cérébraux et cognitifs que les autres catégories de preuves ne nous donnent pas. D'abord, parce que les techniques d'enregistrement neuronal capturent directement l'activité des neurones (Bunge & Kahn, 2009), mais, aussi, parce qu'elles le font en temps réel (Lehnertz et al. 2021). En effet, ces techniques n'ont pas à utiliser un proxy pour évaluer l'activité neuronale comme le font des techniques de neuroimagerie telles que la fMRI avec le signal BOLD ou la DTI avec le mouvement Brownien. Elles peuvent mesurer

¹⁴ Il est à noter que l'utilisation du mot « région », dans ce contexte-ci, n'implique pas que les techniques d'enregistrement neuronal doivent absolument se limiter à la capture d'une région cérébrale prédéfinie telle qu'on le retrouverait dans un atlas neuroanatomique. Au contraire, des techniques comme l'EEG peuvent capturer l'activité de portions aussi petites que quelques centaines de neurones (Stevenson & Kording, 2011).

directement, et de manière soutenue, les variations d'activité neuronale, ce qui permet, à notre avis, un meilleur *mapping* des relations entre les activités cérébrales et cognitives.

Parce qu'elles capturent directement et continuellement les oscillations dans l'activité neuronale, ces techniques sont bien outillées pour nous informer sur une panoplie de phénomènes, dont certaines pathologies du système nerveux, comme l'épilepsie (Noachtar & Rémi, 2009) ou le Parkinson (Soikkeli, 1991), les états de conscience (Noirhomme et Laureys, 2014; Yoo et al. 2014) et les stades du sommeil (Diykh et al. 2016). Plus intéressant pour notre propos est la possibilité pour l'EEG intracrânien de révéler, via les différents rythmes cérébraux, des mécanismes neuronaux impliqués dans des phénomènes cognitifs comme l'encodage et la récupération de l'information en mémoire épisodique (Johnson et al. 2020). En effet, l'EEG met en lumière des phases de synchronisation et de désynchronisation des rythmes thêta-gamma (Hanslmayr et al. 2016) entre différentes régions cérébrales (cortex préfrontal, hippocampe, etc.) qui pourraient rendre compte du processus par lequel nous formons des souvenirs (Fell & Axmacher, 2011).

Hélas, même si des techniques d'enregistrement neuronal comme l'EEG peuvent être employées à des fins de capture dynamique de l'activité cérébrale (comme c'est le cas avec les études qui portent sur les phénomènes de synchronisation et de désynchronisation), cette catégorie de preuve n'échappe pas à l'emprise du localisationnisme cérébral. Bien des études utilisent ces techniques dans l'espoir de localiser certains processus cognitifs. Opitz et al. (2002) montre, par exemple, que le traitement de la déviance auditive, c'est-à-dire la réaction neuronale face à la transgression d'une règle concernant la présentation de l'information dans la modalité auditive, est assumée en grande partie par le gyrus temporal supérieur et le gyrus frontal inférieur. Mainy et al. (2007), de son côté, fait l'inventaire des régions cérébrales impliquées dans la rétention et la consolidation d'une séquence de lettres en mémoire de travail. Enfin, Zhang et al. (2018) ont employé un combiné d'EEG et de ECoG pour cartographier les régions participant à la récupération de l'information en mémoire de travail.

Évidemment, notre but ici n'est pas de faire un procès d'intention sur ces chercheurs. Il existe, selon nous, une ligne très fine, mais forte d'implications, entre l'acte de *localiser* une fonction cognitive et l'acte d'*identifier* une région cérébrale contribuant à cette fonction; parce que la localisation passe par l'identification de régions fonctionnellement et/ou structurellement différentes et que l'identification d'un phénomène repose en bonne partie sur les aspects

spatiaux – ou temporels – de celui-ci. Toutefois, et à notre avis, l’acte d’identifier un objet n’est pas accompagné du même présupposé réductionniste que le localisationnisme quant à la relation entre l’objet et ses composantes. Nous pouvons identifier les différentes aires cérébrales, une à une, *ad nauseam*, sans pour autant réduire épistémologiquement une fonction à celles-ci (identifier ne veut pas dire qu’on a tous les éléments nécessaires du système). Or, lorsque nous localisons une chose dans le monde, que ce soit dans un seul lieu ou plusieurs à la fois, il semblerait que nous portons un jugement sur la totalité (le nombre) et la pondération des éléments qui composent un objet (si la mémoire de travail est localisée dans le cortex préfrontal dorsolatéral et l’hippocampe, elle doit être entièrement prise en charge par ces structures).

Pour des raisons que nous allons expliciter ci-dessous, nous croyons que l’acte de localiser est inadéquat, voire impraticable en ce qui concerne l’étude de la cognition. Cela est particulièrement vrai avec des méthodes, comme l’EEG, qui peinent à identifier adéquatement les sources exactes des signaux reçus (Dürschmid et al. 2016). Bien souvent, il arrive même que les signaux prélevés à un endroit sur le scalp soient la résultante de l’activité électrique d’amas de neurones éloignés dans le cerveau (Bell & Cuevas, 2012). Nous allons revenir sur la question du localisationnisme dans la section qui suit, mais, le point important, en ce qui concerne ce type de preuve, est que les méthodes d’enregistrement neuronal ne devraient pas s’attarder ou se limiter à une description spatiale d’une fonction lorsqu’elles peuvent nous fournir une myriade d’indices électrophysiologiques pertinentes à la compréhension de la cognition (ex. oscillations des ondes, synchronisation du rythme, *spikes*, etc.).

1.2. Les problèmes avec le localisationnisme en sciences cognitives

Nous l’avons vu dans les dernières pages, plusieurs des preuves mentionnées précédemment en faveur de la réduction ontologique des états mentaux aux états cérébraux reposent sur l’objectif de révéler le lieu d’une activité cognitive: étude des lésions cérébrales, fMRI, sMRI et même des techniques comme l’EEG. Or, malgré le poids cumulé de ces preuves, nous sommes en droit de questionner la pertinence explicative, voire même la praticabilité, de la localisation comme démarche d’accroissement des connaissances en sciences cognitives. Cette section adressera sélectivement quelques-uns des problèmes que nous avons identifiés au sujet du localisationnisme.

1.2.1. Le problème définitionnel: qu'est-ce qui porte la marque du cognitif?

Un problème flagrant avec le localisationnisme est celui de déterminer ce que nous essayons de localiser au juste (Uttal, 2001). Nous savons que ce quelque chose porte le nom de faculté (fonction, capacité, etc.) cognitive ou, de manière plus large, de cognition, mais nous sommes encore loin d'avoir atteint un consensus sur le sens à accorder à ces termes. Est-ce que la cognition se résume simplement à un traitement de l'information effectué sur des symboles et les fonctions cognitives à des subdivisions de cette capacité plus générale? Ou est-ce que la cognition ne serait pas plutôt un phénomène émergent de l'organisation biologique particulière de certains organismes et de leurs interactions spécifiques avec un environnement donné, le tout n'ayant absolument rien à voir avec la manipulation de symboles, comme le défendent certains partisans de l'énaclivisme radical (voir Hutto & Myin, 2013)?

Une autre façon de poser la question est de se demander comment on en vient à conclure qu'un phénomène particulier dans le monde porte la marque du cognitif. En d'autres mots, quel est le critère qui nous dit quand un processus peut être considéré comme ayant le statut de cognitif (Rowlands, 2010)? Il existe différentes définitions de ce qu'est la cognition et chacune d'entre elles est lourde de répercussions à l'égard de ce que nous pouvons considérer comme étant de l'ordre du cognitif (Adams et Aizawa discutent en long et en large de ces questions dans leur ouvrage de 2010).

Par exemple, si nous décidons de définir simplement la cognition en termes de traitement de l'information, comme c'est souvent le cas dans la littérature (pour des exemples, voir: Robinson et al. 2013; Clark et al. 2015; Sayood, 2018; Gordon & Ramani, 2021), il s'ensuit que les objets capables de traiter de l'information pourraient être considérés comme faisant preuve de cognition. Une calculatrice, dans cette perspective, ferait une opération cognitive lorsqu'elle donne la réponse à une équation. De même, un logiciel comme Antidote montrerait de la cognition lorsqu'il propose, en réponse à des *inputs* linguistiques inadéquats, des corrections. Un organe comme le cœur, également, ferait de la cognition lorsqu'il adapte son rythme à l'information entrante.

Nous voyons tout de suite pourquoi une définition comme celle du traitement de l'information est problématique pour capturer ce en quoi consiste vraiment la cognition; elle serait trop

libérale. Certes, on peut dire que la capacité de traiter et de manipuler de l'information est un bon point de départ pour démarquer la cognition. Ce critère serait assez large pour accommoder certaines intuitions que nous avons au sujet de différentes formes d'intelligence, dont la nôtre, celle des animaux, et celle que posséderait, par exemple, les martiens de l'expérience de pensée de Clark (2010). Mais, formulée de manière aussi large, cette définition est également en conflit avec d'autres intuitions que nous avons à propos de ce qu'on pourrait considérer comme des agents cognitifs (ex. calculatrice, cœur, etc.). Comme le soulignent Adams et Garrison (2013), le comportement des agents cognitifs ne se résume pas qu'à un traitement de l'information sensoriel, comme on pourrait l'observer chez une bactérie ou dans le cas d'un organe comme le cœur, mais en un traitement complexe qui porte sur des contenus représentationnels; des croyances, des raisons et des objectifs intrinsèques à l'agent qui organisent ses comportements dans l'environnement.

Néanmoins, accepter l'idée que la vie mentale ne serait qu'une question uniquement de traitement de l'information – au sens large – influence considérablement à la fois l'intention derrière la recherche d'un *locus* pour la cognition et le lieu actuel trouvé pour celle-ci. Comme il serait possible de trouver des exemples de traitement de l'information dans une myriade d'objets variés, allant, comme nous l'avons mentionné, de la calculatrice au cerveau humain. La recherche d'une localisation précise à la cognition ne ferait plus de sens (elle serait partout!).

Adams et Aizawa (2010) résumant d'ailleurs très bien cette idée lorsqu'ils disent:

[I]f one wants to find cognition in new and unexpected places, such as the body and the physical and chemical environment, it turns out to be convenient to have easily satisfied conditions in a theory of the cognitive. The theory that any sort of information processing constitutes cognitive processing is just such a theory. If just any sort of information processing is cognitive processing, then it is not hard to find cognitive processing in notebooks, computers, and other tools (p.11).

En revanche, si on définit plus spécifiquement la cognition comme un type de traitement de l'information exclusif à l'organisation intrinsèque du cerveau, comme un phénomène qui émerge des propriétés uniques des cellules neuronales, des types de connexions entre elles et de leurs regroupements en des communautés structurellement et fonctionnellement différentes, l'*intention* de confiner la cognition à l'intérieur du cerveau ferait du sens. La cognition serait, dans cette perspective, intrinsèquement liée – réductible – aux états cérébraux. Or, le problème

se situe précisément à ce niveau-là. Il n'y pas de raisons *a priori* pour accepter une telle vision de la cognition.

Certes, comme nous venons de le voir avec les types de preuves en neurosciences, il existe un corpus de recherches mastodontes sur les corrélations anatomofonctionnelles entre des régions spécifiques du cerveau et des facultés cognitives. Cependant, de la même façon que le géocentrisme ou la théorie ventriculaire ont persisté pendant des siècles, ce n'est pas parce qu'une chose est vastement défendue sur une longue période de temps, qu'elle est pour autant juste. Avec la contribution constitutive de structures extra-crâniennes (que nous allons voir dans le prochain chapitre), nous avons de bonnes raisons de croire qu'il existe des problèmes avec une définition neurocentrée de la cognition.

Il est intéressant aussi d'évoquer du même coup la définition qu'Adams et Aizawa (2010) proposent de la cognition, comme ils défendent justement une position neurocentrée de la cognition. Ces derniers ne ménagent pas leurs critiques à l'égard des partisans de la cognition étendue tels que Clark et Chalmers, les accusant même de fournir intentionnellement des définitions vagues de la cognition qui accommoderaient leur position théorique. La citation ci-dessus fait partie justement de cette critique.

Pourtant, eux-mêmes proposent une définition floue et problématique de ce en quoi consisterait un phénomène cognitif authentique. Suivant la lancée de Searle (1980), ces derniers soutiennent que c'est la présence et la manipulation de représentations non-dérivées qui permettrait de distinguer un phénomène non-cognitif d'un phénomène cognitif authentique. L'idée derrière cette proposition est que les contenus présents dans la cognition telle qu'on la retrouve naturellement ont une intentionnalité intrinsèque « originale » qui ne dépendrait pas du « contrôle » ou de « l'interprétation » d'un agent extérieur. Pour reprendre les termes d'Adams et Aizawa (2010), ces contenus « *do not require independent or prior existence of other content, representations, or intentional agents* » (p.32).

Or, il n'est pas tout à fait clair sous quelles conditions une représentation pourrait être considérée comme non-dérivée (Kaplan, 2012). Prenons, comme exemple, ce que dit Haugeland (cité dans Dennett, 1987, p.289) à ce sujet: « *[Our artifacts] only have meaning because we give it to them. [...] computers themselves don't mean anything by their tokens (anymore than books do) - they only mean what we say they do* ». Ne peut-on pas dire quelque

chose de similaire sur notre manière d’acquérir une sémantique du monde? Que ces choses, ce que nous considérons comme nos représentations du monde, ont un sens simplement en vertu du fait qu’on nous a transmis, dans une perspective développementale, un réseau de significations à leur attribuer? Que ces choses indépendamment d’un *input* extérieur ne signifierait rien pour nous? Comme en conclut Kaplan, nous sommes loin d’avoir tranché la question des contenus originaux.

Qui plus est, il n’y a pas, comme le soulève Shapiro (2009) de consensus sur les conditions particulières qui donneraient naissance aux représentations non-dérivées à l’intérieur des systèmes neuronaux. Adams et Aizawa admettent eux-mêmes dans un ouvrage précédent, que « *it remains unclear just exactly what naturalistic conditions give rise to non-derived content; hence it remains correspondingly unclear just exactly what objects bear non-derived content* » (2008, p.55). Cet aveu est problématique pour leur argument. En effet, si nous n’avons aucune idée de la façon dont les représentations non-dérivées émergeraient de systèmes naturels comme le cerveau, il apparaît injustifié de les réduire *a priori* à cet organe. Il semble plus avantageux épistémologiquement d’investiguer le rôle combiné de différentes structures afin d’avoir un aperçu plus complet du jeu d’interactions qu’il pourrait y avoir entre celles-ci, plutôt que de réduire d’emblée ce phénomène au cerveau.

Pour ces raisons, nous sommes d’avis pour dire avec Kaplan que cette définition est précaire et insatisfaisante pour une définition qui a la prétention d’établir ce qu’est de l’ordre du cognitif. Une définition satisfaisante *a posteriori* de la cognition doit composer, à notre avis, avec la présence de différentes facultés générales considérées essentielles à celle-ci (ex. la sensation, l’encodage, la rétention, la récupération, la planification, etc.) – sans tomber dans des problèmes de granularité (nous y reviendrons). Nous pouvons nous inspirer des connaissances que nous avons sur les quelques facultés que nous avons répertoriées en observant des agents se comporter dans la nature pour construire une définition opérationnelle temporaire de la cognition qui pourrait être confrontée philosophiquement et scientifiquement par la découverte de nouveaux mécanismes ou de nouvelles facultés qui à leur tour mèneront au développement d’une nouvelle définition¹⁵.

¹⁵ Cette suggestion s’inspire de l’exemple anti-fondationnaliste du bateau de Neurath (Neurath, 1921): un bateau qui, en plein océan, doit être réparé pour poursuivre son trajet. Les matelots à bord de ce bateau n’ont d’autres choix que d’utiliser les fondations déjà existantes de l’embarcation, aussi précaires soient-elles, comme support temporaire pour changer les vieilles pièces. Ce faisant, ils peuvent ensuite utiliser les nouvelles pièces comme

Cela dit, et nonobstant la définition exacte que nous donnons à la cognition, il est pertinent de s'attarder sur les définitions proposées pour cette notion (qu'elles soient éphémères ou non), comme elles influenceront vraisemblablement la démarche localisationniste. Selon qu'on considère la cognition vaguement comme une forme de traitement de l'information ou comme un phénomène inhérent à des types d'organismes particuliers, il pourrait être tentant de rejeter d'emblée l'intérêt de situer celle-ci dans l'espace ou, au contraire, accepter de réduire celle-ci à différents endroits spécifiques du monde, comme le cerveau. Par conséquent, il est de notre avis qu'on devrait accorder plus d'attention aux définitions proposées de la cognition avant de tenter l'identification d'un *locus* définitif dans le monde¹⁶; admettant qu'une telle information pourrait être utile, voire obtainable. Cela n'est pas de dire qu'on doit commencer, dans l'absolu, avec une définition de la cognition, et ignorer les avancés scientifiques qui ont eu lieu en neurosciences et dans les autres domaines, mais plutôt de dire que nous sommes à un moment du cycle où il est temps, avec les nombreuses connaissances dont nous disposons (ex. codage prédictif, affordances, principe d'énergie libre, etc.), de clarifier le sens à accorder à ce terme qui prend de plus en plus d'expansion.

1.2.2. Le problème des frontières de la cognition et de la dispersion spatiale

Même dans un scénario où on pourrait prouver hors de tout doute qu'un critère donné est le bon pour saisir l'essence de la cognition – que ce soit la capacité de traiter de l'information ou la possession de représentations non-dérivées –, il demeure néanmoins difficile de déterminer l'étendue spatiale de ce phénomène dans le monde. La définition proposée par Adams et Aizawa (2010) est un excellent exemple de cela. Nous avons un critère plausible pour déterminer ce en quoi consiste la cognition (nous avons une marque du cognitif), mais nous

support pour changer les pièces qui servaient de fondations, et ce, de manière itérative jusqu'à ce qu'ils arrivent à destination. Nous pensons que le projet de définir la cognition est analogue à cet exemple.

¹⁶ Le lecteur pourrait voir ici une contradiction avec ce qui précède, et objecter que la localisation ferait partie des informations qui pourraient, à chaque cycle, raffiner notre définition de la cognition. Nous croyons davantage que ce cycle repose sur une interaction entre les définitions et les mécanismes proposés pour la cognition. Comme nous allons en parler, un peu plus tard, la localisation des composantes ou des processus n'est pas pertinente pour établir un mécanisme explicatif de l'émergence de phénomènes cognitifs, et, donc, nous ne voyons pas d'intérêt à l'inclure parmi les informations qui pourraient alimenter notre conceptualisation de ces dits phénomènes. Il ne s'agit pas d'une exclusion *a priori*, mais d'une exclusion informée à partir des différents cycles conceptualisation-mécanismes (ex. la relation entre psychologie des facultés et phrénologie, la relation entre la neuropsychologie clinique classique et fonctions cognitives, etc.).

sommes confrontés à un problème qui est d'identifier quels genres de systèmes peuvent produire et/ou contenir des représentations non-dérivées (ex. quels genres de systèmes sont cognitifs et quels ne le seraient pas). Ce problème correspond à la question d'établir les frontières de la cognition.

Un débat qui est au cœur de cette question est celui de déterminer quelles composantes identifiées joueraient un rôle constitutif dans le mécanisme qui donne lieu à la cognition des composantes qui n'auraient qu'un rôle causal (voir Clark & Chalmers, 1998; Menary, 2006; Adams & Aizawa, 2010; Kaplan, 2012; Varga, 2019). Dans le premier cas, on postule que l'élément observé fait partie à part entière du mécanisme qui donne lieu au phénomène d'intérêt, que cet élément produit de la cognition, et, dans l'autre, que l'élément en question ne fait que causer une modification au sein du système sans en faire partie; cet élément est alors non-cognitif. Évidemment, le point pivot de ce débat est de déterminer à partir de quels critères on peut conclure qu'un objet dans le monde fait partie du système cognitif¹⁷. Nous donnerons, dans les paragraphes suivants, deux exemples de propositions au sujet des frontières de la cognition, dont une que nous retiendrons dans notre argumentaire.

Le principe de parité introduit par Clark et Chalmers (1998) est un bon exemple d'une tentative de délimitation des frontières du cognitif. Selon ce principe, si un objet dans le monde produit une fonction ou un processus que nous n'aurions pas de mal à reconnaître comme étant cognitif s'il était produit par le cerveau, alors cet objet peut être considéré comme étant capable de produire de la cognition. L'exemple classique proposé par les deux philosophes est celui de deux personnages: Inga et Otto. Les deux doivent se souvenir d'une date importante, celle d'une exhibition au musée d'art de leur quartier. La mémoire d'Inga est parfaitement normale, alors celle-ci n'a qu'à retenir l'information et l'évoquer au moment où elle en aura le besoin, mais Otto souffre d'Alzheimer, alors il ne peut pas compter sur sa mémoire personnelle pour se rappeler de cette date. Il décide donc de l'écrire dans un calepin qu'il traîne toujours avec lui. Ce calepin remplacerait fonctionnellement, dans cette perspective, le rôle de certaines structures cérébrales associées à des processus comme l'entreposage, la consolidation et la récupération de l'information (ex. l'hippocampe, le cortex associatif, le lobe préfrontal). La question est,

¹⁷ Cette question n'est pas à confondre avec celle de la marque du cognitif, même si les deux sont intimement liées. Pour conclure qu'un objet dans le monde est cognitif, il faut déjà avoir sous la main une définition de ce qu'est la cognition (qu'est-ce qu'il faut comme ingrédients pour qu'un objet donné soit cognitif?). Ici, la question est de déterminer, à partir d'une ou plusieurs définitions de la cognition, quels objets peuvent porter cette marque.

alors, de savoir si ce calepin joue un rôle constitutif ou simplement causal dans la performance cognitive d'Otto.

Tous les participants du débat constitution-causation s'entendent pour dire que les objets extérieurs peuvent causer ou mener jusqu'à un certain point à des modifications dans la cognition (Kaplan, 2012). Le conflit se situe plutôt au niveau de déterminer si des objets, comme le calepin d'Otto, peuvent être considérés comme une composante authentique du système cognitif de ce dernier. Clark et Chalmers soutiennent que si, en vertu du principe de parité. Le calepin en question semble effectivement remplir certaines fonctions analogues à des fonctions mnésiques que nous n'aurions pas de mal à reconnaître comme étant cognitives si elles se situaient dans le cerveau d'un individu, mais leur argument est basé sur une vision large (*coarse-grained*) de la cognition (Sprevak, 2009). Si nous commençons à spécifier les fonctions cognitives, l'analogie s'effondre (Rupert, 2004).

En effet, si on s'attarde plus spécifiquement aux différentes facultés subsumées par la mémoire, la fonction émergente du calepin d'Otto n'est plus comparable à celle qui émerge des structures mnésiques du cerveau à Inga. Par exemple, l'information dans le calepin d'Otto ne fait pas nécessairement l'objet d'une reconsolidation lorsqu'elle est évoquée. Certes, ce dernier peut décider de mettre à jour occasionnellement le contenu de son calepin, lorsque des nouvelles informations parviennent à sa conscience, mais, dans le cas de la mémoire, telle qu'elle se réalise naturellement, l'information qui y est contenu est modifiée indépendamment de la volonté du sujet et son accès peut être facilité ou diminué (Alberini, 2011). C'est d'ailleurs un des mécanismes en cause dans la pathogenèse et la régulation de psychopathologies comme le trouble du stress post-traumatique (TSPT). Dans le TSPT, les souvenirs traumatiques réémergent dans la conscience du sujet, indépendamment de sa volonté, et deviennent de plus en plus accessibles (Kida, 2019), c'est pourquoi des pharmacothérapies ont été développées pour freiner indépendamment la récupération (Dominique & Margraf, 2008) et la reconsolidation (Brunet et al. 2018) des souvenirs traumatiques. Bref, on ne pourrait pas expliquer adéquatement les mécanismes régulateurs de cette psychopathologie en faisant appel à une analogie comme celle du calepin d'Otto.

Un autre exemple de proposition pour démarquer les frontières de la cognition est le critère de manipulabilité mutuelle introduit par Craver (2007)¹⁸. Pour Craver, l'établissement des frontières d'un mécanisme ne devrait pas reposer sur l'organisation spatiale des composantes (chose avec laquelle nous sommes entièrement d'accord), ni sur la présence d'interactions causales entre celles-ci. Comme il le souligne, dans sa discussion de la potentialisation à long terme, il arrive souvent que les parties du mécanisme se déplacent dans des sens opposés et que les relations spatiales entre ces parties soient éphémères. Ces deux points font directement écho à deux aspects qui nous intéresseront au sujet du système cognitif: la directionnalité des processus et le caractère dynamique de l'activité entre les composantes.

Pour remédier aux lacunes de ces deux approches et pour identifier les composantes pertinentes au mécanisme (*constitutive relevance*), Craver propose de reconnaître un élément comme étant constitutif d'un mécanisme lorsque cet élément et le phénomène qui émerge de l'activité du mécanisme sont mutuellement manipulables. Considérer une composante comme étant constitutivement rattachée au mécanisme signifie, comme nous l'avons indiqué plus tôt, qu'elle ne fait pas partie des conditions en arrière-plan, mais qu'elle serait un élément authentique du mécanisme, qu'elle serait véritablement pertinente pour expliquer l'*explanandum*.

Craver distingue trois types de procédures expérimentales qui permettent d'établir la relation entre un élément X du mécanisme et le phénomène Ψ émergent du système S . Le premier, les expériences d'interférence, a comme objectif de diminuer ou éliminer X du mécanisme pour ensuite observer les conséquences de cette intervention sur Ψ . Ce type de procédures comprend les études lésionnelles, dont nous avons parlé, l'inactivation génétique et l'utilisation d'antagonistes pharmacologiques, parmi d'autres. Le deuxième, les expériences de stimulation, utilise, à l'inverse, l'activité de certaines composantes comme témoin de leur implication dans Ψ . Cela peut se faire via la stimulation de régions cérébrales ou via l'utilisation, cette fois, d'agonistes pharmacologiques. Enfin, le troisième type, les expériences d'activation, consiste à faire la procédure inverse par rapport aux deux interventions précédentes, c'est-à-dire d'intervenir sur Ψ pour observer si l'activité de X change. Un exemple de ce type d'intervention serait les études d'fMRI qui porteraient sur la manipulation d'une activité cognitive (ex. la

¹⁸ Il est à noter, toutefois, que Craver ne s'attaque pas spécifiquement à la question des frontières du cognitif, mais plutôt à celle des frontières du mécanisme. Malgré cela, Kaplan (2012) - et subséquemment Varga (2019) - montre comment « *[t]he mutual manipulability criterion is effectively a boundary demarcation criterion* » qui s'applique à la question de déterminer les frontières de la cognition.

mémoire).

Le philosophe soutient que ces types d'interventions prises seules sont insuffisantes pour établir la pertinence constitutive d'une composante putative dans le mécanisme. Dans les deux premiers cas, l'activité de X peut être compensée à la hausse ou à la baisse par d'autres composantes du mécanisme. Il est donc possible, dans les expériences d'interférence, que X joue un rôle important au sein du système, mais que son rôle soit dissimulé par la compensation d'autres parties du mécanisme. De même, il est possible que d'autres composantes que X jouent un rôle important dans le mécanisme, mais que leurs activités soient diminuées, voire interrompues temporairement, en réponse à la suractivation de X . C'est le genre de phénomènes compensatoires qu'on observe avec les processus homéostatiques. Enfin, dans le cas des expériences d'activation, il se pourrait que certaines composantes aient une contribution tonique à Ψ , c'est-à-dire que même si on intervient sur Ψ ces composantes se comporteraient de façon habituelle, sans qu'on puisse observer de différence. Il est également possible que l'intervention sur Ψ produise des épiphénomènes stériles à la compréhension du mécanisme (ex. les fluctuations hémodynamiques qu'on observe en fMRI lorsqu'on mesure la performance à une tâche).

Pour ces raisons, Craver propose que l'appartenance d'une composante putative à un mécanisme devrait être déterminée à partir de deux conditions interreliées: si, dans le cas d'une intervention idéale¹⁹, un changement de la valeur d'une composante X entraîne une modification de Ψ et, qu'à l'inverse, une modification de Ψ amène à une modification de X , on peut dire alors que X appartient au mécanisme du système S qui donne lieu à Ψ , en vertu du critère de manipulabilité mutuelle. Pour la question des frontières de la cognition, cela implique qu'une composante donnée est une composante authentique du système cognitif lorsqu'on peut réaliser des manipulations mutuelles *bottom-up* sur les composantes et *top-down* sur le phénomène d'intérêt.

¹⁹ Pour Craver, une intervention idéale I est une intervention où le changement dans Ψ n'est imputable qu'à un changement de la valeur de X sur laquelle on a réalisé l'intervention. Cela implique, comme le souligne le philosophe, plusieurs choses: 1). I ne modifie pas directement Ψ 2). I ne modifie pas la valeur d'une composante causale intermédiaire entre X et Ψ 3). I n'est pas corrélée avec une autre variable qui pourrait causer Ψ 4). Enfin, I peut activer ou inactiver X indépendamment d'une autre cause.

Reprenons l'exemple du calepin d'Otto, comme il est frais en mémoire. Même si cet objet ne se comporte pas exactement comme la mémoire humaine, dans un cas où on adopterait un niveau de spécificité élevé (*fine-grained*), il est tout de même possible que ce calepin fasse partie des composantes du système cognitif d'Otto. Une manipulation *bottom-up* du calepin (la composante X) modifierait certainement le comportement de ce personnage. Autant dans le cas d'une interférence comme d'une stimulation, le retrait ou l'ajout d'informations dans le calepin pourrait mener Otto à différents endroits, loin ou près de sa destination. Inversement, si on décidait de manipuler le comportement de ce dernier (le Ψ), en lui donnant par exemple de nouvelles adresses pour le musée, nous pouvons compter sur l'idée qu'il voudra consulter et mettre à jour l'information contenue dans son calepin pour se réorienter. Dans cette optique, le calepin d'Otto serait effectivement une composante constitutive de son système cognitif.

Nous croyons, avec Kaplan (2012) et Varga (2019), que la manipulabilité mutuelle est un bon critère pour déterminer les frontières de la cognition pour principalement deux raisons. Premièrement, parce qu'il s'agit, pour reprendre les termes de Kaplan, d'un critère de démarcation générique. La manipulabilité mutuelle ne repose pas sur une prise de position particulière sur l'essence de la cognition. Pour nous, il s'agit d'un énorme avantage, car, comme nous l'avons soulevé dans la section précédente, il existe une myriade de définitions conflictuelles au sujet de la cognition. S'entêter à trancher sur laquelle de ces définitions est la bonne est un projet absolument colossal; un projet qui peut nous dévier d'un objectif important en sciences cognitives: la recherche d'un mécanisme explicatif qui ne dépend pas du localisationnisme. Deuxièmement, et de manière plus importante, ce critère permet de distinguer les effets d'arrière-plan des composantes authentiques du système cognitif. Comme nous l'avons mentionné, une grande partie du débat sur les frontières de la cognition s'articule autour de la question de déterminer les éléments qui ne joueraient qu'un rôle causal de ceux qui joueraient un rôle constitutif. Le critère de manipulabilité multiple nous donnerait un bon outil pour démêler l'un de l'autre.

Enfin, il est bien de noter que le problème de la délimitation des frontières du cognitif va de pair avec l'idée d'une dispersion spatiale des composantes du système cognitif. Comme nous le verrons dans le prochain chapitre, il existe de nombreuses preuves en faveur de l'idée que des structures comme l'univers social (Di Paolo & Jaegher, 2012; Carter et al. 2018), le monde matériel (Sterelny, 2010; Augustinack et al. 2014) et le corps de l'agent (Damasio, 1996; 1998) peuvent contribuer de manière significative à l'émergence et la régulation de plusieurs fonctions

cognitives. Si nous acceptons que ces structures aient une pertinence constitutive au mécanisme explicatif, qu'elles font partie de ce mécanisme, alors les composantes pertinentes à la compréhension de la cognition serait grandement dispersée spatialement (les éléments pourraient s'étendre potentiellement sur une ville complète ou le monde selon la manière que nous concevons le rôle de certains composantes putatives comme les réseaux sociaux). Une telle dispersion – si elle se révèle vraie – mettrait en péril l'intérêt même de localiser les composantes impliquées dans l'émergence de la cognition vu qu'elles pourraient s'étendre à la fois sur l'environnement proximal et distal de l'individu.

1.2.3. Le problème de la taxonomie du mentale et du niveau de spécificité

Advenant qu'on puisse déterminer l'essence et l'étendu d'un système cognitif, deux autres problèmes nous apparaissent à l'égard de la localisation des fonctions cognitives: quelles sortes de fonctions ce système remplit-il et quel degré de spécificité devrait-on adopter lorsqu'on considère ces fonctions? Ces deux problèmes renvoient à la sélection d'une taxonomie du mentale (voir les travaux de Poldrack, Anderson et Cisek) et à la sélection d'un niveau de spécificité pour les fonctions cognitives. Bien que ces problèmes fassent référence à des aspects différents de notre conceptualisation des fonctions cognitives, nous jugeons qu'ils sont tellement enchevêtrés l'un dans l'autre qu'il est acceptable pour les besoins de la discussion de les traiter simultanément dans cette section. En effet, il existe un point de croisement important entre la sélection d'une taxonomie du mentale et le niveau de spécificité qu'on adopte: les deux portent ultimement sur la sélection de facultés cognitives à étudier. Une taxonomie du mentale divise l'esprit en différentes capacités « constitutives »²⁰ (pour des exemples, voir la taxonomie de Gall, 1835, celle de Bloom, 1956 ou, plus récemment, celle de Cisek 2019) et le niveau de spécificité reflète cette division de manière détaillée, selon qu'il soit spécifique (*fine-grained*), ou de manière schématique, selon qu'il soit non-spécifique (*coarse-grained*).

Un bon exemple de la manière dont la taxonomie et le niveau de spécificité sont enchevêtrés est la mémoire humaine. Il est communément accepté, aujourd'hui, que la mémoire se subdivise en un certain nombre de processus: l'encodage, la consolidation, la récupération et la

²⁰ Le terme « constitutif » est utilisé ici dans le sens où l'esprit serait composé de ces capacités, et non dans le sens établi par Craver (2007) lorsqu'il parle des relations constitutives.

reconsolidation en sont des exemples. Mais la mémoire se subdivise aussi en différents types, selon qu'on parle des mémoires temporelles (sensorielle, court terme et long terme) ou des formes de mémoire (mémoire de travail, épisodique, autobiographique, sémantique, etc.). Ces subdivisions peuvent guider notre recherche de mécanismes dans le cerveau (Uttal, 2001), mais elles ne sont pas définitives. Si on augmente le niveau de spécificité, on remarque que certaines de ces sous-facultés ne sont pas étanches et chevauchent d'autres facultés. On retrouve, à ce sujet, beaucoup de cas dans la littérature où la récupération de l'information en mémoire s'entremêle avec les capacités linguistiques (Coughlan & Warrington, 1981; Brackenbury & Pye, 2005) et les capacités d'imagerie mentale (Brewer & Pani, 1996) du sujet, ce qui laisse croire, à plus forte raison, que cette faculté ne serait pas exclusivement qu'un processus mnésique tel qu'on le conçoit traditionnellement, mais un processus mixte qui reposerait sur l'interaction de différentes fonctions cognitives.

Le cas de la récupération de l'information révèle qu'une taxonomie donnée n'est pas coulée dans le béton, mais qu'elle évolue parallèlement à notre compréhension des fonctions qu'elle postule, compréhension qui s'acquiert via l'observation de phénomènes à différents niveaux de spécificité. Même si la récupération de l'information concerne traditionnellement les différentes formes de mémoire, comme la mémoire épisodique et sémantique, par exemple, il se peut que ce processus fasse l'objet d'une révision théorique pour inclure plus explicitement des formes linguistiques et d'imagerie mentale. Cela influence, bien évidemment, l'identification de *loci* dans le monde. Il semble en effet inapproprié de localiser cette faculté dans une petite portion du lobe préfrontal²¹ si, en réalité, elle émerge d'une interaction complexe avec d'autres fonctions cognitives. De même, si le processus de récupération est initié et façonné par la présence d'indices environnementaux (Metzger, 2002) ou de facteurs propres à l'individu, comme la posture du corps au moment du rappel (Dijkstra et al. 2007) et le rythme respiratoire (Tyng et al. 2017), il est difficile de déterminer où cette faculté se situe réellement.

Cela étant dit, nous croyons que les questions de la taxonomie et du niveau de spécificité devraient être également adressées avant de tenter de localiser une fonction cognitive de

²¹ Il s'agit d'un exemple exagéré pour illustrer de manière simple la relation entre granularité et localisation. Dans les faits, la récupération de l'information en mémoire sémantique produit des cartes d'activités distribuées complexes qui comprend des activations au niveau frontal, temporal et pariétal (Binder & Desai, 2011), ce qui est, en un sens, consonant avec l'idée que cette fonction serait enchevêtrée dans d'autres fonctions cognitives, ou du moins qu'elle n'émergerait pas de manière isolée.

manière définitive. De même que pour la question de la marque du cognitif, cela ne signifie pas qu'on doive commencer avec une taxonomie *a priori* de la vie mentale, mais que nous devrions concentrer nos efforts sur une définition adéquate de la cognition et des facultés subsumées par celle-ci avant de proposer une localisation définitive. Dire cela n'exclut pas l'intérêt de développer des modèles explicatifs (ex. des mécanismes), car il est de notre avis que ces modèles contribuent à nourrir notre compréhension des phénomènes cognitifs, mais plutôt que la question de la localisation – s'il est même possible de localiser les fonctions cognitives – devrait être mis au second plan; elle ne devrait être pratiquée que si on est certain de disposer d'une bonne définition et d'une bonne taxonomie de la vie mentale.

1.2.4. Le problème de la réalisabilité multiple et du caractère dynamique de la cognition

Enfin, même lorsque nous avons surmonté chacun des problèmes précédents, que nous avons une définition claire de la cognition, de ses frontières et de ses subdivisions, il demeure des problèmes encore plus importants contre le localisationnisme: le problème de la réalisabilité multiple et celui du caractère dynamique de la cognition. L'idée derrière ces deux questions est que la performance cognitive n'est pas quelque chose de statique, mais un phénomène en perpétuel changement qui émerge de l'activité de différentes structures biologiques et non-biologiques. Nous allons discuter de ces deux questions dans les paragraphes qui suivent et de leurs implications pour le localisationnisme.

D'abord, parlons de la question de la réalisabilité multiple. Cette thèse postule qu'un état ou une propriété mentale donnée peut être le produit de différents substrats physiques (Bickle, 2020). Plusieurs exemples sont donnés dans la littérature pour appuyer la réalisabilité multiple, comme l'exemple classique de Putnam (1967) sur la douleur. Ce dernier note que cette sensation semble être présente dans une grande variété d'espèces, et ce malgré des dissimilarités évidentes dans l'organisation de leur système nerveux. Il en conclut, sur cette base, qu'il y a au moins des situations où nous ne pouvons pas réduire un type d'état mental donné à un type d'état cérébral précis.

Une version plus contemporaine de cette thèse est le concept de la dégénérescence dans les sciences biologiques. Selon cette hypothèse, une fonction donnée, dans l'organisme, peut être

prise en charge par des combinaisons de composantes biologiques dissemblables à différents niveaux d'organisation (Edelman & Gally, 2001). Appliquée aux sciences cognitives, la dégénérescence suggère qu'il n'y a pas de relations un à un, c'est-à-dire des relations *token-token*, entre certaines structures cérébrales et certaines facultés cognitives, comme tend à le suggérer les preuves basées sur l'étude des lésions. Plutôt, on observe des cas de doubles dissociations où une lésion cérébrale à une structure particulière (ex. le lobe préfrontal gauche) *ne* donne pas nécessairement lieu au profil clinique (ex. aphasie de Broca) attendu à partir du modèle des effets de cette lésion (voir Dronkers et al. 2000). Inversement, un profil clinique qu'on aurait rattaché à une structure particulière du cerveau peut être lié à la détérioration de structures autres que celle présumément impliquée dans la fonction normale.

Ces doubles dissociations s'expliquent, comme le soulève Figdor (2010), par le fait que certaines fonctions émergeraient de combinaisons différentes entre 1) des structures cérébrales unipotentes 2) des structures pluripotentes 3) des structures uni- et pluripotentes. Il est important de dissoudre d'entrée de jeu l'ambiguïté autour du concept d'unipotentialité. Dire qu'une région ne contribue qu'à une seule fonction ou qu'elle est spécialisée fonctionnellement ne signifie pas que la fonction qu'elle accomplit est localisée dans cette région. Comme l'explique Figdor, la spécialisation fonctionnelle n'est pas à confondre avec le localisationnisme. Même si une région est spécialisée hypothétiquement pour une fonction donnée, il est possible que d'autres aires cérébrales contribuent à la réalisation de cette fonction, soit en étant redondante fonctionnellement ou en étant complémentaire. C'est l'idée que capture la distinction que nous avons fait plus tôt entre localiser et identifier: nous pouvons identifier des régions qui contribueraient au mécanisme explicatif, sans pour autant réduire la fonction d'intérêt à l'une ou l'autre de ces régions.

Non seulement y-a-t-il de la redondance au sein des réseaux cérébraux (voir Ghanbari et al. 2022), mais il existe aussi des régions pluripotentes, c'est-à-dire des régions qui pourraient, selon le contexte, contribuer à différentes fonctions cognitives. L'aire de Broca, qu'on croyait jadis dévoué uniquement à la production du langage, est maintenant reconnue comme une région pluripotente. Différentes études (voir Marshall & Fink, 2003; Grodzinsky & Santi, 2008) ont d'ailleurs montré son rôle dans une myriade de capacités, dont la reconnaissance et la production de syntaxes musicales, la perception du rythme, la capacité à imaginer des trajectoires, etc. De même, il a été montré que des régions comme l'amygdale, qu'on croyait principalement impliquée dans l'expression de la peur (Ohman, 2005), participe à des processus

aussi variés que la mémoire émotionnelle, la production d'attentes et la cognition sociale (Gothard, 2020). Il va de soi que le fait que des fonctions cognitives puissent émerger de combinaisons entre des structures redondantes et/ou pluripotentes compromet sérieusement la possibilité de localiser ces fonctions. Une même fonction pourrait être prise en charge à différents moments (ou chez différentes personnes) par des structures dissemblables, faisant de sorte que l'endroit où se situerait l'activité observée ne serait jamais exactement le même²².

Pour les besoins de notre argument, nous utiliserons les relations de réalisabilité multiple et de dégénérescence de manière interchangeable lorsque nous parlerons des structures biologiques intrasomatiques qui contribuent à la cognition. Nous sommes conscients que la dégénérescence est un cas particulier de la réalisabilité multiple qui s'applique spécifiquement aux sciences biologiques, alors nous limiterons son utilisation aux cas qui concernent l'organisation fonctionnel du cerveau (ex. les régions et réseaux neuronaux qui contribueraient à la réalisation des fonctions cognitives) et aux cas qui touchent la contribution de l'appareil sensori-moteur et des viscères à la cognition.

Nous pouvons maintenant passer au prochain problème contre le localisationnisme: le caractère dynamique de la cognition. L'idée est que la cognition n'est pas un phénomène statique qui prendrait toujours la même forme indépendamment du contexte, mais une faculté qui s'adapte aux demandes de l'environnement interne et externe de l'individu. Smith (2005) résume parfaitement notre position lorsqu'elle dit que la pensée « [...] *is an in-the-moment unique event, open to a continually changing world, and the product of the intrinsic dynamics of a nonstationary system* » et que « *all knowledge [is] inseparable from real time processes of perceiving, remembering, attending and acting* » (p.279). Le lecteur l'aura compris à partir de ces deux passages, un aspect essentiel des systèmes dynamiques est la temporalité des processus (Metzger, 2017). Non seulement l'ordre des opérations est important pour comprendre le processus étudié, mais aussi les transitions entre celles-ci (Perone & Simmering, 2017).

En effet, selon la perspective des modèles dynamiques de la cognition, nous ne pouvons pas capturer adéquatement les processus cognitifs en ne prenant que des instantanées d'activité (ex.

²² Figdor illustre très bien ce point lorsqu'elle parle des résultats incohérents observés dans les études fMRI. Ces études, pour reprendre son propos, montrent, pour une même tâche, des activations régionales différentes au sein du même individu à différents temps et entre les individus. Souvent, hélas, ces résultats sont interprétés comme étant du bruit de fond ou des erreurs d'interprétation de données, alors qu'ils pourraient en réalité refléter la dégénérescence de la fonction en différents sous-systèmes neuronaux (Price & Friston, 2002).

récupération mnésique à T_1 , récupération mnésique à T_2 ...) comme le font plusieurs devis expérimentaux traditionnels (Metzger, 2017). Il faudrait plutôt observer le phénomène en temps réel si l'on espère comprendre la manière dont il s'organise. Prenons l'exemple de la reconnaissance. Un modèle statique de cette faculté diviserait celle-ci en différentes étapes clés: perception, comparaison de l'information actuelle avec l'information en mémoire, jugement sur l'équivalence de l'information, etc. Chacune de ces étapes pourrait faire l'objet d'une instantanée, mais nous n'aurions pas d'information sur les transitions entre ces étapes. Dans un modèle dynamique, en revanche, la reconnaissance est un processus qui fait appel à une boucle récursive entre des séquences de perception et de reconnaissance partielle de l'information où l'élément récupéré en mémoire est confronté ou enrichi, en temps réel, par des nouvelles entrées dans le système cognitif (Cox & Shiffrin, 2012). Il n'y a pas de divisions temporelles, dans cette optique, mais une série continue non-linéaire d'activité qu'il faut observer pour comprendre le mécanisme qui donne lieu à cette faculté.

Si les facultés cognitives se réalisent dynamiquement dans le sens que postule le modèle de la reconnaissance, alors tenter de localiser ces facultés, sur le plan ontologique, relèverait de l'impossible. Ces facultés n'auraient pas de lieu d'émergence (comme si elles étaient achevées); elles seraient l'aspect observable d'une succession de processus enchevêtrés continu (*ongoing*). Nous pourrions en revanche établir, sur le plan épistémologique, un critère pour diviser cette succession en des instantanées observables²³. Mais, même dans un tel cas, l'acte de localiser serait impraticable, car il reviendrait à sélectionner arbitrairement, parmi toutes les configurations d'activités neuronales – et extraneuronales – impliquées dans l'émergence et la régulation du processus, laquelle de ces configurations seraient le siège, l'endroit, où se réaliserait cette capacité. Ce qui, on nous l'accordera, est inconcevable.

²³ Comme nous allons le voir dans le chapitre 3, on peut visualiser les fonctions cognitives comme des configurations d'activités entre différentes composantes ontologiques au sein de réseaux complexes. L'aspect intéressant d'une telle approche est qu'il est possible de représenter successivement ces configurations d'activités hypothétiques pour les comparer entre elles et pour en dégager les composantes et les processus modifiés entre les configurations.

Conclusion

Nous avons examiné au début de ce chapitre les différentes preuves en faveur du localisationnisme cérébral et, plus spécifiquement, de la réduction ontologique et épistémologique des états mentaux aux états cérébraux qui l'appuie. Le point culminant de cette réduction est l'acceptation d'une équivalence exacte entre des régions cérébrales spécifiques et des facultés cognitives précises (relation occurrence-occurrence). Nous avons vu à ce propos que les preuves fournies par des méthodes d'investigation comme l'étude des lésions, les techniques d'imagerie et d'enregistrement cérébrale soutiennent effectivement l'idée d'une contribution neuronale à la performance cognitive, mais qu'elles sont sujettes à différents problèmes méthodologiques et conceptuels, limitant l'exactitude des inférences produites au sujet de l'équivalence entre les états mentaux et les états cérébraux. Malgré cela, il est de notre avis qu'une combinaison judicieuse de ces outils (ex. fMRI, DTI et EEG) pourraient pallier les limitations individuelles rencontrées avec ceux-ci, et que ces outils constituent des éléments méthodologiques cruciaux à la compréhension de la réalisation des phénomènes cognitifs. Le point focal de cet exposé est que ces méthodes ne devraient pas être mises au service de la question du « où » de la cognition, mais du « quoi » et du « comment ».

Nous avons également soulevé à travers ce chapitre un certain nombre de problèmes associés aux tentatives de localiser la cognition. Parmi celles-ci, les questions de la marque du cognitif, des frontières de la cognition et de la sélection d'une taxonomie du mental ont retenu notre attention. Le problème commun à ces questions est qu'une définition opérationnelle partagée de la cognition et des processus cognitifs devrait précéder les tentatives de localiser ces processus; advenant qu'il soit possible de le faire. Nous avons fourni d'ailleurs, à ce sujet, des arguments importants, avec les questions de la variété ontologique des éléments qui participent à la cognition, la réalisabilité multiple et le caractère dynamique des processus cognitifs, contre la possibilité de localiser définitivement les phénomènes cognitifs. Ces phénomènes se réalisent en faisant appel à des structures variées, dispersées, interchangeable et dont l'activité se reconfigurent moment-par-moment pour répondre aux besoins de l'environnement de l'individu. Bref, il existe de bonnes raisons de croire que la localisation est impossible et que, même si elle l'était, elle ne serait pas d'une grande importance en sciences cognitives.

Chapitre 2

Au-delà de la boîte noire, les échafauds extraneuronaux de la cognition

Nous l'avons vu dans le chapitre précédent, le localisationnisme en neurosciences cognitives est rudement mis à l'épreuve par un certain nombre de problèmes. Non seulement il n'existe pas d'unanimité sur la manière de définir la cognition, mais il n'y a également pas d'entente sur le degré de spécificité (*fine-grained* contre *coarse-grained*) à adopter lorsqu'on traite des fonctions cognitives. Ces problèmes font en sorte qu'il est difficile de délimiter spatialement les phénomènes cognitifs (ex. d'établir des frontières à la cognition et aux fonctions cognitives), comme il n'existe pas encore d'épistémologie stable à leur sujet. Advenant même qu'une telle épistémologie survienne, il existe, comme nous l'avons soulevé, des problèmes plus importants encore auxquels font face le localisationnisme: ceux de la dispersion spatiale des composantes putatives, de la réalisabilité multiple et du caractère dynamique du système cognitif. Ce chapitre est consacré à l'étude de ces deux premiers problèmes. Nous verrons en effet, dans les prochaines pages, qu'il existe des relations de manipulabilité mutuelle entre des structures extraneuronales et des phénomènes de l'ordre du cognitif; de sorte que nous pouvons considérer ces structures comme étant des composantes authentiques du système cognitif. Plus que cela, nous verrons également des scénarios où certaines de ces structures assumerait la charge computationnelle de régions cérébrales, appuyant l'idée que non seulement les fonctions cognitives se réalisent de plusieurs façons sur le plan neuronal, mais aussi, et surtout, qu'elles émergent de différentes configurations d'activités entre des structures variées ontologiquement (ex. le corps, le monde social et le monde matériel).

2.1. Le cadre théorique de l'échafaudage actif de l'esprit

2.1.1. Préambule: qu'est-ce que l'échafaudage actif?

Pour mener à bien ce projet, nous utiliserons tout au long de ce chapitre le cadre théorique de l'échafaudage actif de l'esprit développé par Varga (2018)²⁴. Pour saisir en quoi consiste ce

²⁴ Il est bien de noter que la notion d'échafaud n'est pas spécifique à Varga. Il existe une littérature importante autour de celle-ci (voir Clark 1997; Sterelny, 2010; Maeise, 2016; Krueger, 2018).

cadre, deux éléments doivent être définis: les « échafauds » et ce que le philosophe entend par « actif ». Pour commencer avec le premier, Varga définit un échafaud comme une structure matériellement neutre, c'est-à-dire une structure qui n'est pas rattachée à des *a priori* sur la marque du cognitif et la matière qui composerait l'esprit (ex. unités neuronales, physiques, sociales, etc.), qui peut servir de support substantiel à une habileté ou une fonction. Cette structure peut être une ressource présente naturellement dans l'environnement de l'individu tout comme elle peut être également conçue artificiellement pour soutenir ou augmenter la performance de certaines fonctions cognitives (ex. le calepin d'Otto). Si la notion d'échafaud semble vague à première vue, il s'agit d'un choix délibéré de la part de Varga pour permettre une application plus flexible du cadre théorique à différents cas de contribution extraneuronale aux processus cognitifs, et pour éviter aussi, comme Varga l'indique, quelques débats stériles sur la nature de la cognition.

Il faut savoir néanmoins que les échafauds ne sont pas nécessairement des parties constituantes du système cognitif *a priori*. Comme le précise Varga, « *the term scaffolding includes a rather large number of relationships of dependence, [that's why] it is necessary to exclude trivial ones* » (p. 51). Les échafauds peuvent aussi bien faire référence à des conditions de possibilité pour la réalisation de certaines facultés (ex. il est nécessaire d'avoir un organisme pour réaliser une fonction cognitive), à des conditions facilitatrices (ex. un bon niveau de glucose sanguin peut faciliter telle ou telle fonction) ou à des relations de constitution. Il est donc important de se prémunir d'un critère pour distinguer les échafauds réellement importants à l'émergence des processus cognitifs des échafauds qui ne jouent qu'un rôle trivial, celui de la manipulabilité mutuelle introduit dans le chapitre précédent. Nous y reviendrons.

Il existe trois types d'échafauds différents selon Varga: l'échafaud intrasomatique, l'échafaud intersomatique et l'échafaud extrasomatique qui comprend l'échafaud précédant. Pour faire simple, le premier fait référence à la contribution de structures non-neuronales de l'organisme à la cognition, dont l'appareil sensori-moteur, les viscères, le système hormonal, etc. Varga donne plusieurs exemples d'échafaudage intrasomatique dans son ouvrage, dont ceux de l'injection de Botox et de la prise de Tylenol pour illustrer différents niveaux de complexité au sein des échafauds. Dans l'exemple du Botox, qu'il qualifie de « simple », les muscles faciaux paralysés par l'injection de ce produit sont directement liés à l'amélioration de symptômes dépressifs. L'hypothèse évoquée est que les muscles zygomatiques de la bouche seraient impliqués dans différentes émotions déplaisantes présentes dans la dépression, telles que la

colère, la peur et la tristesse, et que l'inhibition de ces muscles contribuerait à diminuer la sensibilité de l'amygdale à des stimuli négatifs, en raison du circuit qu'ils partagent (Magid et al. 2016 dans Varga, 2018).

Dans le deuxième exemple, qui est plus complexe, la consommation de Tylenol réduirait la douleur sociale par le biais d'une structure médiatrice. En effet, et contrairement à l'injection de Botox où l'activation des muscles zygomatiques atténueraient directement les symptômes dépressifs, le Tylenol affecterait indirectement la douleur sociale via les circuits neuronaux en charge de traiter des signaux nociceptifs en provenance du corps. Ces circuits, comme l'indique Varga, échafauderaient la sensation de douleur associée au rejet et à l'isolation sociale (cette sensation partagerait des racines communes avec la douleur corporelle), mais ces circuits sont eux-mêmes échafaudés sur l'activité de l'organisme en entier. Il y a là une espèce d'emboîtement des échafauds, d'où la complexité présumée de cette intervention.

Même si Varga prend le soin de distinguer différents niveaux de complexité avec les exemples de l'injection de Botox et du Tylenol, il n'est, en revanche, pas exactement clair dans quelle mesure les autres exemples d'échafaudage qu'il donne à travers son ouvrage appliquent cette distinction (ils semblent pour la plupart être du second type), comme il ne mentionne que très laconiquement le motif derrière celle-ci, disant dans un court passage: « *it is not hard to see that thinking in terms of simple scaffolding will not do justice to the Tylenol case. Instead, an explanation in terms of complex scaffolding is suitable* » (p.58). Ce que nous pouvons retirer des propos de Varga, avec une once de spéculation, est que l'étiquette « simple » s'applique à des cas où:

- 1) Il y a une relation relativement directe, telle qu'on pourrait la représenter avec un lien entre deux nœuds d'un graph, entre la structure qui échafaude et l'effet qui lui est attribué.
- 2) (qui découle naturellement du point (1) précédent) Il n'y a pas un autre type d'échafaud interposé ou une structure médiatrice entre la structure identifiée et son effet.
- 3) Cet effet ne devrait pas être attribué à un autre ou plusieurs autres échafauds ou structures que celui identifié initialement.

Nous n'appliquerons cependant pas cette distinction dans ce qui suit, car nous croyons que, vu l'enchevêtrement complexe des structures qui participent à la cognition, qu'il est difficile de départager des cas d'échafaudage simples de ceux qui seraient complexes. Par exemple, il est possible, dans le cas de l'injection de Botox, que des structures médiatrices non mentionnées

entre les muscles zygomatiques et l'amygdale, comme certaines régions pariétales et motrices, modulent la relation entre ces structures.

Avec cette clarification en tête, poursuivons avec les différents types d'échafauds. Le deuxième, l'échafaud intersomatique, renvoie aux interactions motrices, affectives et sociales qui peuvent émerger entre des organismes. Un bon exemple de ce type d'échafaudage est le phénomène de synchronisation intersomatique dont parle Varga. L'idée derrière cette notion est que des agents, des entités pour reprendre les termes du philosophe, peuvent se coordonner mutuellement et dynamiquement pour former des organisations comportementales autonomes (p.157) et soutenir la régulation d'une activité. Cette idée n'est pas à confondre avec de la simple imitation ou avec une action commune entreprises par les agents. La synchronisation est l'aboutissement d'un jeu d'interactions complexes de coordination dynamique entre ces agents lorsqu'ils prennent part à une même activité. Aussi, et contrairement à l'utilisation usuelle de ce terme, la synchronisation, telle que l'entend Varga, « *need not involve perfectly symmetric timing* » (p.162). Il est même important, selon ce dernier, qu'il y ait des moments d'inadéquations entre les états des agents pour promouvoir l'adaptabilité et la créativité, bien qu'il n'élabore pas davantage sur ce point.

La littérature scientifique regorge d'exemples de synchronisation intersomatique. Par exemple, nous tendons spontanément à synchroniser nos mouvements lorsque nous marchons avec d'autres personnes (Bernieri & Rosenthal, 1991; van Ulzen et al. 2008). Il en va de même aussi pour d'autres types de mouvements, comme les mouvements qui consistent à se balancer ou à taper des doigts de manière rythmée (Richardson et al. 2008; Oullier et al. 2008 dans Varga, 2018). Mais la synchronisation se joue aussi à des niveaux comportementaux et cognitifs plus complexes, comme dans la régulation émotionnelle. Par exemple, on observe des interactions réciproques coordonnées dans la dyade mère-enfant qui aurait comme objectif la régulation étendue des états émotionnels du nourrisson (Feldman, 2003; Reck et al. 2004; Varga, 2018). En effet, bien que les bébés peuvent employer jusqu'à un certain point des stratégies d'auto-apaisement (Adams et al. 2022), ces stratégies sont souvent insuffisantes pour réguler adéquatement les émotions qu'ils vivent. La synchronisation avec le parent, dans cette perspective, permettrait de distribuer la tâche complexe de la régulation émotionnelle du nourrisson sur les interactions réciproques continues avec la figure d'attachement. De même, nous voyons également un phénomène de synchronisation dans les relations amoureuses. En effet, les partenaires romantiques peuvent servir, de manière analogue à la figure parentale dans

la dyade mère-enfant, de ressource étendue pour la régulation des états émotionnels de leur partenaire, et même compenser pour des déficits émotionnels associés à certaines psychopathologies (Marroquin & Nolen-Hoeksema, 2015 dans Varga, 2018). Nous reviendrons plus en détail sur la question de la synchronisation dans la section dédiée à l'échafaud intersomatique.

Enfin, le dernier échafaud, l'échafaud extrasomatique, renvoie de manière plus générale à l'ensemble des structures extérieures à l'agent, comme les différents objets et contextes qui composent son monde immédiat, ce qui inclut par ailleurs les interactions intersomatiques dont nous venons de discuter. Toutefois, dans le but d'éviter des confusions potentielles entre les deux types d'échafauds, nous réserverons l'utilisation de ce terme aux objets et aux contextes matériels qui peuvent influencer l'individu. Cela dit, l'exemple du calepin d'Otto que nous avons mentionné plus tôt dans ce mémoire illustre bien ce en quoi consiste ce type d'échafaud. En effet, on peut voir ce calepin comme une structure qui contribue substantiellement aux capacités mnésiques et agentives d'Otto. Cet objet informe Otto à différents moments de son parcours vers le musée, de manière analogue à un processus cognitif comme la récupération de l'information sémantique en mémoire. Mais il n'y a pas que le calepin d'Otto qui peut soutenir la cognition. Il existe une myriade d'autres exemples d'échafauds extrasomatiques: le papier et le crayon pour soutenir la réalisation d'opérations arithmétiques (Clark, 1989), la planche de Scrabbles pour le rappel des mots (Kirsh, 1995), les téléphones portables pour une panoplie de fonctions comme la récupération de l'information et la navigation spatiale (Bruineberg & Fabry, 2022), etc.

Cela étant dit, nous pouvons maintenant poursuivre avec la question de ce que Varga entend par *échafaudage actif*. Pour le philosophe, une structure donnée échafaude *activement* une fonction lorsque l'échafaud et l'échafaudé entretiennent des relations de manipulabilité mutuelle (MM). Pour rappel, une relation de manipulabilité mutuelle signifie que, dans un système S , une intervention idéale sur une composante X modifie le phénomène émergent Ψ de S , et, inversement, une intervention idéale sur Ψ modifie la valeur de X . Varga spécifie toutefois que les relations MM dont il parle dévient du sens originel que leur accorde Craver (2007) sur un aspect. Pour Craver, il ne peut y avoir des relations MM entre une composante X et Ψ que si X fait partie de S . Or, comme l'indique Varga « [...] *the exact nature of the requirement that putative components be contained within S has remained underspecified and runs the risk of becoming too inclusive or exclusive* » (p.121).

Cette déviation du sens originel n'est pas incohérente avec notre position initiale sur la valeur de MM comme critère de démarcation de la cognition, tout au contraire. Nous croyons que « X fait partie de S » ne devrait pas être une condition préalable à l'identification de relations MM, mais plutôt la résultante de l'application de MM comme critère de démarcation. L'un de nos objectifs est, après tout, d'identifier quelques-unes des composantes qui participent constitutivement à l'émergence et la régulation des processus cognitifs. S'il était facile de déterminer les éléments constitutifs du système cognitif *a priori*, nous n'aurions pas à réaliser ce travail. Mais, comme nous l'avons indiqué dans le chapitre précédent, l'un des problèmes auquel font face les sciences cognitives est justement de déterminer à partir de quels critères on peut conclure que de tels éléments appartiennent réellement au domaine du cognitif.

Varga précise aussi que les relations de MM se doivent d'inclure des relations de *nature-dependance* et de *difference-making*. Dans l'optique du *nature-dependance*, l'échafaudé – la fonction cognitive – ne dépend pas trivialement de l'échafaud pour son existence (ex. la cognition dépend du fonctionnement du cœur, car sinon il n'y a pas d'organisme en vie pour produire de la cognition). L'échafaud contribue carrément à la nature même du phénomène échafaudé, c'est-à-dire à sa forme et à ses propriétés. Par exemple, le corps n'est pas qu'une condition triviale d'existence pour la récupération de souvenirs émotionnels (ex. sans organisme, il n'y a pas d'agent pour récupérer quoi que ce soit), mais jouerait un rôle actif dans la capacité de l'individu à se remémorer de ce genre d'informations. Riskind (1983) et Wilson & Pepper (2004) ont d'ailleurs montré à ce sujet que les postures corporelles et les expressions faciales qu'on adopte lorsqu'on tente de se rappeler d'une expérience personnelle influencent le genre de souvenirs rappelés en mémoire. En effet, selon que les participants, dans les deux études, adoptèrent une posture « verticale », confiante, ou une posture « abattue », avec la tête inclinée vers le sol, ces derniers se souvenaient différemment d'expériences positives et négatives. Dans la première condition, les participants se souvenaient davantage d'expériences positives, alors que les participants de la seconde condition se souvenaient généralement plus d'expériences négatives.

Cet exemple illustre également le principe de *difference-making*, c'est-à-dire la possibilité de manipuler le phénomène échafaudé en faisant des interventions *bottom-up* sur l'échafaud. En effet, on voit qu'une manipulation de l'échafaud – ici, la posture, – contribue à façonner le genre de souvenirs qui seront remémorés; qu'en manipulant certains éléments posturaux, comme en indiquant au sujet de baisser sa tête, qu'il est possible de moduler la probabilité de

rappel d'un souvenir avec une valence négative. Plus que cela, on peut dire qu'il existe une relation MM entre la posture et la valeur émotionnelle des souvenirs. Dans un sens, manipuler la posture des participants affecte la valeur émotionnelle des souvenirs récupérés, mais, dans l'autre sens, manipuler les émotions affectent également, en retour, la posture du corps, comme on peut le voir dans les cas de dépression (Dael et al. 2011; Peper & Lin, 2012).

2.1.2. Pourquoi des échafauds?

Bien que le cadre théorique de l'échafaudage actif ne soit pas la seule option à notre disposition pour exposer le rôle de l'environnement extraneuronal dans l'émergence et la régulation des processus cognitifs, il est préféré à d'autres approches de la cognition 4E pour différentes raisons. D'abord, parce qu'il répond mieux aux besoins de la critique qui est articulée dans ce mémoire. En outre, il est de notre avis que les échafauds tels qu'ils sont présentés par Varga ont, chacun, une extension conceptuelle mieux définie que les thèses traditionnelles des approches 4E prises à part (ex. *embodiment*, *cognition embedded*, *énactivisme*, etc.). Mais aussi, parce que « *[the] claims that mental processes are embodied, enacted or extended have become interpreted in so many different ways that it is, perhaps, no longer advisable to define one's view in terms of them* » (Rowlands, 2015, p. 311).

Prenons, par exemple, la cognition *embedded*. Selon cette approche, la cognition n'est pas un événement détaché de l'univers physique, social et culturel de l'individu, mais le produit d'une rencontre située et contextuelle de ces aspects (Van der Schyff et al. 2018). C'est effectivement une position que nous endossons personnellement, mais celle-ci est, à notre avis, trop inclusive conceptuellement pour une analyse individualisée des composantes extraneuronales, en englobant à la fois contexte matériel et social. On peut penser que ces domaines sont logiquement indépendants l'un de l'autre, comme il serait théoriquement possible pour une personne d'adhérer à l'un tout en rejetant l'autre. Si cette personne, par exemple, endosse une position purement développementale et socio-constructiviste de la cognition, elle pourrait être amenée à accepter l'influence de l'environnement social comme, au moins, causalement pertinente à la performance cognitive, mais refuser d'accorder un statut particulier à l'univers matériel à l'égard de la cognition.

Il en va de même pour la cognition étendue où la démarcation entre ces dimensions ne sont pas toujours claires (Varga, 2018). Non seulement cela, mais, comme le soulève Sterelny (2010), plusieurs partisans de la thèse de la cognition étendue considèrent que les objets du monde extérieur sont des composantes authentiques du système cognitif, que la cognition encapsulerait en quelques sortes ces objets par défaut. À l’opposé, l’idée que la cognition est échafaudée n’implique pas, comme nous l’avons mentionné plus tôt, que les structures échafaudantes appartiennent *a priori* au domaine du cognitif (c’est précisément pourquoi Varga introduit le concept d’échafaudage actif pour distinguer les cas de conditions de possibilité et de conditions facilitatrices des cas de constitution). Plusieurs objets de notre quotidien soutiennent différents types de processus (ex. cognitifs, moteurs, sociaux, etc.), sans qu’on puisse admettre qu’ils font partie constitutivement de ces processus. Sterelny en donne d’ailleurs un excellent exemple avec l’analogie de la digestion étendue. En effet, notre capacité à manger et à digérer efficacement de la nourriture est grandement soutenue et façonnée par les outils que nous avons à notre disposition pour manipuler, cuire et agrémenter les différents aliments qui composent notre régime alimentaire. Toutefois, même si nous utilisons des instruments comme des couteaux pour dépecer la viande que nous consommons ou des poêlons pour cuire nos aliments, il est difficile d’affirmer que ces objets font partie constitutivement de notre système digestif. Il en va de même pour d’autres structures échafaudantes de notre environnement. Une souris d’ordinateur nous permet de naviguer avec aisance les longues pages d’un document, à la manière que nos doigts peuvent guider visuellement la lecture, mais il est difficile d’inclure cet appareil dans notre système visuel.

Cela nous ramène au point que nous avons soulevé dans la section précédente: l’identification des composantes constitutives du système cognitif. Les échafauds ne contribuent pas tous de la même façon à la réalisation des fonctions cognitives (ex. à titre de condition de possibilité, condition facilitante, etc.), et que certains d’entre eux n’ont que des rôles causaux de second-plan, voire même qu’ils ne contribuent pas du tout à une fonction particulière - ou du moins qu’ils contribueraient si indirectement ou minimalement qu’il serait impertinent de les prendre en considération dans une analyse qui viserait à expliquer le phénomène Ψ . Non seulement cela, mais certains éléments (somatiques, intersomatiques et extrasomatiques) peuvent à certains moments participer constitutivement à la réalisation d’une performance cognitive et, à d’autres moments, être absents des composantes participantes. Parfois, le partenaire d’un couple peut servir de support authentique à la récupération de l’information (voir Harris et al. 2014 pour une discussion du rappel partagé), et, d’autres fois, cette même fonction peut être soutenue par

un autre élément de l'environnement social et matériel, comme un cellulaire (Bruineberg & Fabry, 2021).

Ce qu'il faut retenir est que la réalisation et la régulation d'une fonction cognitive est, comme nous l'avons mentionné dans le chapitre précédent, un phénomène dynamique (Smith, 2005; Metzger, 2017), qui implique l'activité interactive de différents échafauds à différents moments pour achever sa finalité (ex. l'information récupérée en mémoire). Dans cette perspective, il est mieux, selon nous, de voir la question de la constitution, non pas dans l'optique de la réalisation d'un processus achevé (ex. X fait partie du système S pour la récupération de l'information comme processus achevé), mais dans l'optique d'une succession d'instantanées qui mènent à la performance cognitive observée (ex. X fait partie de S à un moment donné du processus de récupération de l'information).

Bien sûr, la tâche de retracer les configurations d'activités successives entre les différents échafauds responsables de l'émergence d'une performance cognitive donnée est excessivement complexe. Il s'agit d'une tâche qui, comme nous le soutiendrons dans le chapitre suivant, pourrait faire l'objet d'un domaine de recherche entier: ce que nous nommerons provisoirement les neurosciences cognitives réseautiques étendues. Pour nous simplifier la vie, nous proposerons dans les prochaines pages des exemples de processus achevés pour parler des relations de manipulabilité mutuelle et de constitution avec l'idée que si certains échafauds contribuent constitutivement à un moment important du processus cognitif (ex. un moment qu'on jugerait essentiel d'identifier pour expliquer l'émergence de la récupération de l'information²⁵), il est alors possible d'affirmer que l'émergence du dit processus repose de manière non-triviale sur des composantes extraneuronales.

Cela étant dit, nous pouvons passer aux différents types d'échafauds. Les prochaines sections offriront, à cet égard, un regard plus approfondi sur les éléments encapsulés par les échafauds intra- inter- et extrasomatiques, leurs ramifications, et la possibilité que certaines de ces structures appartiennent authentiquement au système cognitif, via le critère de démarcation MM.

²⁵ Quelques exemples de moments cruciaux dans la récupération de l'information en mémoire seraient, par exemple, le moment où la recherche de l'information est initiée, l'identification de l'information (lorsque celle-ci est trouvée dans les structures mnésiques) et le rappel conscient de cette information (que ce soit sous une forme imagée ou linguistique).

2.2. Échafaud intrasomatique: comment le corps contribue à la cognition?

Nous avons présenté sommairement quelques exemples d'échafauds intrasomatiques dans la dernière section avec les exemples de l'injection de Botox pour la dépression et la prise de Tylenol pour la douleur sociale, mais il existe une panoplie d'autres exemples dans la littérature qui soutiennent le rôle de cet échafaud dans l'émergence et la régulation des processus cognitifs. Cette section présentera en détail deux exemples d'échafaudage intrasomatique qui concernent des structures du corps et des processus cognitifs différents. Nous nous efforcerons de montrer, à chaque fois, que ces structures participent à titre de composantes constitutives à la réalisation des phénomènes cognitifs présentés.

Commençons avec un premier exemple. Des participants d'une étude sont installés devant l'écran d'un ordinateur, et on leur donne l'instruction de reproduire un modèle composé de blocs de différentes nuances de gris à l'aide de blocs présents dans une réserve (voir la figure 2 ci-dessous). Quelle stratégie utiliseront-ils? S'efforceront-ils de mémoriser laborieusement l'emplacement de chacun des blocs qui composent le modèle? Ou, peut-être, parviendraient-ils à grouper certains blocs ensemble en fonction de leur similarité pour réduire la complexité de la tâche? Notre intuition voudrait que ces participants utilisent d'une façon ou d'une autre une stratégie mnésique cérébrocentrée, qu'ils parviennent à maintenir en tête sur une période de temps suffisamment longue l'agencement des blocs pour reproduire le modèle avec un minimum d'interactions avec celui-ci. Or, comme le montre l'équipe de Ballard (1995; cité par Clark, 2008), ce n'est pas exactement le cas. Ils s'avèrent que, dans la plupart des cas, les participants de leur étude n'utilisaient pas des stratégies aussi complexes pour réaliser la tâche. Plutôt, on remarque qu'ils ont fait de nombreuses saccades oculaires entre les différentes sections de l'écran, avant et après le déplacement de chacun des blocs, ce qui suggère que les participants n'utilisaient pas de stratégies mnésiques intensives pour retenir l'information, mais plutôt qu'ils filtraient l'information (ex. couleur ou position du bloc) par le regard en fonction des besoins de la situation.

Comme le résume Clark (2008): « *when glancing at the model, the subject stores only one piece of information: either the color or the position of the next block to be copied (not both). [E]ven when repeated saccades are made to the same site, very minimal information is retained. Instead, repeated fixations provide specific items of information “just in time” for use* ». La

manipulation de l'information, ici, n'est donc pas une question de rétention en mémoire, mais une question de mouvements oculaires répétés où l'information est sélectionnée au gré des besoins de l'individu alors qu'il accomplit la tâche.

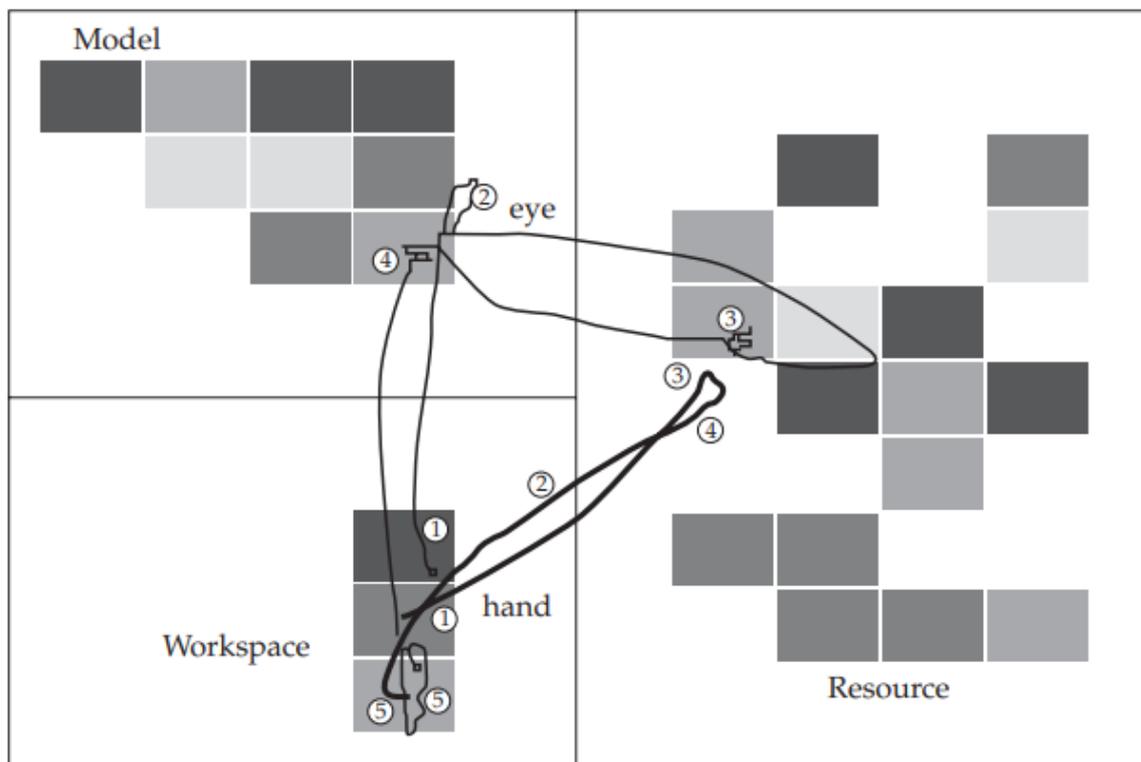


Figure 2.1 - Schéma illustrant la tâche que devaient réaliser les participants de l'étude de Ballard et al. (1995). Dans le coin supérieur gauche, on voit le modèle à reproduire, dans le coin inférieur gauche, l'espace dont disposent les participants pour le reproduire, et, enfin, dans la case de droite, les différents blocs nécessaires pour reconstruire le modèle. On remarque, par les lignes noires, que différentes saccades oculaires ont eu lieu entre les cases, indiquant que les participants ont régulièrement fait appel aux informations visuelles présentes dans leur environnement pour compléter la tâche.

Malgré leur rôle dans la capacité à reproduire le modèle, on peut se demander si les saccades oculaires peuvent être considérées comme une composante authentique de la cognition dans ce contexte. L'application du critère de démarcation MM nous révèle que si. Nous voyons, dans l'expérience de Ballard, que les saccades oculaires contribuent de manière *top-down* à la capacité des participants de reproduire le modèle. Elles permettent efficacement de filtrer l'information visuelle présente à l'écran pour retenir que les éléments nécessaires moment par moment. Mais, inversement, on peut croire qu'une intervention *bottom-up* sur cette faculté affecterait significativement la capacité des participants à accomplir la tâche. C'est précisément ce que Ballard et son équipe ont montré dans la suite de leur étude. Dans cette deuxième expérience, les participants étaient assignés à une condition où ils devaient reproduire le même

modèle que dans la première expérience, à l'exception qu'ils devaient, cette fois, le faire en fixant visuellement le centre de l'écran. Les saccades oculaires étant éliminées, leur performance chuta nettement. Le temps pour accomplir la tâche était approximativement trois fois plus grand que celui requis dans la condition où les mouvements oculaires étaient permis^{26 27}.

Un deuxième exemple d'échafaud intrasomatique nous provient de l'hypothèse, maintenant bien connue, des marqueurs somatiques de Damasio (1996; 1998; 2018) qui sont des cartes somatosensorielles transitoires de l'activité viscérale et musculosquelettique de l'organisme. Dans le cadre de cette hypothèse, Damasio fournit des preuves bien articulées sur l'interaction entre la régulation homéostatique, qui est à l'origine de la production et la différenciation des émotions, et la prise de décision rationnelle, traditionnellement considérée comme le paroxysme de l'activité cognitive et intellectuelle. Il commence par décrire son expérience clinique avec des patients qui ont subi des lésions au niveau du lobe frontal ventromédian. Cette structure, normalement, achemine les informations émotionnelles en provenance du système limbique vers les centres cérébraux responsables de la prise de décision et la planification.

Or, chez ces patients, à qui cette connexion neuronale fait défaut, il leur est impossible ou très difficile de prendre une bonne décision, surtout une qui est bénéfique sur le long terme. Ceux-ci vont généralement privilégier des choix « payants » immédiatement, même s'ils présentent un risque démesuré, ou ils se retrouveront devant l'incapacité de prendre une décision tout court, en étant bloqués sur un raisonnement trop fastidieux de coûts-bénéfices. L'hypothèse avancée par Damasio pour expliquer ce profil cognitivo-comportemental particulier est que l'activité des viscères, qui rend compte de la bonne ou mauvaise régulation homéostatique au sein de l'organisme, n'informe plus la délibération. L'individu, pour le dire autrement, est déconnecté

²⁶ Il est bien de noter que Ballard et al. (1995) ont mené d'autres expériences afin de contrôler quelques-unes des variables qui auraient pu affecter le temps requis pour compléter la tâche, comme la diminution de la résolution visuelle associée à la fixation.

²⁷ Même si les participants réussissent la tâche sans faire appel à des saccades oculaires, cela ne signifie pas que ces mouvements sont facultatifs - ou qu'ils ne joueraient qu'un rôle facilitateur dans la réalisation de cette tâche. Ces résultats suggèrent plutôt que la capacité à reproduire un modèle pourrait être réalisée de différentes façons est que nous avons développé des stratégies élaborées pour compenser la perte de certaines capacités essentielles, en l'occurrence ici les mouvements oculaires. On pourrait penser dans cette optique que les saccades oculaires serait spontanément la stratégie employée par la plupart des participants - comme elle est moins énergivore en termes de ressources cognitives -, mais que des processus mnésiques de plus haut niveau peuvent intervenir pour appuyer ou pour remplacer fonctionnellement cette capacité en cas de besoin.

des marqueurs somatiques, et, conséquemment, n'est plus en mesure d'attribuer une valence à ses choix.

Loin d'être une perspective cérébrocentrique, ces cartes sont le fruit d'une boucle continue avec le corps (*body loop*); elles sont la codification des différents signaux somatiques en un langage neuronal qui pourra ensuite être utilisé dans les computations nécessaires à la prise de décision (ex. en permettant de distinguer la valeur positive ou négative des choix offerts à l'individu). Ces cartes peuvent également être récupérées à des moments ultérieurs pour simuler l'activité du corps (*as-if body loop*) et guider, encore une fois, le processus de décision. Toutefois, ce scénario ne signifie pas que la prise de décision est entièrement assumée par cet organe. Certes, le cerveau jouerait un rôle important, comme coordinateur des marqueurs somatiques, mais l'homéostasie est un processus dynamique qui a continuellement lieu en arrière-plan dans notre organisme (Billman, 2020). Il est difficile de croire alors que les boucles simulées par le cerveau seraient la source première d'information homéostatique pour la prise de décision. Il est plus raisonnable de penser que les boucles interagissent de manière dynamique sur différents plans pour informer l'individu, et que les boucles *as-if* seraient principalement utilisées dans des contextes où l'information homéostatique en provenance du corps serait insuffisante ou perturbée, comme dans un contexte où l'intéroception serait affectée par une maladie (Bonaz et al. 2021).

Néanmoins, qu'il s'agisse de boucle corporelle ou de boucle simulée, il apparaît clair que les viscères jouent un rôle crucial dans la prise de décision. Leurs activités ne font pas qu'influencer distalement le processus décisionnel, mais font partie constitutivement de celui-ci, en vertu du critère MM. D'un côté, l'activité des viscères guide l'individu en colorant émotionnellement les différents choix qui se présentent à lui, lui permettant ainsi de filtrer les nombreuses options et de trancher sur celles qui, à la lumière de son état homéostatique présent et anticipé²⁸, offrent potentiellement la meilleure valeur adaptative. Mais, de l'autre côté, les différentes options que l'individu pèse dans son analyse affectent également en temps réel sa régulation homéostatique (Joffily & Coricelli, 2013; Korn & Bach, 2015) et émotionnelle (Gosling et al. 2020). Celui-ci peut éprouver de la peur après avoir considéré une option désavantageuse ou de la joie pour une

²⁸ Joffily & Coricelli montent dans leur article de 2010 comment les émotions positives et négatives servent à indiquer le taux de changement dans l'énergie libre. Traduit à la question de l'homéostasie, les émotions vécues permettraient d'anticiper la valeur de certaines actions sur la régulation homéostatique de l'individu.

option prometteuse. Ce nouvel état émotionnel peut alors être utilisé dans une nouvelle boucle corporelle pour orienter l'individu vers un choix spécifique ou l'en écarter, jusqu'à ce qu'il soit fixé sur le « bon » choix à prendre. Dans cette optique, il est faux de considérer la prise de décision comme un processus statique qui n'a lieu que dans le cerveau, il s'agit plutôt d'un processus interactif dynamique entre les signaux viscéraux, la régulation homéostatique et l'analyse que l'individu fait de sa situation.

2.3. Échafaud Intersomatique: comment l'univers social contribue à la cognition?

Passons maintenant à un cas plus complet d'échafaud intersomatique. Nous avons donné quelques exemples de ce type d'échafaudage avec la notion de synchronisation intersomatique, dont celui de la régulation émotionnelle dans la dyade mère-enfant, mais cet échafaud ne se limite pas qu'à ce genre de cas. La littérature regorge d'exemples sur le rôle des interactions sociales comme support à la cognition. Nous allons examiner en détail dans les prochains paragraphes un de ces exemples: celui des systèmes de mémoire transactionnelle. Nous prendrons soin, tout comme dans la section précédente, de montrer la présence de relations MM entre les structures identifiées et le phénomène qu'elles soutiennent.

Commençons avec notre exemple. En 1986, le psychologue social Daniel Wegner publia un ouvrage intitulé « *Transactive memory: a contemporary analysis of the group mind* ». Dans cet ouvrage, le psychologue décrit un phénomène qui va à contresens des idées de son temps sur les effets de la pensée de groupe (Theiner, 2013). En effet, la psychologie sociale était marquée à l'époque par l'idée que la pensée de groupe était un phénomène nuisible qui mine le jugement et, de manière plus large, les capacités rationnelles des individus au sein du groupe (voir les exemples classiques de Asch (1956) et de Milgram (1963); voir également Allport, 1968). Or, Wegner défendit le contraire: le groupe peut soutenir efficacement la réalisation de tâches organisationnelles et cognitives (ex. mnésiques). En effet, le psychologue observa que des systèmes de mémoire peuvent émerger à partir des procédures collaboratives - les transactions - entre les membres d'un groupe, les systèmes de mémoire transactionnelle. Ces procédures sont l'ensemble des activités par lesquelles des informations sont échangées, encodées, entreposées et récupérées entre les membres d'un groupe.

Il ne faut pas se méprendre sur la conception que Wegner a des systèmes de mémoire transactionnelle. Ces systèmes ne se réduisent pas à un échange d'information entre plusieurs systèmes isolés, mais sont la résultante d'une collaboration coordonnée entre les membres du groupe par laquelle de *nouvelles* informations et habiletés sont rendues accessibles aux membres (Theiner 2013). Autrement dit, ces systèmes sont en mesure de produire des nouvelles informations qui ne peuvent pas être retracées à la mémoire individuelle de ses membres.

Prenons l'un des cas discutés par Wegner comme exemple:

Imagine [...] that a couple is leaving a party. At different times, they each talked to Tex. The male notes that Tex was depressed this evening: he stared at the floor and barely talked. The female says that Tex was not at all depressed; in fact, she saw him for quite a while early in [said] party and he seemed unusually frisky and friendly. The male recalls that Tex said he was thinking about separating from his wife. And in short order, the couple reaches a conclusion: Tex was flirting with the female and feeling embarrassed about it in the presence of the male (p.267).

On voit, dans cet exemple, comment la conclusion tirée par le couple, celle que Tex séduisait les femmes présentes à la fête, n'était initialement pas accessible aux membres du couple séparément. L'homme avait sa propre perspective sur la soirée où il se souvenait d'un Tex morose et distant socialement. Sa copine, quant à elle, se souvenait plutôt de Tex comme un personnage enjoué et amical. C'est en combinant les informations, de fil en aiguille, qu'ils sont parvenus à la nouvelle information. On pourrait évidemment appliquer le même principe à des exemples moins anodins, comme dans le cadre d'une investigation policière où des enquêteurs doivent collaborer pour identifier un criminel²⁹. Dans ce scénario réaliste, les enquêteurs se sont divisé les tâches pour accélérer l'enquête. L'un d'eux a prêté une attention particulière au lieu du crime et a retenu la présence d'un objet *X* jugé pertinent à l'enquête. Un autre s'est intéressé à la liste de personnes ayant fait l'acquisition de *X* récemment. Enfin, le dernier enquêteur a épiluché la liste de suspects potentiels à partir des alibis présentés. Ils se réunissent un coup l'enquête préliminaire effectuée et mettent en commun les différentes informations obtenues. Ce faisant, ils parviennent à déduire que c'est le suspect *Y*, qui a commis le crime basé sur son

²⁹ On retrouve un exemple similaire avec l'étude anthropologique détaillée de Hutchins (1996) sur l'équipe de navigation du *Palau*, un navire de type porte-avion rattaché au Navy. Dans cette étude, l'anthropologue s'est intéressé aux relations informationnelles et fonctionnelles entre les membres de l'équipage, et, plus spécifiquement, sur la manière qu'ils s'organisent entre eux pour réaliser des comportements complexes, comme la navigation maritime.

achat de X , la présence de X sur les lieux du crime et l'absence d'alibi. Ils ont fait ensemble ce qu'on pourrait appeler de la « déduction logique étendue », et sont parvenus à une nouvelle information qui n'est pas retraceable à aucun des enquêteurs pris seul.

Toutefois, ce n'est pas tous les groupes qui sont capables de former un système de mémoire transactionnelle efficace. Dans certaines circonstances, on observe que le groupe ne contribue pas à améliorer la performance individuelle des membres par rapport à ce que prédit le modèle des effets de la collaboration entre ceux-ci sur la performance. Dans ce cas, le groupe « *serves as an accurate conduit for individual member performance but nothing more* » (Pavitt, 2003). Il se peut même, dans d'autres circonstances, que le groupe nuise à la performance finale des membres; ce qu'on peut qualifier d'inhibition collaborative. La présence de ces deux scénarios ne signifie pas que les procédures collaboratives entre les membres sont généralement inefficaces ou qu'elles contribuent de manière instable à la réalisation des fonctions cognitives, mais que certaines conditions doivent être respectées pour que la performance du groupe soit optimale. Nous pouvons les définir comme:

1) La présence de caractéristiques idiosyncrasiques (ex. traits de personnalité) ou contextuelles qui motivent les membres du groupe à s'impliquer et à partager efficacement des informations. Comme le soulève Theiner (2013), plusieurs caractéristiques relatives à la personnalité des membres peuvent influencer leur capacité à partager efficacement leur expertise: leur capacité à s'affirmer, leur motivation (ex. certains membres peuvent croire qu'il est inutile de transmettre certaines informations ou, dans un contexte compétitif, qu'il est plus avantageux de garder ces informations pour eux-mêmes), leur rôle conversationnel (ex. est-ce que les membres qui possèdent les expertises nécessaires sont plus du genre à communiquer les informations ou à écouter passivement les autres?), la taille du groupe (ex. s'il est trop large, il se peut que les membres perdent la sensation d'être utile individuellement ou qu'ils distribuent la responsabilité épistémique aux autres membres), etc. Clairement, ces caractéristiques jouent un rôle essentiel dans les procédures collaboratives entre les membres, et peuvent affecter la performance globale du groupe.

2) Le niveau d'expertise des membres et le degré de chevauchement épistémique.

Le groupe bénéficie adéquatement de l'expertise de chacun de ses membres s'il existe un bon équilibre entre la ségrégation de l'expertise et la redondance (Theiner, 2013). Autrement dit, un groupe ne devrait pas être composé que d'experts avec rien en commun; on se retrouverait dans

un scénario où l'information serait ségréguée et ne pourrait pas être partagée correctement entre ceux-ci par manque de terrain épistémique commun. Nous ne voulons pas non plus, en revanche, que les membres d'un groupe aient trop d'informations en commun, car on tomberait à ce moment-là dans un scénario de redondance où il n'y aurait pas de progrès épistémique. Il faut donc un juste milieu entre les deux.

3) La stabilité de l'environnement et des membres présents dans le groupe.

L'environnement dans lequel se produisent les procédures collaboratives doit rester suffisamment stable à travers le temps pour permettre aux membres de se rassembler et d'apprendre à se connaître (ex. domaines d'expertise, traits de personnalité, rôle conversationnel, etc.)³⁰. Il en va de même pour la composition du groupe. Les membres ne sont pas inter-substituables, au sens des systèmes agrégatifs (Theiner, 2013; voir aussi Wimsatt, 1986); ils apportent chacun une contribution spécifique par le biais de leurs connaissances propres et de leurs traits de personnalité. Enlever une personne, en ajouter une ou en échanger une pour une autre est autant de facteurs qui peuvent modifier la dynamique du groupe et influencer la capacité des membres à collaborer de manière fructueuse.

4) Le groupe ne doit pas être trop large.

Wegner ne spécifie pas exactement le nombre idéal de membres dans un groupe pour maintenir un système de mémoire transactionnelle authentique, mais la plupart de ses exemples classiques portent sur des dyades, comme des couples amoureux, des relations patient-docteur ou, encore, des relations enseignant-étudiant. Plus récemment, des travaux ont montré que cette notion s'applique bien à des groupes de travail triadiques (Liang et al. 1995) et à de petites organisations (Lewis, 2003; 2011). Cela dit, Wegner considérait qu'il était peu probable « *that larger social ensembles would possess the requisite organizational capacities to construct and maintain a genuine [transactive memory system]* » (dans Theiner, 2013). Combiné avec les

³⁰ Cela dit, les procédures collaboratives entre les membres sont aussi fructueuses dans des contextes où l'environnement est dynamique (Argote & Ren, 2012). Ren et al. (2006) ont montré par exemple que les systèmes de mémoire transactionnelle étaient plus efficaces lorsque les problèmes auxquels faisaient face les membres du groupe changeaient. Lorsque les problèmes restaient stables, en revanche, l'information échangée entre les membres diminuait. L'idée d'un environnement dynamique n'est pas nécessairement en contradiction avec celle d'une stabilité. Un environnement peut combiner à la fois stabilité et dynamisme. Par exemple, il y a plusieurs caractéristiques importantes de notre appartenance qui restent stables à travers le temps, même si nous y apportons des modifications quotidiennes (voir Clark, 1996). Le genre de stabilité auquel on fait appel concerne davantage les caractéristiques spatiales de l'environnement (ex. le lieu, son organisation, etc.), alors que les études susmentionnées portent sur l'évolution des problèmes et de l'information dans le temps. Ce qu'on peut retenir de ça est qu'il est important pour le groupe de trouver un équilibre sain entre un environnement physique stable et un environnement intellectuel dynamique.

conditions précédentes, on pourrait dire que le groupe ne devrait pas excéder une dizaine de personnes pour réaliser optimalement sa capacité à soutenir les fonctions mnésiques de ses membres.

L'idée importante à retenir ici est que, similairement aux autres échafauds qu'il faut trier, ce n'est pas tous les systèmes de mémoire transactionnelle qui sont susceptibles de soutenir efficacement la réalisation des fonctions cognitives. Certains de ces systèmes peuvent produire des résultats neutres équivalents à la performance individuelle des membres du groupe, et, d'autres, peuvent nuire à cette performance. Cet état des choses n'est pas problématique pour notre argument, car on peut voir le même genre de relations entre la réalisation de certaines fonctions cognitives et l'activation de régions cérébrales. En effet, même dans un modèle cérébrocentré de la cognition, la réalisation adéquate des processus cognitifs repose sur un équilibre délicat dans l'activité cérébrale; un équilibre quant aux régions activées, leurs propriétés spécifiques et leurs contributions respectives. Ce qui ne manque pas de nous rappeler analogiquement les conditions nécessaires pour que le groupe soutienne correctement la performance cognitive de ses membres.

Cela étant dit, nous pouvons voir assez aisément comment des relations de manipulabilité mutuelle pourraient apparaître entre les membres du groupe et les systèmes de mémoire transactionnelle. D'un côté, la manipulation des caractéristiques du groupe affecte considérablement l'efficacité des systèmes mnésiques émergents. Augmenter la ségrégation de l'expertise au sein du groupe, par exemple, conduirait potentiellement à un « gouffre épistémique » où l'information ne serait plus partagée adéquatement entre les membres. De même en ce qui concerne des caractéristiques comme la stabilité et la taille du groupe. Ajouter, éliminer ou échanger un membre pourrait modifier l'équilibre épistémique et la dynamique générale du groupe de sorte à diminuer la récupération et le partage d'information entre les individus (ex. en changeant les rôles conversationnels, la motivation, la confiance des membres, etc.).

De l'autre côté, on voit également qu'une manipulation des processus englobés (ex. encodage, entreposage, récupération, etc.) par les systèmes de mémoire transactionnelle affecte la performance individuelle des membres. En effet, supposons que nous prenions un groupe d'individus qui performant bien ensemble. On apprend en analysant la dynamique conversationnelle du groupe qu'une des raisons pour lesquelles ce groupe encode et récupère

bien l'information est que les membres partagent un même cadre théorique de référence. Sachant cela, supposons maintenant que nous prenions les individus à part, que nous leur transmettions des informations contradictoires relativement à ce cadre théorique et qu'ensuite nous leur donnions l'occasion de discuter ensemble des informations obtenues. Il y a fort à parier que si nous leur demandions d'utiliser ce cadre théorique la performance mnésique générale du groupe sera amoindrie. C'est le phénomène bien documenté de l'inhibition rétroactive où l'apprentissage d'une nouvelle information nuit à la récupération d'une information apprise antérieurement (Smirnov, 1973; Royer et al. 1977) - mais à l'échelle du groupe.

2.4. Échafaud extrasomatique: comment le monde matériel contribue à la cognition?

Enfin, nous pouvons passer au dernier type d'échafaud, l'échafaud extrasomatique. Nous avons donné précédemment quelques exemples de son rôle avec le calepin d'Otto et l'utilisation des téléphones cellulaires, mais cet échafaud ne se limite pas qu'à des objets physiques particuliers. Il peut inclure également des espaces et des contextes de la vie de tous les jours. En effet, que ce soit au travail ou à notre domicile, nous avons tendance à structurer les espaces que nous fréquentons régulièrement pour distribuer et faciliter la régulation de nos états affectifs. Krueger (2018) résume cette idée avec éloquence:

We routinely individualize our affective niches [...] via manipulative strategies that highlight how self-selected items of material culture act directly on aspects of our affective states, modulating their character and function—including the bodily processes (e.g., changes in autonomic nervous system activity, behavioral expressions, etc.) states of action readiness (dispositions to act in certain ways), evaluations or appraisals, and phenomenal experiences that make up the complex architecture of our affective states.

Un exemple de cette utilisation de notre espace matériel est l'implémentation d'intention. Nous pouvons structurer nos espaces quotidiens de sorte à soutenir la réalisation de certaines fonctions cognitives, comme la récupération de l'information en mémoire (pour plus d'information, voir Orbell et al. 1997; Gollwitzer, 1999; Prestwich et al. 2015). L'idée derrière la notion d'implémentation d'intention est simple: lorsque nous ne sommes pas certains de se souvenir ultérieurement de nos intentions actuelles concernant certaines tâches à accomplir,

nous pouvons disposer les objets dans notre environnement pour qu'ils permettent le rappel de nos intentions passées.

Un étudiant craignant de ne pas se souvenir de remettre sa dissertation à temps pourrait laisser traîner volontairement un exemplaire de son travail sur sa table pour que la vue de celui-ci lui rappelle à un moment futur l'obligation de le remettre. Il pourrait alors vaquer à d'autres tâches, dans sa journée, sans avoir à fournir un effort conscient pour se souvenir de son objectif, et, au moment opportun, se souvenir de soumettre son travail lorsqu'il verra la copie traîner. Dans cette perspective, on pourrait dire que la copie de l'étudiant remplacerait fonctionnellement les régions du lobe frontal en charge d'initier la récupération de l'information (pour une discussion plus approfondie de ces régions, voir Dickerson & Eichenbaum, 2010; Rugg & Vilberg, 2013). La seule différence est qu'à l'instar de ces régions cérébrales l'élément qui initie le processus de récupération est exogène à l'organisme.

Encore une fois, nous pouvons montrer la présence de relations MM entre l'échafaud et le phénomène échafaudé. Ici, il s'agit respectivement de l'objet placé dans l'environnement et de la récupération ultérieure de l'information (qui est l'intention). Si nous reprenons notre exemple, nous observons qu'une manipulation de l'objet dans l'environnement peut provoquer la récupération de l'information si cet objet entretient un lien avec l'intention à rappeler (ex. la copie de la dissertation pour se souvenir de remettre ce travail). Qu'en est-il par contre du cas inverse? Est-ce que le rappel de l'intention peut modifier les objets implémentés? Oui, c'est possible. À différents moments de la journée, l'étudiant de notre exemple pourrait se souvenir de certains détails à la vue de son travail sur la table. Il pourrait se rappeler que le travail est à remettre virtuellement, et, ensuite, décider de déplacer sa copie sur le clavier de son ordinateur pour augmenter les chances qu'il s'en souviene. Il pourrait notamment se souvenir que ledit appareil n'est pas connecté à l'imprimante défectueuse de son appartement, et que, la raison pour laquelle il était aussi stressé quant à la remise de son travail est qu'il devait se rendre rapidement à son Université pour numériser son travail. À ce moment-là, il pourrait décider de mettre directement sa copie dans son sac à dos et laisser traîner cet objet devant la porte d'entrée de son appartement le temps qu'il finisse de se préparer.

Plus que de dégager une relation de manipulabilité mutuelle entre l'échafaud et l'échafaudé, cet exemple illustre un aspect intéressant de la façon dont des capacités mnésiques comme la récupération de l'information pourraient être échafaudées par l'environnement de l'individu.

En effet, il révèle comment un tel processus pourrait être la résultante d'interactions réciproques complexes et variées entre ce dernier et son univers matériel; que l'objectif « se souvenir de remettre sa dissertation » peut être décomposé en une longue chaîne d'opérations cognitives soutenues différenciellement par la niche écologique de l'individu (ex. plusieurs moments de rappel indicé par différents objets, plusieurs instances de planification et de modification de l'environnement, etc.). À un moment du processus, c'est la copie de la dissertation qui traîne sur la table qui pourrait initier le rappel de l'intention, et, à un autre, la vue du sac à dos près de la porte. Il se pourrait même que ce soit la copine de l'étudiant - qui appartient à l'échafaud intersomatique - qui lui rappelle de numériser son travail. L'idée est que l'atteinte de l'objectif, ici, repose sur une succession de ce qu'on pourrait appeler des *configurations d'activité transitoires* entre différentes ressources neuronales (ex. les régions sensorielles, motrices, exécutives de l'étudiant, etc.) et environnementales (ex. la dissertation, le sac, une personne, etc.). Nous aborderons cette question plus en détail dans le chapitre suivant lorsqu'on introduira la notion de TALoPASS.

Cela dit, les échafauds ne se limitent pas qu'aux ressources immédiatement accessibles à l'individu, comme tendent à le montrer les différents exemples que nous avons utilisé jusqu'à présent (ex. les viscères, nos collègues de travail, les différents objets qui tapissent nos espaces du quotidien, etc.). Ils peuvent inclure des structures qui ne font pas directement partie de notre environnement de la vie de tous les jours, comme des institutions. En effet, Gallagher et Crisafi montrent dans leur article de 2009 comment certains processus cognitifs comme la prise de décision peuvent être externalisés et soutenus à différents niveaux par les institutions en place dans notre société (ex. le système juridico-légal, les écoles, etc.). De manière analogue à Clark & Chalmers (1998), ils nous invitent à considérer trois exemples qui reflètent des niveaux différents de distribution de la tâche cognitive:

- 1). Dans le premier cas, Alexis, une personne quelconque, doit évaluer la légitimité d'une affirmation, dans ce qu'on peut comprendre comme un cadre juridico-légal, et prendre une décision basée sur l'analyse des faits qui lui sont présentés. Elle doit ensuite formuler trois questions au sujet de ces faits et y répondre sans avoir recours à de l'aide extérieure.
- 2). Dans le deuxième cas, Alexis est dans une situation similaire, sauf que, cette fois, elle est accompagnée d'un groupe d'experts qui l'aide à formuler les trois questions. Ce groupe lui fournit également un ensemble de réponses possibles à partir duquel elle peut choisir. Elle peut décider cependant de formuler ses propres réponses.

3). Enfin, dans le dernier cas, Alexis est encore accompagnée par le groupe d'experts qui l'aide à formuler ses questions. Cette fois, en revanche, les experts l'informent de certaines règles préétablies qu'elle doit suivre lorsqu'elle répondra aux questions. Alexis doit choisir seulement une réponse parmi deux choix pour chacune des questions, et elle ne peut pas formuler de réponses personnelles.

La question qui se pose à partir de ces exemples est la suivante: dans lequel de ces scénarios y a-t-il plus d'activité cognitive? Gallagher et Crisafi affirment que l'effort est équivalent entre ceux-ci. Il n'y a pas un scénario « moins cognitif » qu'un autre. C'est simplement que le processus de prise de décision est de plus en plus distribué au fil des exemples pour inclure des ressources intersomatiques, les experts, et extrasomatiques, les règles institutionnalisées. Dans le premier cas, Alexis est en mesure de formuler - difficilement - les questions et d'y répondre à l'aide de ses propres moyens. Elle aurait pu très bien faire appel à une feuille de papier, et nous aurions accepté qu'il s'agit d'un effort cognitif de sa part. Dans le deuxième cas, la démarche d'Alexis est soutenue par le groupe d'experts d'une manière analogue à ce qu'on a décrit avec les systèmes de mémoire transactionnelle. L'information sur les questions et les réponses n'était pas initialement présente dans le cerveau d'Alexis, mais lui ont été rendues accessibles par le biais des procédures collaboratives entre cette dernière et les experts.

Enfin, dans le dernier cas, le processus de décision est assisté par les différentes règles préétablies par l'institution juridico-légale. Ces règles, soutiennent les deux philosophes, sont le résultat d'une longue délibération collective par laquelle nous avons externalisé une partie de l'effort cognitif nécessaire pour accomplir certaines tâches, comme celles qui concernent les décisions juridiques. Plutôt que de générer des règles particulières propres à chaque cas spécifique qui se présente, nous avons établies d'avance des règles pour baliser le traitement de ces cas (ex. lorsque le meurtre est prémédité, la personne reçoit une sentence X , mais, s'il y a présence de circonstances atténuantes $A, B, C \dots$, alors il se peut que cette personne reçoive une sentence amoindrie Y), et nous les avons consignées à travers des ouvrages de référence et des formations qui servent à préparer et à réguler les compétences des personnes qui exercent des professions dans le domaine juridico-légal (ex. législateur, avocat, juge, etc.).

Nous pouvons comparer le rôle de l'institution légale dans la prise de décision à celui qu'a le système académique, au sens large, dans la résolution d'un problème mathématique. Lorsque nous sommes confrontés à la tâche de résoudre un problème de ce genre, nous n'avons pas à

inventer de nouvelles formules à chaque fois, ni à en démontrer la validité. Cet effort cognitif a été accompli par de nombreux mathématiciens (ex. les experts) avant nous et a été consigné dans des ouvrages, des programmes d'enseignements et des exigences ministérielles qui incitent et facilitent la transmission de ces connaissances. Parce que cet effort cognitif a été réalisé à notre place, nous pouvons résoudre des problèmes de plus en plus complexes, et ce, en épargnant beaucoup de temps sur des longues démonstrations. Nous savons d'avance que pour un certain problème, étant donné les conditions spécifiées, une méthode de résolution est plus appropriée qu'une autre. Cette méthode nous est donnée de manière analogue aux règles que le système juridique confère à Alexis dans l'exemple ci-dessus, et sa transmission accomplit un rôle fonctionnel similaire: connaître d'avance une méthode ou une règle spécifique pour traiter un problème limite le nombre d'options à considérer et diminue la charge cognitive qui serait associée à ce travail.

Cela étant dit, nous pouvons nous poser la question à savoir si les institutions susmentionnées ne joueraient pas simplement qu'un rôle facilitateur (ex. diminuer la charge cognitive accaparante du sujet, ce qui lui permettrait en retour de réaliser plus rapidement ou plus efficacement la tâche). Encore une fois, l'application du critère MM nous montre qu'on peut considérer ces institutions comme des composantes authentiques du système cognitif. D'un côté, des institutions, comme le système scolaire, soutiennent la résolution de problèmes en fournissant sur mesure des connaissances et des règles qui façonnent la démarche de résolution de problèmes mathématiques. Mais, inversement, il se peut également qu'en tentant de résoudre ces problèmes en faisant appel aux institutions que nous sommes confrontés à des impasses et qu'en bout de compte les règles données ne soient pas justes. Dans un tel scénario, nous pouvons modifier le contenu des institutions conformément à nos nouveaux besoins épistémiques et sociétaux, modifications qui en retour offriront des opportunités cognitives aux individus et qui pourraient mener à d'autres révisions de ces institutions.

2.5. Objections possibles à l'échafaudage de l'esprit

Une présentation du cadre théorique de l'échafaudage actif de l'esprit ne serait pas complète sans discuter minimalement de quelques objections possibles. Une première pourrait concerner la nature vague des échafauds (ex. comme structures), objection tout à fait valable

épistémologiquement. Il est, en effet, difficile de départager les échafauds qui contribuent réellement à la réalisation des processus cognitifs de ceux qui jouent un rôle secondaire causal ou facilitateur. C'est la raison pour laquelle Varga se dote du critère MM, et c'est, en partie, pour la même raison que nous l'avons utilisé. Cependant, et étant donné la visée de ce travail, nous ne voyons pas cet état des choses comme étant problématique. Nous présenterons dans le chapitre présent une proposition de modélisation qui se veut flexible quant aux structures à retenir dans le mécanisme explicatif, et la notion d'échafaud, telle que définie par Varga, accommode ce désir de flexibilité. Un modélisateur quelconque pourrait décider de retenir autant de structures qui l'intéresse pour expliquer un phénomène, que ces structures n'aient qu'un rôle facilitateur ou qu'ils jouent un rôle constitutif dans le mécanisme; en autant qu'il possède de bonnes justifications pour retenir ces éléments (ex. le critère MM).

Une deuxième objection pourrait s'appuyer empiriquement sur des contre-exemples où certains échafauds sembleraient absents ou inaccessibles à une personne, et montrer que cette personne performe bien cognitivement. Nous pouvons penser à des cas spécifiques comme celui du tristement défunt Stephen Hawking ou celui d'Helen Keller³¹, à qui plusieurs éléments de l'échafaud intrasomatique faisaient défaut. La réponse à cette objection passe à nouveau par la question de la réalisabilité multiple. Les processus cognitifs peuvent être réalisés et soutenus à différents moments par des échafauds variés, de sorte que même en étant incapable de tirer profit de certaines composantes intrasomatiques (ex. appareil locomoteur, saccades oculaires, etc.), ces personnes ont réussi à produire efficacement de la cognition, assez pour atteindre des statuts prestigieux.

On pourrait dire, dans le cas d'Hawking, que les échafauds intersomatiques et extrasomatiques ont compensé largement pour les déficits intrasomatiques que lui ont causé sa condition. Dans le premier cas, on sait, à partir de sa bibliographie (2013), que plusieurs personnes ont joué un rôle clé dans sa réussite, notamment sa première femme lorsqu'est venu le temps de rédiger sa thèse doctorale. Ils s'étaient dotés ensemble d'un système précis d'interprétation des mouvements faciaux à partir desquels sa femme pouvait rédiger sa pensée. Leur relation ne manque pas de nous rappeler les procédures collaboratives au sein des systèmes de mémoire transactionnelle. Du côté de l'échafaud extrasomatique, on sait aussi que Stephen Hawking

³¹ Helen Keller (1880-1968) est une autrice et conférencière née sourde et aveugle. Elle est, selon l'histoire, la première personne handicapée à avoir obtenu un diplôme universitaire.

disposait d'appareils variés, comme une chaise roulante et un synthétiseur vocal. Ces appareils remplaçaient fonctionnellement certaines habiletés typiquement intrasomatiques, comme la locomotion et la production de la parole, de sorte qu'il pouvait tout de même bénéficier du support de ces habiletés.

Une dernière objection, plus grave que les précédentes, serait d'attaquer directement la validité du critère MM. Cette objection consisterait à démontrer que ce critère n'est pas fondé empiriquement, et que les objets identifiés diffèrent trop quant à leur nature pour appartenir au système cognitif. Nous croyons que cette objection serait largement infondée, car le critère MM reprend informellement, dans son essence, la pierre angulaire à partir de laquelle s'articule le réductionnisme psychoneural. En effet, nous avons vu, avec l'étude des lésions, qu'un argument important en faveur de la réduction onto-épistémologique des états mentaux aux états cérébraux est la présence de doubles associations (quand X est activé, il y a Y , et quand il y a Y , X est activé) et dissociations (quand X est lésionné, Y est affecté négativement, et quand on observe une perturbation d' Y , on voit qu' X est lésionné) entre des régions du cerveau et des fonctions cognitives. Or, c'est précisément ce que révèle la manipulabilité mutuelle: des doubles associations entre des composantes putatives et des fonctions cognitives. Rejeter le critère MM, dans cette optique, impliquerait de démontrer que la présence de doubles associations anatomofonctionnelles entre le cerveau et la cognition est insuffisante pour établir le rôle constitutif des régions cérébrales, et que cet organe serait le seul *loci* de la cognition en vertu de propriétés intrinsèques précises; ce qui on l'admettra n'est pas gagné d'avance.

Conclusion

Dans ce chapitre, nous avons abordé le cadre théorique de l'échafaudage actif de l'esprit présenté par Varga (2018). Ce cadre, comme nous l'avons défendu, présente une bonne voie d'entrée pour s'attaquer aux questions de la variété ontologique des composantes qui participent à la cognition, de leur dispersion spatiale et de la réalisabilité multiple, sans être teinté par certains *a priori* sur la nature de la cognition (ex. définition et marque du cognitif). À travers cet exposé, nous avons vu, à l'aide du critère de manipulabilité mutuelle, que la cognition peut être soutenue constitutivement par des structures variées et dispersées dans l'espace – les échafauds intra-, inter- et extrasomatiques –, mais aussi, dans certains cas, que des processus,

comme la mémorisation des caractéristiques spatiales d'un modèle peut être réalisée de différentes façons, soit par des structures cérébrales en charge de maintenir activement ces caractéristiques à l'esprit et/ou par les saccades oculaires (ces choix ne sont pas exhaustifs ni mutuellement exclusifs). De même, nous avons vu, avec la notion d'implémentation d'intention, que l'initiation de la récupération de l'information en mémoire peut être réalisée par des composantes variées comme des objets disposés dans l'environnement et/ou des personnes. Quoi qu'il en soit, la chose importante à retenir de ce chapitre est que les processus cognitifs ne découlent pas systématiquement et exclusivement de l'activité cérébrale. Chacun des exemples que nous avons abordé questionne la primauté absolue du cerveau dans les mécanismes qui expliquent les processus cognitifs: saccades oculaires, marqueurs somatiques, systèmes de mémoire transactionnelle, objets et institutions diverses, etc.

Chapitre 3

La cognition comme réseau complexe distribué et dynamique

Nous avons, jusqu'à présent, fait état de différents problèmes qui impactent la possibilité de localiser la cognition dans l'espace, et, plus spécifiquement dans le cerveau, dont le problème de la distribution spatiale des composantes putatives du système cognitif, de la réalisabilité multiple et du caractère dynamique de la cognition. Nous avons d'ailleurs montré, dans le chapitre précédent, comment la cognition peut être réalisée et régulée par différents types de structures non-neuronales, intrasomatiques, intersomatiques et extrasomatiques; concluant sur l'idée que, en raison de l'hétérogénéité ontologique et la dispersion spatiale des composantes qui participent à la production et la régulation des fonctions cognitives, il est impraticable de situer spatialement la cognition. Les fonctions cognitives ne sont pas des phénomènes statiques qu'on peut simplement situer dans des espaces circonscrits définitifs, comme des espèces d'instantanées isolées, mais des processus qui émergent d'interactions complexes et dynamiques entre différents types d'objets dans le monde.

Le point culminant de ce mémoire est que l'entreprise localisationniste de la cognition, particulièrement le pan qui tente de situer cette dernière dans le cerveau, est vouée à l'échec. La cognition n'est pas un phénomène statique qui émergerait toujours de la même façon (ex. à partir des mêmes structures), mais un processus qui se réalise dynamiquement en recrutant différentes structures neuronales et extraneuronales selon les besoins et les contraintes de l'agent. Aucune structure ciblée ne produit à elle seule la faculté étudiée (le lobe préfrontal médian n'est pas l'unique responsable du rappel épisodique) à toutes les fois où la fonction est réalisée, et aucune structure ne produit que la faculté en question (le lobe préfrontal médian est impliqué dans d'autres fonctions que le rappel). Qui plus est, certaines de ces structures, comme nous l'avons vu dans le chapitre précédent, peuvent être interchangées fonctionnellement par d'autres qui peuvent ou non être relativement présentes dans l'environnement immédiat de l'individu, sans affecter les propriétés essentielles du système, créant une multiplicité d'espaces contextuels propres à chaque événement cognitif. Il n'y aurait donc aucune localisation définitive pour la cognition, au-delà de localisations éphémères inintéressants pour une science qui tente d'expliquer les mécanismes exacts derrière les phénomènes cognitifs.

Cet état des choses nous amène au point central de ce dernier chapitre: si la question du « où » de la cognition est mésadaptée pour comprendre et expliquer la manière qu'émerge cette capacité, quelles questions devrions-nous nous poser alors en sciences cognitives? Nous suggérons, à titre de conclusion de ce mémoire, que la question de la localisation des fonctions cognitives devrait être remplacée, comme nous l'avons suggéré dans le chapitre 1, par les questions du « quoi » et du « comment ». De *quoi* la cognition est composée pour prendre la forme particulière qu'elle a, et, par extension, quel est le rôle et l'importance relative de chacune des composantes identifiées dans l'émergence moment-par-moment de celle-ci? Dit autrement, dans quelle proportion les différents échafauds identifiés jusqu'à présent, les différents objets subsumés par ceux-ci et les interactions possibles avec ces objets contribuent à produire et réguler l'état cognitif dans lequel se trouve un individu particulier? C'est en orientant la recherche sur les composantes de la cognition et de leurs interactions dynamiques, le comment, que nous parviendrons à dégager de celle-ci des informations pertinentes sur son fonctionnement, et non par le biais de la localisation.

Évidemment, nous l'avons soulevé d'orchestre avec Uttal (2001), les facultés mentales sont des entités extrêmement difficiles à étudier comparativement à leurs équivalents matériels (dans une perspective localisationniste classique portée sur le cerveau) qui se prêtent avec beaucoup plus de facilité à l'observation et aux mesures directes. En l'absence de ces équivalents matériels - cérébraux -, les meilleurs témoins de la cognition se trouvent dans le comportement - de manière large - des individus, et dans notre capacité à inférer à partir de celui-ci la présence d'un procédé « intelligent », la cognition, qui le produirait (Wilkinson & Halligan, 2004; Rowe & Healy, 2014), chose qui est loin d'être infallible conceptuellement et méthodologiquement. Sans le recours à la localisation, nous risquerions fort bien d'être confrontés aux mêmes problèmes qui ont affligés le béhaviorisme logique (pour une discussion plus approfondie, voir Putnam, 1967; Ribes-Iñesta, 2004), ou pire de nous trouver devant une métaphysique insurmontable de la vie mentale.

L'objectif de ce chapitre est de sauver, dans une certaine mesure, les sciences cognitives de ces destins funestes. Nous allons voir à travers les pages qui suivent qu'il est possible de modéliser la cognition de manière à en faire un sujet d'investigation scientifique légitime (dans les règles de l'art), sans se limiter méthodologiquement à l'étude des comportements, mais, surtout, et de manière plus importante, sans faire appel à des *localisations définitives* et *cérébrocentrées*. La cognition prend naissance dans le monde, est observable et est également mesurable *en soi* (pas

par le biais d'un proxy); il ne s'agit pas d'un phénomène dont on postulerait « arbitrairement » l'existence faute de données empiriques. Nous pouvons l'étudier concrètement via des *occurrences transitoires d'activité*, soit des configurations particulières d'interactions relativement bien délimitées temporellement entre différentes structures neuronales et extraneuronales qui correspondent à – sont – des moments particuliers dans la réalisation des états cognitifs (ex. des moments clés pour comprendre l'émergence du processus), et surtout via les transitions entre celles-ci.

Ces *occurrences* auraient, certes, une localisation momentanée dans le monde. Elles auraient une signature spatio-temporelle bien précise qui nous permettrait de les saisir et de les étudier comme tout autre phénomène de la physique, mais, il ne faut pas se méprendre, ce n'est plus la question de la localisation en elle-même qui nous intéressera, contrairement à l'approche cérébrocentrée à laquelle nous avons été exposé jusqu'à présent dans ce mémoire, mais les principes organisateurs qui feraient que des structures de différents types (ou d'un même type) s'unissent d'orchestre pour produire et réguler un état cognitif particulier. Autrement dit, ce sont les *occurrences transitoires d'activités* entre les structures neuronales et les échafauds intra-, inter et extrasomatiques qui nous intéressera. L'étude de ces structures et de leurs interactions nous révéleront les composantes et les propriétés, le « quoi » et le « comment », importantes à l'émergence de la cognition.

On retrouve un excellent exemple de ce genre de configurations particulières d'activités dans le concept de « *transiently assembled local neural subsystems* » (TALoNS) développé par Anderson (2014)³². Les TALoNS rompt avec la vision traditionnelle de la spécialisation fonctionnelle où chaque région du cerveau aurait une fonction pour laquelle elle serait spécialisée (pour plus d'information, voir Kanwisher, 2010). En effet, selon Anderson, des assemblées neuronales locales formeraient des coalitions fonctionnelles transitoires dispersées dans le cerveau pour produire une fonction donnée. Ces coalitions émergeraient de l'activité locale de groupes de neurones à différents endroits du cerveau qui ont l'habitude d'être coactifs lors de certains comportements, et seraient également modulées dynamiquement par des

³² Buzsáki (2010) présente une notion similaire à celui du TALoNS en parlant d'assemblage cellulaire et synaptique, excepté que celui-ci ne parlera pas des mécanismes menant aux partenariats fonctionnels observés entre les neurones, comme le fait Anderson, mais de règles syntaxiques dont l'objectif serait de mettre sous silence l'activité des assemblages cellulaires. En un sens, on peut donc dire que les deux chercheurs vont, à partir d'une même idée, dans deux directions opposées; l'un s'intéresse aux mécanismes sous-jacents à l'apparition d'une activité organisée transitoire et l'autre aux mécanismes qui y mettent fin.

facteurs extrinsèques aux neurones, comme l'expression de gènes, les cellules gliales et, plus indirectement, des facteurs corporels et matériels, ce qui ferait que les réseaux d'activités changeraient rapidement.

Une idée importante pour comprendre le cadre dans lequel s'inscrivent les TALoNS est que les assemblées neuronales locales sont pluripotentes, c'est-à-dire qu'elles peuvent donner lieu à plus d'une fonction selon leur organisation, et qu'elles peuvent être recyclées dans des processus différents – ce qu'Anderson nomme la réutilisation neuronale. Les TALoNS reflètent comment il est impossible, à différentes échelles, d'assigner une localisation définitive à un processus, précisément en raison de la dispersion et de la réalisabilité multiple que nous avons évoqué plus tôt.

Clark (2008) propose, pour sa part, la notion de TECS, pour *transient extended cognitive systems*, notion qui capture davantage le type d'interactions qui nous intéressent. Tout comme les TALoNS, les TECS sont des configurations d'activité temporaires entre différentes structures qui sont assemblées pour réaliser une tâche spécifique (ex. une fonction cognitive). Toutefois, à l'instar des TALoNS, les structures agglomérées dans les TECS peuvent être autant des structures neuronales qu'extraneuronales. Pour reprendre les termes de Clark, « [*TECS are*] *transient creations, geared toward a specific purpose (doing the accounts, writing a play, locating a star in the night sky), and combine core neural resources with temporary add-ons such as pen, paper, diagrams, instruments, and so on* » (p.158). Ces ajouts, comme il le dit, correspondent à des « *elements of local cognitive scaffolding* », bien qu'il ne précise pas dans quelle mesure ces éléments échafaudants contribuent à au fonctionnement du système en entier (ex. condition facilitante? élément constituant?).

Les concepts de TALoNS et de TECS résonnent bien avec notre conception de la cognition, et particulièrement avec les idées qui se rapportent au caractère distribué et dynamique des processus cognitifs, mais ils ne capturent pas tout à fait le projet que nous avons en tête pour ce chapitre. Bien que les TALoNS soient compatibles avec ce que nous entendons par « occurrences transitoires d'activité », nous croyons que cet acronyme n'est pas adéquat pour rendre hommage à l'idée de configurations d'activité entre des composantes ontologiquement variées, et ce malgré l'importance qu'accorde Anderson aux composantes extraneuronales. L'utilisation d'un tel acronyme semblerait suggérer que nous limitons ces configurations à des activités neuronales, alors que ce n'est pas le cas. Le concept de TECS remédierait partiellement

à ce problème, mais Clark l'utilise dans l'optique d'un système complet dédié à une fonction/tâche prédéfinie. Notre utilisation des occurrences transitoires d'activité divergent de cette conception en ce qu'elle insiste sur la succession de différentes configurations d'activité alors qu'un processus se développe et se régule dans le temps.

Pour ces raisons, nous croyons qu'il est préférable d'employer un nouvel acronyme pour désigner les occurrences transitoires d'activité, un qui pourrait combiner les intuitions d'Anderson et de Clark et qui pourrait capturer l'apport dynamique des différents types d'échafauds à l'émergence et la régulation des processus cognitifs. Nous allons donc dès-à-présent utiliser l'acronyme TALoPASS pour désigner ces occurrences. TALoPASS signifie *transiently assembled local physical and social subsystems*, et englobe la contribution des échafauds physiques (intra- et extrasomatiques) et sociaux. Il ne faut pas se méprendre, cependant, sur l'utilisation du terme « local », comme il peut prêter à confusion. Le « local » dont il sera question est un endroit relativement circonscrit dans le monde en lien avec la fonction observée, et non un plaidoyer en faveur d'un espace restreint unimodal.

3.1. Une introduction à la science des réseaux

Il est de notre avis que la meilleure façon de capturer et d'investiguer ces TALoPASS est de faire appel, en continuité avec les TALoNS d'Anderson, à une modélisation par réseaux. Les réseaux, vu leur organisation particulière et les techniques et algorithmes qui existent pour visualiser, isoler et simplifier l'information qu'ils représentent (nous allons en parler en plus détail un peu plus tard dans ce chapitre), permettent de révéler commodément les composantes impliquées dans le phénomène étudié ainsi que leurs interactions entre elles. De plus, la modélisation par réseaux, nous permet de dégager avec une plus grande aisance des modifications d'activité – des transitions – à travers le temps. En effet, en représentant la cognition comme un réseau d'activité entre différentes structures, il est possible de voir comment la connectivité entre ces structures peut changer à mesure qu'un état cognitif particulier transitionne vers un autre.

Conformément à Schermerhorn & Cummings (2008), Perone & Simmering (2017), Spivey (2020), et bien d'autres théoriciens, nous pensons que l'intérêt ne réside pas dans un état particulier du système étudié, une *occurrence* prise seule, mais dans les transitions dynamiques

qui s'opèrent entre ces états. Cela signifie, pour notre propos, que l'étude d'une faculté cognitive ne devrait jamais s'articuler autour que d'une seule occurrence transitoire d'activité – ce qui reviendrait en un sens à s'intéresser à un *locus* –, mais de plusieurs. C'est à travers la lentille de ces transitions qu'on pourra saisir les principes organisateurs qui lient et délient, à différents moments, les structures impliquées dans la production et le maintien d'une activité cognitive.

Il est bien de noter, avant de poursuivre, que l'idée de modéliser la cognition via la science des réseaux n'est pas une proposition originale. Plusieurs chercheurs ont fait des suggestions similaires, dont Sutton et al. (2006), Thomas & DaSilva (2009), Zhao et al. (2012), Siew et al. (2019), Smyth (2020) et Poirier & Faucher (2020) pour ne prendre que quelques exemples. Sutton et al. (cité précédemment) proposent, par exemple, un modèle réseautique pour rendre compte des étapes successives par lesquelles un agent en vient à apprendre de ses actions dans le monde via la notion de cycle cognitif. Ce cycle commence par l'observation de l'environnement, l'orientation, la planification, la décision, l'action, pour, enfin, aboutir à un apprentissage basé sur l'interaction de l'agent avec le monde qui orientera celui-ci, à son tour, vers de nouvelles observations, et le début d'un nouveau cycle itératif (voir **Figure 3.1**).

Poirier & Faucher (2020) proposent, quant à eux, un nouveau modèle causal intégré (ICM) pour rendre compte de la variabilité architecturale de la cognition chez l'humain qui repose sur l'idée d'un dialogue inter-domaine entre les théories des sciences cognitives énaactives, incluant les sciences biologiques et neurologiques, et les théories des systèmes développementaux, dialogue qui est représenté sous la forme d'un réseau complexe hiérarchique et modulaire.

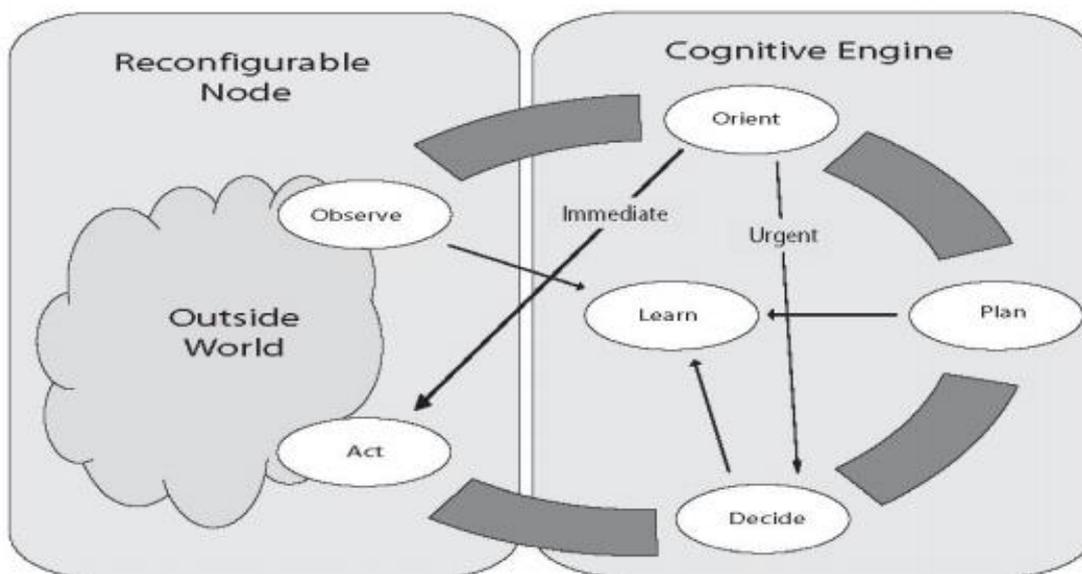


Figure 3.1 - Illustration montrant le cycle cognitif tel qu'il est présenté dans l'article de Sutton et al. (2006).

La proposition qui sera faite ici se démarquera de ces dernières en ce qu'elle tentera d'intégrer en un réseau complexe les différents domaines d'activités et de relations capturés par les échafauds intra-, inter- et extra-somatiques. Cela implique, il en va de soi, que le réseau que nous présenterons sera distribué; il possédera des composantes horizontales importantes décentralisées, des structures qui correspondront aux viscères de l'organisme, aux individus dans le domaine social et aux objets matériels du monde, qui formeront à leur tour des sous-réseaux modulaires avec leurs propres unités et leurs propres modalités d'interactions (sans être pour autant cloisonnés, loin de là). Aussi, ce réseau, bien qu'il sera présenté sous la forme d'instantanées, ne devrait pas être interprété comme une représentation statique, ou pire comme une représentation d'un phénomène statique. L'un des objectifs primordiaux de ce travail est, après tout, de saisir les transitions dynamiques entre les états du système cognitif, les TALoPASS, et non de s'attarder à des états particuliers fixes. En somme, nous tenterons de proposer une modélisation de la cognition sous l'angle d'un réseau complexe distribué et dynamique. Mais avant de précipiter les choses, commençons par introduire quelques notions de base sur la science des réseaux, question d'être sur la même longueur d'onde.

3.1.1. Théorie des graphes et science des réseaux

On peut retracer l'origine de la science des réseaux à la naissance de la théorie des graphes, à qui l'on doit l'apparition à Leonhard Euler, un mathématicien Suisse qui vécut au XVIII^e siècle (Bula et al. 1994; Alexanderson, 2006). Celui-ci proposa, pour la première fois, une réponse définitive au problème des ponts de Königsberg fondée non-pas sur des calculs ardu, mais plutôt sur la position relative des points d'intérêts. Ce problème va comme suit: quatre masses de terre sont séparées par un cours d'eau et reliées par des ponts qui sont au nombre de deux par îlot, à l'exception de deux îlots interconnectés qui en possèdent trois (en incluant le pont qui les lie entre-eux), de sorte que la première île possède un pont qui mène à la suivante et un autre qui mène à la même ou à une autre île, et ainsi de suite pour chacune des îles, pour un total de sept ponts. Le problème, alors, est de déterminer si une personne pourrait passer par chaque pont et retourner au point de départ sans utiliser le même pont plus d'une fois. Pour une représentation visuelle claire de ce problème, vous pouvez consulter la **figure 3.2** ci-dessous.

Afin de résoudre ce problème, Euler proposa de représenter graphiquement les relations spatiales entre les masses de terre et les ponts à l'aide de différentes lettres (A, B, C et D) pour les îlots et de liens pour les ponts, formant ainsi le graphe que nous pouvons voir à la **figure 3.2** à droite. Cette représentation lui permit de visualiser plus aisément les trajectoires possibles entre chaque îlot et de conclure qu'il était impossible, en fait, de faire le trajet complet sans passer deux fois sur le même pont.

L'aspect le plus intéressant du graphe d'Euler, outre que de simplifier le traitement d'un problème mathématique, est que les chemins possibles ne dépendent pas de notre capacité à les examiner et à les identifier (thèse épistémologique), mais qu'ils sont des propriétés encodées dans la structure du graphe elle-même (Barabási, 2014). Ces trajectoires existent toutes parallèlement, ou, du moins, sont toutes des existences possibles liées à l'organisation structurelle du graphe. Appliquée à la question qui nous intéresse, cette idée suggère que les TALoPASS identifiés existeraient réellement, qu'ils sont des occurrences particulières qui tirent leur origine des propriétés organisationnelles et fonctionnelles de la structure dans laquelle ils s'inscrivent.

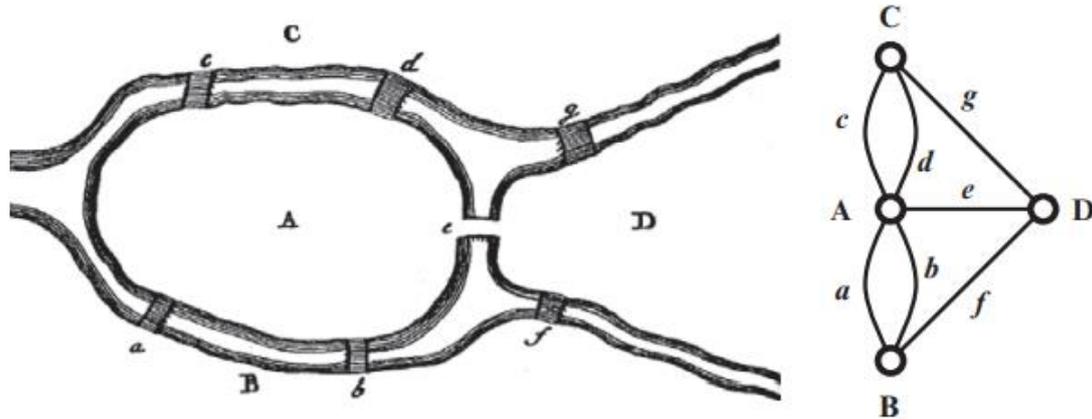


Figure 3.2 - Représentation visuelle du problème du pont de Königsberg tirée des travaux d'Euler (1736). À gauche, un dessin illustrant la position des masses de terre et des ponts qui les lient entre eux. À droite, une représentation graphique du même problème où les masses de terre et les ponts ont été remplacés par des nœuds et des liens respectivement. Source: Barabási, 2014.

Il ne faut pas, en revanche, confondre les graphes avec les réseaux. Malgré leur apparence similaire, le premier fait référence spécifiquement à une modélisation mathématique d'un phénomène donné (Barabási, 2014); un outil épistémologique par lequel on schématise de l'information sans nécessairement capturer celle-ci telle qu'elle se présenterait dans le monde³³. Les unités, dans un graphe, n'ont pas à faire référence à un contenu réel pour être reliées entre elles; elles peuvent l'être en vertu de différentes règles logico-mathématiques. Le réseau, en revanche, fait référence spécifiquement à un système réel. Ces unités décrivent des éléments et des interactions qui existent dans le monde. Par exemple, le réseau social d'un individu représente concrètement les personnes « importantes » au sein de son cercle social et le type de relations que cet individu a avec celles-ci (ex. relation de parenté, d'amitié, de collègue de travail, etc.). Soutenir qu'on peut modéliser la cognition à l'aide d'un réseau, plutôt que par le biais d'un graphe, n'est donc pas sans importance: il s'agit d'une prise de position ontologique à l'égard de celle-ci.

³³ Il faut noter que la distinction qui est proposée ici entre graphes et réseaux est hautement simplifiée, et qu'elle vise simplement à départager deux types de modélisation pour des fins de clarté. Dans les faits, graphes et réseaux se chevauchent très souvent dans la littérature, si bien qu'on parle dans certaines études de *graph neural networks* (Zhou et al. 2020; Liu et al. 2021). Les propos de Newman (2003), par ailleurs, montre comment la ligne qui démarque ces deux types de modélisation est fine: « *mathematics of networks has been driven largely by observations of the properties of actual networks and attempts to model them [...]* (sous-ligne ajoutée) ».

3.1.2. Quelques notions de base sur les types de réseaux

Pour mener à terme notre projet – celui qui consiste en la modélisation réseautique de la cognition via les TALoPASS – il est absolument essentiel que nous présentions quelques notions de base sur les réseaux. Un réseau classique contient deux éléments de nature différente: des nœuds, ce qui correspond aux masses de terre dans le dessin de droite dans la **figure 3.2**, et des liens qui correspondent aux ponts de ce même dessin. Les nœuds peuvent faire référence à toute sorte d'objets dans le monde. Pour un réseau social, comme nous venons de le voir, il s'agit de personnes. Pour un réseau de citations, il s'agit d'articles scientifiques. Pour un réseau neuronal, il s'agit de neurones, de voxels, de ROI (pour *regions of interest*, voir, pour plus d'information, Sun et al. 2019), etc³⁴. On peut penser qu'il existerait au minimum autant de types de nœuds variés qu'il y aurait de réseaux différents dans le monde.

Aussi, un nœud peut représenter un réseau tout entier (ex. le cortex préfrontal dans le réseau du mode par défaut) qui peut, à son tour, être décomposable en différents sous-réseaux, et cela récursivement jusqu'à des nœuds qui ne posséderaient pas de structure réseautique en eux-mêmes (Poirier & Faucher, 2020). À l'un des extrêmes, on pourrait, par exemple, concevoir un réseau microscopique, dans le monde, dont les nœuds seraient si élémentaires qu'ils ne posséderaient pas de sous-unités en interaction (ex. un réseau d'interactions subatomiques), et, à l'autre, un réseau encapsulant immense qui comprendrait l'ensemble des réseaux existants (ex. l'univers). On peut voir la chose de manière analogue à ce que décrivent Kirchhoff et al. (2018) en parlant de la relation entre les couvertures de Markov de la vie, c'est-à-dire comme une succession hiérarchisée de systèmes auto-organisés et auto-différenciés à différentes échelles physiques – biologiques, dans leur cas.

Cela dit, il ne faut pas croire qu'un nœud donné n'appartienne qu'à un seul réseau en tout temps. Au contraire, un même nœud peut appartenir à différents réseaux sur un même plan horizontal (ex. le cortex préfrontal appartient simultanément au réseau du mode par défaut et le réseau exécutif central; voir Chen et al. 2019) et, aussi, à des réseaux de différents ordres hiérarchiques (ex. le cortex préfrontal peut figurer simultanément dans le réseau du mode par défaut, le réseau

³⁴ Barabási, dans son ouvrage de 2014, fait une excellente recension de plusieurs types de réseaux qu'on peut étudier: réseaux internet, réseaux électriques, réseaux de collaboration scientifique, réseaux d'interactions protéiques, etc. Cela ne fait que montrer la variété des types de réseaux existants et, par ricochet, la variété des types de nœuds différents.

cérébral tout entier et – on pourrait le croire – dans des réseaux plus complexes *brain-body-environnement* pour prendre l’expression de Sheya & Smith, 2019).

Dans les prochaines sections, nous nous intéresserons précisément à ce niveau d’analyse – celui du *brain-body-environnement* –, sauf que cet environnement sera divisé en une sphère sociale, intersomatique, et une sphère matérielle, extra-somatique. Il est bien de garder en tête, aussi, que, malgré cette résolution particulière, ce « zoom » sur le phénomène, les nœuds étudiés auront leur propre structure réseautique qui sera tout aussi pertinente à prendre en compte dans l’étude des processus dynamiques derrière la vie mentale. Donc, même si, à une plus grande échelle, le cerveau n’est représenté que par un nœud, il faut se souvenir que ce nœud comprend en lui-même un système d’unités complexe en interaction, et, il en va de même pour les autres nœuds qu’on pourrait identifier dans l’organisme et dans l’environnement de l’individu.

Maintenant que nous avons clarifié la notion de nœuds, on peut parler du deuxième type d’éléments qu’on retrouve dans les réseaux: les liens. Selon la forme qu’ils ont au sein du réseau (ligne continue, pointillée, grasse, fléchée, etc.), les liens font référence à différents types de connexions entre les nœuds; différentes modalités d’interactions pour le dire autrement. Prenons, à titre d’illustration, la **figure 3.3** tirée du texte de Newman (2003). Dans celle-ci, on peut voir 4 réseaux distincts: un premier réseau, en A, qu’on pourrait qualifier de « simple » et de non-dirigé, un deuxième, en B, avec différents types de noeuds et des liens conditionnels, un troisième, en C, avec des noeuds et des liens qui possèdent un poids différent (*weighted*) et, enfin, un quatrième, en D, composé de liens dirigés. Nous allons nous pencher sur chacune de ces variantes dans les paragraphes qui suivent.

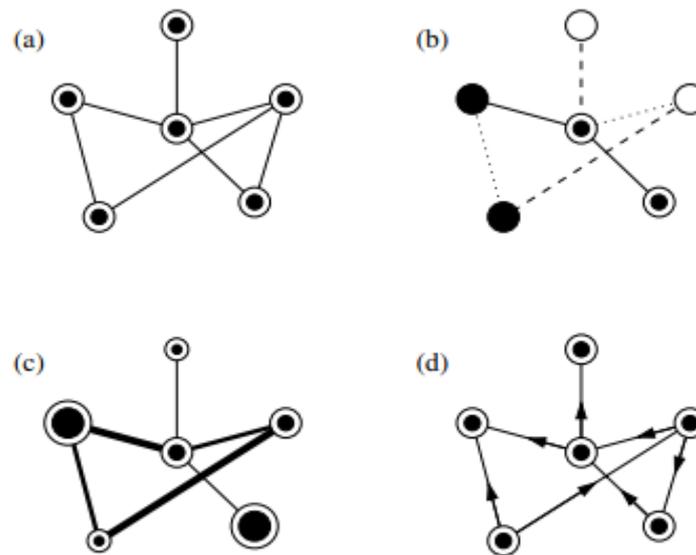


Figure 3.3 – Une illustration montrant une variété de réseaux distincts. En A, un réseau non-dirigé avec un seul type de nœuds et de liens. En B, un réseau illustrant divers nœuds et liens, dont des liens non-dirigés et des liens conditionnels (les lignes pointillées). En C, un réseau constitué de liens et de nœuds possédant des poids différents. En D, un réseau montrant une série de liens dirigés. Source: Newman, 2003.

Ce qui est intéressant avec cette illustration est que chacun des réseaux qui y figurent met en relief un type particulier de liens qui nous sera utile pour modéliser les interactions entre les échafauds de la cognition. Le réseau en A affiche une série de nœuds connectés par des lignes continues qui font référence à des liens non-dirigés, c'est-à-dire des liens qui ne vont pas dans une direction unique, ou, pour le dire autrement, des liens qui désignent une connexion réciproque entre deux nœuds (Barabási, 2014). Dire qu'un nœud x est lié à un nœud y via un lien non-dirigé, c'est donc dire que ces deux nœuds sont coactifs, qu'ils s'influencent mutuellement à une même échelle temporelle. On peut retrouver des exemples de ce genre de lien dans différents domaines: dans des relations de couple où x sort avec y et vice-versa, dans des réseaux cinématographiques où deux acteurs auraient joué dans un même film, ou dans des cas plus complexes d'interactions comme dans ceux qu'on retrouve entre certaines régions cérébrales qui seraient actives en même temps: le cortex préfrontal et le précuneus dans le réseau du mode par défaut par exemple (Utevsky et al. 2014; Chen et al. 2019).

Contrairement, au réseau en A, celui de D présente des lignes fléchées qui renvoient à des liens dirigés unidirectionnels - c'est-à-dire des liens qui ne vont que dans un sens spécifique

déterminé. Ces lignes indiquent que la communication entre les nœuds se fait de manière causale³⁵. Dans ce schéma, l'activation d'un nœud x pourrait influencer l'activité d'un nœud y , mais pas l'inverse. Le nœud y pourrait produire, à son tour, l'activation d'un autre nœud dans le même réseau, ou, s'il se trouve au sommet du graphe, l'activation d'un nœud présent dans un autre réseau. On parlera alors d'un réseau dirigé acyclique si les liens de ce réseau partagent la propriété d'être dirigés et si celui-ci ne boucle pas sur lui-même, et d'un réseau cyclique si celui-ci, au contraire, boucle sur lui-même (Newman, 2003).

On peut retrouver des exemples intéressants de ce genre de liens dirigés au sein du cerveau. Lv et al. (2019) montrent, à ce propos, qu'il existerait des patrons de connectivité dirigée des régions postérieures jusqu'aux portions antérieures du cortex cérébral qui seraient impliqués dans le contrôle moteur derrière la précision et la force de préhension des objets. Liao et al. (2011) parlent même de zones de convergence modulaires, dans le réseau du mode par défaut, qui seraient composées entièrement de liens dirigés, notamment le précuneus, le cortex cingulaire postérieur et le noyau thalamique médio-dorsal.

Il n'existe pas que des liens dirigés unidirectionnels dans les réseaux. Il existe également des liens qui vont dans les deux sens, sans pour autant qu'on puisse les qualifier de non-dirigés. Ces liens, bidirectionnels, indiquent que deux nœuds peuvent causer mutuellement leur activation à des moments différents. À un moment, l'activité de x peut causer celle de y et, à un autre, c'est celle y qui pourrait causer l'activité de x . La transmission du potentiel d'action illustre à merveille ce genre de liens. En temps normal, le potentiel d'action traverse l'axone du neurone pré-synaptique pour se rendre au neurone post-synaptique via les dendrites de celui-ci, mais, dans plusieurs cas, il peut, retourner à rebours vers les dendrites à l'origine de celui-ci dans le neurone post-synaptique. C'est ce qu'on appelle la rétropropagation neuronale (pour plus d'information sur le sujet, voir Stuart et al. 1997). Lorsque nous tenterons de modéliser la cognition, nous utiliserons des flèches à double sens pour représenter ce type de lien et des flèches à sens unique pour représenter les liens dirigés qui seraient unidirectionnels.

³⁵ Parler de causation, ici, est une simplification qu'on s'accorde pour éviter d'alourdir le texte. Dans les faits, les relations entre un nœud X et un nœud Y connectés par un lien dirigé peuvent être beaucoup plus complexes. L'activité de X pourrait faire partie des nombreuses causes qui mèneraient à l'activation de Y (on parlerait alors de causalité multiple). Aussi, il pourrait y avoir des instances où l'activité de X causerait Y , mais que, pour une raison particulière - en raison de l'activité d'autres éléments dans le réseau -, Y ne s'activerait pas (son activité pourrait être inhibée à l'avance ou contrecarrée). Enfin, il n'est pas dit que, parce que l'activité de X cause Y , Y ne peut pas causer X en retour par son activité. Il s'agit du cas des liens dirigés bidirectionnels dont nous parlerons en plus grand détail dans les prochains paragraphes.

Le réseau en B, quant à lui, présente une collection de nœuds de différentes couleurs (ceux en noir, ceux en blanc et les mixtes), en plus de liens qu'on pourrait qualifier de structurels et de fonctionnels, lesquels sont symbolisées par des lignes continues et des lignes pointillées respectivement. Des quatre réseaux qui nous sont présentés dans la **figure 3.3**, celui-ci est très certainement le plus important pour notre propos, car il reflète – de manière simplifiée – la variété des éléments et des interactions qu'on pourrait rencontrer en essayant de modéliser la cognition via les réseaux. En effet, il est difficile de concevoir une modélisation complète de la vie mentale qui serait monolithique, c'est-à-dire qui ne ferait appel qu'à un seul type de nœuds et de liens.

Dans les faits, un modèle satisfaisant des facultés cognitives devrait présenter, au minimum, une combinaison de ces types de nœuds et de liens, en plus de spécifier l'importance relative - le poids - de chacun de ceux-ci dans le réseau étudié (nous en reparlerons avec le réseau en C), pour représenter les différentes propriétés et opérations intra-niveaux et inter-niveaux qu'on peut observer entre les structures qui donneraient « lieu » à la cognition. Ce modèle devrait, à un niveau d'explication particulier, représenter, à l'aide de différentes formes et couleurs, la variété des types de cellules neuronales et gliales qui contribuent à un processus observé au sein du cerveau, et, à l'aide de différentes lignes la directionnalité de ces opérations et les types de connexions qu'il y aurait entre ces cellules (e.g. structurelle, fonctionnelle et effective). De même pour les autres niveaux hiérarchiques d'explications (intégratif, systémique, intra-somatique, etc.).

À ce sujet, il serait bien de prendre quelques instants pour clarifier le sens à accorder aux liens structurels et fonctionnels dont il a été fait mention, comme il s'agira d'éléments importants pour apprécier le caractère dynamique des TALoPASS (surtout le deuxième). Pour commencer, les liens structurels représenteront des instances où deux objets - deux nœuds - seraient *constamment* connectés entre eux. Dans le contexte présent, *constamment* ne signifie pas que le lien est permanent ou qu'il est invariablement présent; il pourrait y avoir des instances où un lien qui est généralement présent dans les diverses opérations d'un réseau disparaît temporairement ou indéfiniment, ou serait réutilisé ailleurs dans celui-ci (c'est notamment le cas à quelques détails près de la réutilisation neurale). Ce qui est important est que ce type de lien nous permette, à titre d'heuristique, de considérer deux nœuds comme étant connectés entre eux la plupart du temps ou sur une période de temps suffisamment longue de sorte à pouvoir en étudier la relation et en extirper un mécanisme.

Les liens fonctionnels, quant à eux, renvoient à des instances où l'activité de nœuds connectés entre eux covarieraient pour une fonction cognitive donnée (ex. on observe que l'activité du lobe occipital varie en fonction de la complexité d'un stimulus visuel), et non qu'il s'agit de connexions probabilistes, bien que nous exploiterons cette idée. Nous pensons que, vu la nature transitoire et inter-substituable du réseau de la cognition, la grande majorité, si ce n'est pas l'entièreté, des liens seraient probabilistes, et donc qu'il n'est pas important, dans notre cas, de réserver un terme pour ce genre de connexions (l'idée de connexions impliqueraient des probabilités d'activation en elles-mêmes).

Les liens fonctionnels peuvent également être conditionnels, c'est-à-dire qu'ils n'apparaîtraient qu'en vertu de la présence de certaines conditions (e.g. l'activation d'une région cérébrale spécifique ou d'un ensemble de neurones, la présence d'un objet particulier dans l'environnement de l'individu, etc.). Elles seront actives à certains moments et, à d'autres, non. Ces liens visent à capturer certaines variétés de connexions dynamiques que nous retrouvons, par exemple, avec le concept de partenariats fonctionnels dont parle Anderson (2014) ou les assemblages cellulaires dont fait mention Buzsáki (2010) et qui sont au cœur de la notion de TALoPASS que nous tenterons d'élaborer ici.

Enfin, le réseau en C affiche une collection de nœuds et de liens de différentes tailles qui renvoient à des poids différents, symbolisés par des ronds noirs plus visibles et des traits gras. Lorsqu'une chose possède un poids plus grand qu'une autre, cela signifie simplement que cette chose joue un rôle plus important que celle-ci quant à la réalisation d'un phénomène observé. Pour un nœud, le poids peut être déterminé par différents facteurs, comme le degré de connexion de celui-ci avec d'autres nœuds, particulièrement la centralité de celui-ci au sein d'un réseau (pour une discussion approfondie de ce facteur, voir Thomas et al. 2015) ou par la force de celui-ci (il peut être hautement influent en vertu d'une propriété intrinsèque qui ne serait pas lié à son degré de connectivité avec d'autres nœuds). Pour un lien, le poids indiquerait une probabilité plus élevée que les deux nœuds qu'il connecte soient coactivés (il serait plus fort). Accorder un poids différencié aux nœuds et aux liens nous permet de creuser la question du « quoi », d'identifier respectivement les objets les plus importants d'un réseau donné, les objets sans lesquels celui-ci ne pourrait maintenir adéquatement son activité ou sa finalité.

Dans la recherche en neurosciences, on rencontre souvent des exemples de réseaux pondérés (traduit de *weighted network* en anglais) dont les éléments possèdent un poids³⁶. Schedlbauer et al. (2014) illustre à l'aide du degré de connectivité des neurones les régions susceptibles d'être les plus importantes dans la récupération de souvenirs spatio-temporels. De manière analogue, Gu et al. (2015) montrent quelles sont les régions du cerveau qui ont le plus d'influence dans le contrôle cognitif et dans les mécanismes permettant la transition d'un état cognitif à un autre. Enfin, Nicolini et al. (2017) ont mis au point une méthode permettant d'identifier les communautés neuronales – c'est-à-dire des regroupements de neurones qui sont plus densément connectés entre eux qu'avec le reste des neurones du réseau – à partir de la connectivité pondérée de certains ensembles de neurones.

3.2. Complexité et problèmes potentiels à modéliser la cognition via une architecture réseautique

3.2.1. Un mot sur les systèmes complexes

Maintenant que nous avons présenté les éléments de base des réseaux et, plus spécifiquement, les types de nœuds et de liens dont nous aurons besoin, nous sommes mieux outillés face à l'épreuve qui se dresse devant nous, celle de modéliser la cognition sans faire appel à des localisations définitives cérébrocentrées. Toutefois, nous ne serons pas prêts à affronter cette épreuve tant et aussi longtemps que nous n'aurons pas confronté certains problèmes inhérents à la modélisation que nous suggérons: les problèmes liés à la complexité du réseau. En effet, représenter l'activité cognitive comme une TALoPASS, une configuration distribuée et transitoire d'activité à travers différents types de structures neuronales et extraneuronales physiques et sociaux, comporte son lot de difficultés: réseaux denses, variés et enchevêtrés, nœuds divers et pondérés, types de connexions différents et co-existants, etc. Nous ne pouvons pas faire l'économie de ces éléments dans une modélisation inclusive qui se veut informative sur les mécanismes contribuant à l'émergence et la régulation de l'activité cognitive.

³⁶ Il est à noter que ces exemples ne sont pas nécessairement des exemples de réseaux tels que nous l'entendons dans la science des réseaux; ils prennent la forme de réseaux (neuronaux), mais les unités à l'intérieur de ceux-ci ne sont pas désignées par la même terminologie qui est employée dans la science des réseaux (e.g. nœuds, liens, etc.). Ceux-ci n'ont également pas exactement la même organisation schématique et ne sont pas non plus étudiés de la même façon. Nous pouvons néanmoins reprendre certaines des idées qu'on retrouve dans les études autour des réseaux neuronaux pondérés dans un modèle réseautique « authentique ».

Lorsque nous parlons de complexité, nous pouvons l'entendre comme le sens commun nous le prescrit, c'est-à-dire comme un objet ou un sujet décomposable en plusieurs éléments, potentiellement hétérogènes, mais la complexité dont nous parlerons dans ce chapitre est – sans vouloir faire un mauvais jeu de mots – plus complexe que cette version plus courante du terme. Nous ferons spécifiquement référence au sens qu'accorde à cette notion l'étude des systèmes complexes. Dans ce domaine, un système est considéré comme complexe lorsqu'il présente un certain nombre de propriétés, comme la présence de parties, d'éléments simples qui composent le système en question, la présence d'interactions non-linéaires entre les composantes, une décentralisation des opérations et l'émergence de comportements (Mitchell, 2010; Sieniutycz, 2020).

Nous pouvons également considérer la complexité d'un objet sous l'angle de l'informatique théorique. De manière non-exhaustive, il existe deux mesures complémentaires de la complexité dans ce domaine: la complexité de Kolmogorov-Solomonoff et la profondeur logique de Bennett. La première s'appuie sur la longueur de l'algorithme nécessaire pour créer un objet donné, dans un programme informatique, afin d'évaluer sa complexité (Uspensky, 2001; Zenil & Tegner, 2014), et la seconde sur le temps requis pour le traiter (Bennett, 1988; Ay et al. 2010). On peut dire que ces deux mesures sont complémentaires, car elles capturent deux aspects opposés de la complexité: la complexité aléatoire (la complexité qui découle de l'agencement stochastique des parties), pour la première, et la complexité organisée (un objet peut être complexe, même si toutes ses parties sont bien structurées), pour la seconde (Delahaye & Vidal, 2016).

Comme il ne s'agit pas d'un exposé en informatique théorique, nous ne creuserons pas davantage la question des différences qui séparent ces deux mesures de la complexité. Seulement, nous pouvons retenir de notre détour dans ce domaine l'idée que l'étude de la complexité ne s'arrête pas aux propriétés susmentionnées, que celle-ci peut prendre différentes formes (e.g. la complexité aléatoire et la complexité organisée), notamment à travers le langage utilisé pour créer des objets en informatique et, de manière analogue, pour décrire des objets existants (e.g. plus la description d'un objet serait longue – sans l'ajout d'éléments superflus – et plus il serait probable que l'objet décrit soit complexe, si on reprend une version altérée de la complexité de Kolmogorov). Pour notre part, cependant, les propriétés que nous avons mentionnées précédemment seront amplement satisfaisantes pour notre analyse, comme elles concernent directement notre tentative de modélisation (particulièrement la question de la

décentralisation et celle de l'émergence). Nous allons nous pencher plus en détail sur ces propriétés dans les prochains paragraphes.

Il existe plusieurs exemples de systèmes complexes dans le monde. L'organisation des fourmilières, des cellules du système immunitaire, du génome humain en sont quelques-uns. Dans ces systèmes, on observe l'émergence d'un comportement complexe qui ne peut pas être retracé individuellement aux parties qui les composent (Mitchell, 2010). Dans le cas du système immunitaire, par exemple, on peut voir la gestion des corps étrangers comme l'aboutissement d'un orchestre complexe d'opérations entre une variété de types de cellules (les leucocytes, les lymphocytes et les phagocytes, parmi d'autres) qui possèdent des fonctions différentes au sein de ce système (Ahmed & Hashish, 2006). Certaines de ces cellules s'occupent de détecter les corps étrangers, d'autres de les neutraliser, et au milieu de ces deux pôles se trouvent une myriade d'autres cellules qui accomplissent des rôles tout aussi importants dans le fonctionnement global de ce système, comme l'activation d'autres cellules du système immunitaire, la facilitation de la communication entre celles-ci, l'isolation des agents pathogènes, etc.

On retrouve clairement à l'intérieur du système immunitaire les différentes propriétés propres aux systèmes complexes (ou du moins ceux qui sont typiquement associés à ceux-ci dans la littérature, voir Mitchell, 2010; Sieniutycz, 2020), soit la présence de parties. Le système immunitaire est décomposable en de nombreuses parties simples: les différents types de cellules dont nous venons de parler. Soit la présence d'interactions non-linéaires entre les composantes du système. L'activité et le produit final du système immunitaire ne peuvent pas être réduits simplement à la sommation de l'activité séparée des cellules; ils sont plutôt la résultante d'interactions dynamiques complexes à plusieurs niveaux spatiaux et temporels entre les parties. Soit une décentralisation des entités et des opérations. L'activité du système immunitaire ne repose pas sur l'activité d'un seul siège de commande où seraient concentrées toutes les familles de cellules immunitaires et où auraient lieu toutes les opérations nécessaires à son bon fonctionnement. Plutôt, les cellules sont dispersées dans le corps humain à différents endroits, comme dans le thymus, la moelle osseuse, la rate, les ganglions lymphatiques, etc. Les opérations sont également décentralisées; elles apparaissent à différents endroits du corps selon l'étape de la gestion des corps étrangers (détection, recrutement, isolation, élimination, nettoyage, etc...). Enfin, de cet orchestre complexe, on observe l'émergence d'un

comportement qui ne pourrait être produit par l'activité simple des composantes individuelles: l'élimination localisée et bien encadrée d'un agent pathogène.

On peut maintenant mieux voir, à la lumière de l'exemple du système immunitaire, comment la cognition peut être considérée comme un système complexe. La cognition, de manière analogue au système immunitaire, comporte une variété de composantes hétérogènes (e.g. structures somatiques, intersomatiques et extrasomatiques) en interaction, et cela à plusieurs niveaux (cellulaires, organiques, systémiques, etc.). Ces composantes sont – comme nous l'avons vu dans les chapitres précédents – décentralisées; il n'y aurait pas un siège unique où se produirait les facultés mentales, comme pourrait le croire un cérébrocentriste, mais plutôt plusieurs centres importants, représentés par les différents types d'échafauds, qui auraient leur fonction propre dans l'émergence de celles-ci. C'est l'interaction dynamique et non-linéaire de ces centres, par la nature des composantes qui les composent et le type d'opérations qui a lieu entre elles et les parties des autres centres, qui donnent lieu aux TALoPASS, des particuliers de l'expérience cognitive, que nous pouvons capturer à l'aide de la modélisation par les réseaux.

3.2.2. Techniques de réduction de la complexité

Si on accepte le postulat ontologique que la cognition est un système complexe au même titre que le système immunitaire et d'autres systèmes qu'on rencontre dans la nature (e.g. les fourmilières) et dans la société en général (e.g. système économique), et que ce système contient non seulement le cerveau, mais les autres organes du corps, des personnes et même des objets extérieurs, il s'ensuit des problèmes épistémologiques auxquels nous devons faire face. L'un d'entre eux pourrait être formulé de la façon suivante: « si la cognition est, en réalité, un agrégat d'unités variées et décentralisées sur plusieurs structures intra-, inter- et extrasomatiques en interaction, sur quoi devrions-nous porter notre attention pour étudier et comprendre les phénomènes qui la sous-tendent? ». Sur chacune de ces structures? Sur une combinaison spécifique de certaines de ces structures à un moment et sur une autre combinaison à un autre moment donné? Ces questions nous rapportent à celle que nous avons proposée au début de ce chapitre: « quelle est l'importance relative de chacun des échafauds et de leurs composantes dans l'émergence et la régulation de la cognition? ». C'est en répondant à cette question que nous réglerons – au moins de manière temporairement satisfaisante – le problème épistémologique de l'identification des structures clefs de la cognition.

La réponse à cette question est déjà, en fait, amorcée. En effet, nous avons vu dans les chapitres précédents que des critères comme celui de la manipulabilité mutuelle peuvent nous aider à identifier les structures susceptibles de contribuer significativement à la réalisation des processus cognitifs. Cela dit, il ne faut pas rejeter d'emblée les structures qui ne satisfassent pas à ce critère en pensant que celles-ci seraient inutiles à la compréhension du phénomène étudié, car l'inclusion de ces structures dans les TALoPASS pourraient fournir des informations additionnelles sur le ou les mécanismes responsables du phénomène. L'ajout de ces structures pourraient également nous permettre de tester ou de consolider de nouveaux critères. Par exemple, si une structure que nous croyons initialement être tertiaire au mécanisme apparaît de manière constante dans plusieurs configurations d'activités disparates, il se pourrait à ce moment-là que cette structure occupe en réalité un rôle fonctionnel plus important au sein du système à l'étude.

L'idée, en présentant des relations MM, était de montrer qu'au moins quelques échafauds peuvent contribuer constitutivement à la cognition, mais pas de limiter l'étude de la cognition à ces structures. Après tout, l'un des objectifs derrière les TALoPASS est précisément celui d'identifier, pour une faculté cognitive spécifique (e.g. le rappel en mémoire épisodique), une configuration distribuée d'activité entre différents types de composantes, une qui serait suffisamment large pour inclure l'ensemble des composantes et mécanismes présumés nécessaires à l'émergence du dit phénomène cognitif, mais aussi assez circonscrite pour ne pas être triviale. Éliminer *a priori* des composantes potentielles de ces configurations pourrait s'avérer problématique dans cette perspective, car on se priverait de certaines relations et de certaines informations sur le fonctionnement du mécanisme. Nous suggérons que ce travail devrait être effectué que lorsqu'on est certain qu'une composante et les relations qu'elle entretient à d'autres sont impertinentes à la compréhension du phénomène ou lorsqu'on connaît *a priori* les effets de son absence dans le mécanisme.

Néanmoins, vu la quantité massive d'information que les TALoPASS contiendraient (ex. les différentes composantes et leur appartenance à certains sous-réseaux, le degré de connectivité individuel entre ces composantes, le poids des connexions entre celles-ci, etc.), un objectif important pour nous sera de trouver un moyen de compresser le réseau excessivement complexe et « indéfini » (au sens où ne connaît pas très bien ses frontières) de la cognition en des configurations individualisées d'activités entre le cerveau et les différents échafauds selon la fonction étudiée, sans couper dans les éléments importants. Cela nous permettra non seulement

de circonscrire des TALoPASS individuels, mais, de manière plus importante, de comparer, un coup bien délimitées, ces configurations entre elles, à travers le temps, pour en dégager les éléments et les opérations (e.g. les connexions) qui resteraient stables et ceux qui changeraient. Cette étude des différences entre les TALoPASS pourrait ultimement nous informer sur les processus dynamiques qui sous-tendent la cognition et la façon dont les processus cognitifs pourraient être interconnectés entre eux.

C'est ici que nous viennent en aide les techniques de réduction de la complexité. Ces techniques permettront précisément de compresser le modèle réseautique de la cognition en diminuant (ex. en priorisant certaines informations) et en simplifiant la lecture de l'information présente dans celui-ci (via des techniques de visualisation des réseaux par exemple). Cela nous permettra de réduire l'opacité inhérente à la quantité massive d'information que ce réseau contient et nous permettra de mettre en relief les TALoPASS intéressants pour les phénomènes cognitifs étudiés. Le reste de cette section va examiner quelques-unes de ces techniques. Nous ne pourrons évidemment pas examiner en détail chacune d'entre elles. Un tel travail nécessiterait des livres entiers. Nous allons nous contenter, seulement, de mentionner l'existence de certaines avenues intéressantes relativement à notre projet et de montrer comment celles-ci pourront faciliter l'identification des TALoPASS.

Nous pouvons commencer par préciser ce que nous entendons par « techniques de réduction de la complexité ». Wade et Heydari (2014) définissent la réduction, dans le cadre de l'étude des systèmes complexes, comme « *the process of removing superfluous elements from the system, either in practice or in implementation, and/or limiting the context under which the system is allowed to operate and reducing the state space to something which is understood* ». C'est dans ce sens que nous comprenons la réduction de la complexité dans un réseau, sauf que nous incluons également dans cette définition les processus visant à modifier la présentation de l'information dans le modèle du réseau - la visualisation de celle-ci - de manière à simplifier le traitement et la compréhension de l'information présente.

Seligman et al. (2015) suggèrent quelques options intéressantes pour réduire la complexité du modèle. Bien qu'elles soient appliquées à la vérification de propriétés formelles sur des plans d'exercices, on peut très bien appliquer ces techniques à la modélisation de la cognition, comme les chercheurs traitent ces plans comme des systèmes complexes (e.g. le plan d'un circuit pour réguler les lumières de signalisation routière). Nous pouvons sommairement diviser les options

proposées par ces derniers en deux catégories: une *underconstraint* et une autre *overconstraint*. Dans le premier cas, il s'agit de réductions qui rendent le modèle plus approximatif en mettant sous silence des parties de celui-ci. Cette catégorie comprend deux options principales: créer des *black boxes*, c'est-à-dire des segments ou des régions du modèle qu'on va délibérément *ignorer* (un sous-réseau ou un agrégat de nœuds interconnectés), ou couper des points (des nœuds ou des liens spécifiques). Dans notre situation, ça pourrait être d'*ignorer* à un moment donné le rôle d'un organe particulier ou d'éléments contextuels qu'on jugerait moins pertinents dans l'émergence de l'activité cognitive analysée. La **figure 3.4**, ci-dessous, donne un aperçu de ces deux options.

Dans le deuxième cas, celui du *overconstraint*, on tenterait plutôt de rendre le modèle plus spécifique. Cela peut se faire de moult façons, notamment en désactivant sélectivement des parties du modèle. Ces parties ne seraient pas ignorées, comme dans le premier cas, mais effacées temporairement, c'est-à-dire qu'elles n'auraient plus d'influence sur les autres composantes du système. Une autre option est d'assigner un intervalle fixe de valeurs à une partie du système ou carrément d'assigner une valeur constante. Dans la modélisation des TALoPASS, l'application de cette option pourrait se traduire de la manière suivante: plutôt que de considérer individuellement, liens par liens, la probabilité qu'un nœud en active un autre qui est connecté à celui-ci, nous pourrions assigner des intervalles qui désigneraient la qualité de la connexion ou la force d'influence d'un premier nœud sur le second. Un intervalle de 80% à 100% inclusivement signifierait, par exemple, que la force « causale » du premier nœud est très forte relativement au second, un entre 60% et 79% serait forte, et ainsi de suite jusqu'à un intervalle dont la borne inférieure serait de 0%.

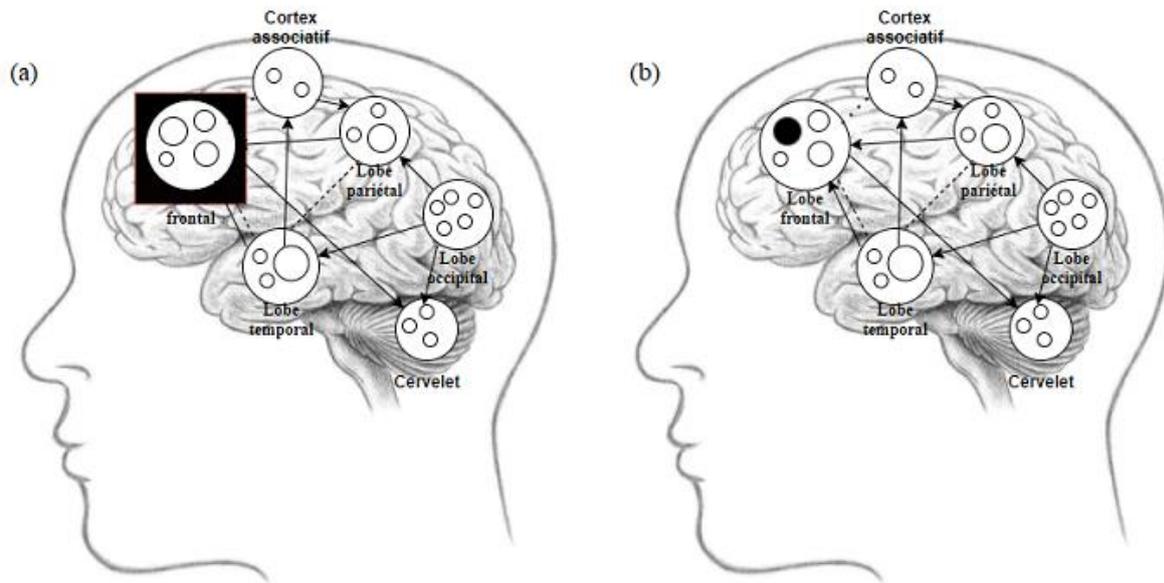


Figure 3.4 - Représentation schématique de ce à quoi ressemblerait l'application des techniques de *black boxing* et de *point cutting* sur des réseaux. Il est à noter que les nœuds et les liens présents dans les deux schémas ci-dessus sont une approximation grossière et ne reflètent en aucun cas la quantité de nœuds qu'on pourrait retrouver au sein des réseaux cérébraux, ni la complexité des interactions possibles entre ceux-ci. Cela étant dit, l'illustration (a) représente l'application d'une procédure de *black boxing*. À noter que la région frontale n'est pas *effacée* du réseau, mais simplement isolée par la boîte noire. On pourrait prendre en compte les interactions que celle-ci aurait avec d'autres nœuds (elle existe encore dans le réseau), mais pas les interactions au sein de celle-ci. L'illustration (b) met en relief le même principe, mais à échelle réduite: on peut avec le *point cutting* ignorer une toute petite section du réseau, en l'occurrence, dans ce cas-ci, un nœud spécifique du lobe frontal. Enfin, il est possible de représenter les techniques d'*overconstraint* d'une manière analogue à celle du *black boxing*, sauf que, dans ce cas-ci, la section encadrée n'exercerait plus aucune influence sur les régions avoisinantes dans le modèle.

Évidemment, ces deux catégories d'options sont accompagnées de problèmes non-négligeables. Mettre sous-silence certaines régions du modèle ou certains éléments spécifiques de celui-ci présupposent déjà une bonne compréhension des interactions qui pourraient émerger à partir d'eux et d'autres sections du modèle. Or, s'il est possible d'identifier des éléments qui ne participeraient pas de manière significative à une fonction cognitive donnée au sein d'une région spécifique, il est beaucoup plus difficile – voire impossible – d'affirmer, en bloquant l'influence de plusieurs régions et éléments individuels du modèle, qu'aucun d'entre eux contribuerait à l'émergence de cette fonction. Au mieux, nous pouvons soutenir qu'il est plausible, en raison des évidences qui nous sont présentement accessibles, que ces parties ignorées ou désactivées ne soient pas essentielles pour comprendre la faculté à l'étude, mais nous ne sommes pas autorisés à conclure qu'elles sont strictement inutiles tant aussi longtemps que nous n'aurons pas investigué l'univers quasi-illimité des interactions que ces parties pourraient avoir avec d'autres éléments du modèle.

Ainsi, en négligeant ou en désactivant certaines parties du réseau de la cognition, nous sommes toujours à risque de créer des gouffres épistémiques dans notre modèle. Il se peut très bien que le ou les phénomènes d'intérêts soient le résultat distal d'une interaction tacite qui aura été masquée par la procédure de mise en boîte noire. Dans d'autres cas, il se pourrait que le phénomène en question soit présent, même si une région jugée essentielle à son émergence aurait été désactivée, simplement parce que la charge computationnelle derrière celui-ci aurait été distribuée sur d'autres éléments du réseau. On observe d'ailleurs un phénomène assez similaire avec la récupération fonctionnelle suivant une lésion cérébrale où la fonction affectée est reprise par d'autres réseaux neuronaux (pour plus d'information, voir Stein, 2017). Enfin, il y a également un risque que notre modèle perde en validité écologique. En effaçant l'influence de certaines parties de notre modèle, on risque de ne plus représenter adéquatement la manière dont les différentes structures dans le monde contribueraient à produire de la cognition.

Mis-à-part les options présentées, il est bien de noter également qu'il existe une grande variété d'algorithmes et de logiciels qui permettent d'organiser les éléments et les opérations présents à l'intérieur d'un modèle réseautique comme le nôtre (McGuffin, 2012 en fait une excellente recension). Quelques exemples de ces algorithmes incluent l'algorithme de Sugiyama-Tagawa-Toda qui permet de repositionner les nœuds désordonnés de manière à former une hiérarchie verticale (voir Nikolov, 2015) – ce qui rend le réseau plus simple à lire, l'heuristique barycentrique qui permet d'ordonner les nœuds d'un réseau sur des lignes horizontales et verticales pour limiter le croisement des liens (Mäkinen & Siirtola, 2005), ou, encore, l'algorithme de Louvain qui permet de regrouper aisément des nœuds en fonction de leur force de connexion entre eux (Blondel et al. 2008).

Du côté des logiciels, il existe une panoplie d'options pour faciliter la visualisation des réseaux³⁷. Ces options sont attrayantes et permettraient vraisemblablement une visualisation

³⁷ Parmi ces logiciels, on peut parler du logiciel en ligne *Flourish* qui permet de créer un modèle réseautique à partir d'une base de données interne (celle-ci peut être créée directement dans le programme ou importée). Ce qui est pratique avec ce logiciel est qu'il est possible de créer différents groupes d'objets dans la base de données pour les nœuds et pour les liens et de leur assigner ensuite une couleur différente pour les différencier dans la représentation graphique du réseau. Non seulement, les objets appartenant à des groupes différents seront différenciés par leur couleur, mais le programme les groupera automatiquement ensemble, de sorte qu'ils ne soient pas éparpillés dans le réseau (pour plus d'information, vous pouvez consulter la page: [Make interactive network visualizations without coding](#)). Le logiciel *Plotly* est également une option intéressante pour organiser et simplifier la représentation de l'information au sein d'un réseau. *Plotly* est un peu plus complexe que son prédécesseur, *Flourish*, comme il requiert de créer manuellement les nœuds, les liens ainsi que leurs propriétés à partir du langage de programmation *Python*, mais il a l'avantage potentiel de ne pas être limité quant au nombre d'éléments à représenter dans le réseau.

organisée et accessible d'un réseau complexe comme celui dont on postule l'existence ici – à condition bien sûr de cibler un bon niveau d'analyse pour les différentes sections du modèle (ex. au niveau systémique pour des organes comme le cœur, au niveau intégratif pour le cerveau, etc.). Toutefois, elles ne seront pas retenues, car notre but, dans ce mémoire, n'est pas de tracer le réseau complet de la cognition – est-ce même une tâche possible? – ni même, en fait, de dresser un réseau si complexe qu'il nécessiterait ce genre de techniques pour être compréhensible. Une telle entreprise ne saurait être l'objectif d'une seule personne, dans un ouvrage, ni même d'un groupe de personnes restreint, mais d'une discipline à part entière, une qu'on pourrait qualifier de « neurosciences cognitives réseautiques étendues » où un regroupement de chercheurs et de théoriciens de plusieurs orientations différentes (biologie, psychologie, informatique cognitive, philosophie, etc.) pourraient se consacrer à la délimitation du réseau complexe de la cognition. L'objectif, en parlant de ces techniques ici, était simplement de montrer comment la complexité n'est pas un obstacle insurmontable à une modélisation de la cognition par réseau, qu'il est possible d'atteindre un niveau satisfaisant d'inclusivité (un niveau d'information satisfaisant) à l'égard des composantes et des interactions présentes dans le réseau sans pour autant compromettre son intelligibilité.

Enfin, nous parlions, plus tôt, « d'identifier les éléments clés de la cognition ». L'idée n'est pas de présenter ici une liste exhaustive de ces éléments, mais de montrer surtout comment la notion de TALoPASS, appliqué à un certain nombre de composantes importantes (cérébrales et extra-cérébrales), pourrait mettre en relief les organisations d'activité temporaire entre celles-ci (le quoi) et les transitions dynamiques qui existeraient entre ces organisations au sein du réseau de la cognition. Cela signifie évidemment que l'identification et l'inclusion des composantes de la cognition dans un réseau est une démarche importante préalable à l'application des TALoPASS. Il faut alors identifier un nombre n d'éléments jugés pertinents dans le réseau – le plus possible en fonction des composantes identifiées dans la littérature –, identifier les degrés de connectivité entre ces éléments (les liens qui les unissent), et, enfin, la force de connexion entre ceux-ci. Cela peut être accompli, en partie, en appliquant comme nous l'avons fait dans le chapitre précédent le critère de démarcation MM.

Cependant, il n'est pas faux d'affirmer que les TALoPASS contribueraient également à identifier des composantes à leur tour. La modélisation de la cognition par réseau nous donne un aperçu particulier des interactions entre certains nœuds qui pourraient surgir comme des nouveaux éléments pour la fonction observée. Les opercules frontaux et l'aire préfrontale

gauche impliquées présumément dans la récupération de l'information en mémoire (voir chapitre 1) sont individuellement des composantes du réseau de cette fonction cognitive, mais, ensemble, si nous prouvons qu'elles sont coactivées la majorité du temps d'une manière significativement forte, elles pourraient, avec leurs connexions, être considérées comme une nouvelle composante à part entière du réseau. Ce groupement pourrait faciliter la représentation visuelle des modélisations futures, si, évidemment, le travail est bien fait. Aussi, l'identification de configurations d'activité dans le réseau de la cognition pourrait en retour servir de critère pour élaborer de nouvelles divisions des facultés mentales – répondre au problème de la granularisation – en fonction des éléments ou des activités qui restent stables à travers le début présumé du processus observé et le moment où il se terminerait. Si la configuration change du tout au tout dans cet intervalle de temps, il est possible que nous ayons à faire en réalité à une situation où une fonction cognitive a transitionné vers une autre.

Dans un scénario idéal, il faudrait que le réseau créé soit le plus complet possible pour rentabiliser l'utilisation des TALoPASS – pour les rendre plus informatifs. Toutefois, la décision d'incorporer un nombre n d'éléments ou une quelconque autre propriété spécifique, comme les degrés de connectivité et les forces de connexion précises, dans le modèle réseautique de la cognition repose ultimement sur la discrétion du modélisateur en fonction de ses objectifs de recherche et les connaissances à sa disposition. Il se peut qu'un élément soit jugé d'emblée insignifiant au réseau et qu'il ne soit pas incorporé dans celui-ci. C'est un risque que le modélisateur doit prendre en connaissance de cause, car l'information qu'on peut retirer des TALoPASS est proportionnelle à la quantité de composantes et d'interactions qu'il incorporera dans le modèle réseautique de la cognition.

Dans notre cas, cependant, nous allons présenter un modèle simplifié de ce à quoi ressembleraient les TALoPASS à l'aide de l'exemple de la récupération de l'information en mémoire épisodique élaboré dans le premier chapitre. Ce modèle, vu les limites prescrites de ce travail, ne pourra pas être élaboré en détail. Il sera à bien des égards sélectif et approximatif – surtout quant aux connexions qui existeraient entre les différentes structures à l'intérieur de celui-ci. Le but, nous le rappelons, n'est pas de reproduire exhaustivement le réseau de la cognition ici, mais de montrer comment les TALoPASS pourraient servir à identifier des organisations d'activité particulière entre des structures variées ontologiquement et à apprécier les transitions qui auraient lieu entre ces organisations.

La prochaine section sera consacrée à la présentation de l'exemple de la récupération de l'information en mémoire épisodique. Cet exemple nous permettra de mettre à l'épreuve la notion de TALoPASS, mais, plus important encore, il nous donnera l'occasion de contraster la version échafaudée de la récupération de l'information en mémoire avec la version cérébrocentriste présentée dans le chapitre 1.

3.3. Intégration: le cas de la récupération de l'information en mémoire épisodique

Nous avons utilisé l'exemple de la récupération de l'information (épisodique et sémantique) à différents moments de ce mémoire. Dans le chapitre 1, pour illustrer certains problèmes au sujet du localisationnisme: frontières du cognitif, taxonomie du mental, niveau de spécificité, etc. Nous avons continué à travailler cet exemple, dans le chapitre 2, avec les notions de systèmes de mémoire transactionnelle et d'implémentation d'intention. Ce fut l'occasion de mettre en relief la contribution des échafauds inter- et extrasomatiques à cette faculté. Dans cette section, nous poursuivrons et achèverons ce travail avec l'application de la notion de TALoPASS. Nous montrerons visuellement, à l'aide d'un exemple détaillé, quel genre de configurations transitoires d'activités pourraient apparaître tout au long de la réalisation de ce processus.

Commençons, d'abord, par présenter le schéma cérébral proposé traditionnellement par les neurosciences cognitives pour la récupération de l'information en mémoire épisodique. Plusieurs régions du cerveau contribueraient à différents degrés à cette fonction: les opercules frontaux bilatéraux, le lobe préfrontal gauche, l'hippocampe, le cortex périrhinal, le gyrus denté, et de manière large les aires associatives du cortex et le lobe pariétal (Dickerson & Eichenbaum, 2010; Rugg & Vilberg, 2013; Addis et al. 2015). En effet, il est présumé que les souvenirs épisodiques sont entreposés parcimonieusement dans le cortex associatif selon les modalités sensorielles concernées par le souvenir (Hasegawa & Miyashita, 1999; Rosen et al. 2018) et que le lobe pariétal contribuerait au moins de manière partiel à l'attention dirigée lors de la récupération des souvenirs. La **figure 3.5** ci-dessous donne un aperçu approximatif de ce à quoi pourrait ressembler le réseau cérébral de cette faculté. À noter qu'il s'agit d'un schéma excessivement simplifié.

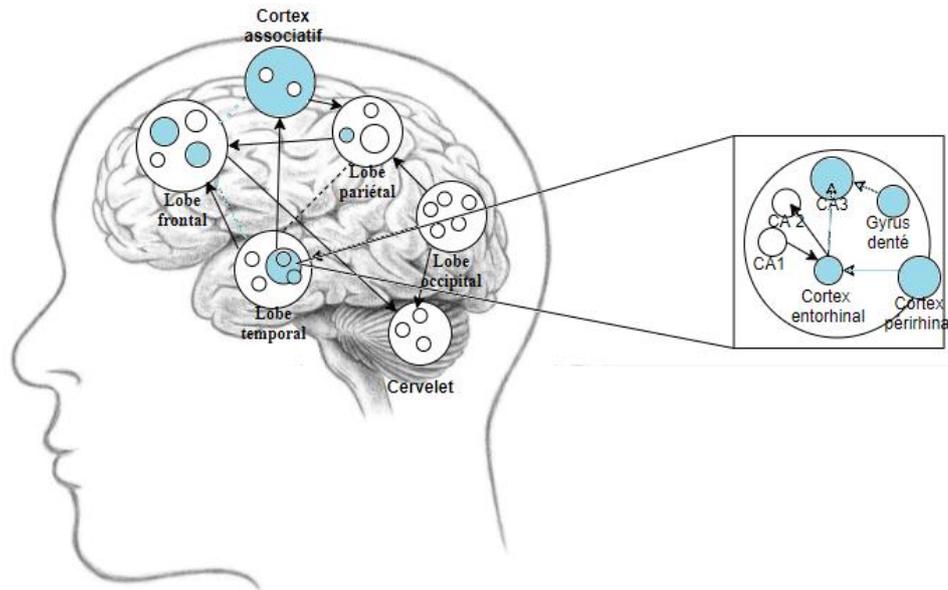


Figure 3.5 - Représentation visuelle³⁸ extrêmement simplifiée d'un réseau cérébral pour le processus de récupération de l'information en mémoire épisodique. Au niveau du lobe frontal, nous pouvons apercevoir, en bleu, deux nœuds qui correspondent respectivement aux opercules frontaux et à la région préfrontale. Il est à noter, bien qu'il ne soit pas indiqué dans le schéma, qu'il existe des relations entre ces structures qui auraient pu être représentées à leur tour par des liens, mais, comme nous manquons d'espace et que l'objectif n'est pas de représenter complètement le réseau pour cette fonction, ils ont été omis ici. Au niveau du lobe temporal, nous voyons l'hippocampe (le grand nœud en bleu) et, à l'intérieur de celui-ci, le gyrus denté (le nœud septentrional). Le nœud bleu qui borde l'hippocampe correspond au cortex périrhinal qui est normalement adjacent à cette structure. Nous voyons plus clairement, dans le carré de droite, certaines des interactions entre ces structures et le cortex entorhinal qui fait le pont entre l'activité de celles-ci.

En ce qui concerne les différents échafauds, commençons avec les éléments qu'on pourrait retrouver dans la section du réseau dédié à l'échafaud intrasomatique. Nous l'avons mentionné dans les chapitres précédents, différents éléments de l'organisme, allant de l'appareil sensorimoteur jusqu'aux viscères, participent à la récupération de l'information en mémoire épisodique. Des éléments proprioceptifs comme la posture d'un individu peuvent faciliter considérablement – voire même provoquer – la récupération d'un souvenir selon que la posture adoptée au moment du rappel soit congruente avec celle adoptée dans la scène faisant l'objet du rappel (Dijkstra, 2007). Il en va de même pour des viscères comme le cœur et les poumons. Le cœur, par exemple, peut contribuer à la fois directement (Sidharta et al. 2013) et indirectement – via les émotions – au rappel de l'information. De même pour les poumons, via

³⁸ Il est important de mentionner que les schémas présentés ici ressemblent à ceux qu'on pourrait retrouver en neurosciences cognitives traditionnelles, où ils illustrent des hypothèses (empiriquement testées) concernant les interactions entre les diverses régions cérébrales lors de l'exécution de tâches, mais que, tels que nous les concevons ici, ces schémas émergeraient plutôt de l'analyse statistique des covariations d'activités entre les différentes régions cérébrales, corporelles et environnementales lors de l'exécution d'une tâche cognitive donnée.

les rythmes respiratoires (Zelano et al. 2016), et pour des espaces intra-somatiques comme le microbiote intestinal (Heyck & Ibarra, 2019; Benmelouka et al. 2019).

Du côté de l'échafaud intersomatique, nous avons vu dans le chapitre précédent, avec les systèmes de mémoire transactionnelle, que des systèmes de mémoire transitoire peuvent émerger à partir des procédures collaboratives entre les individus d'un même groupe. Qui plus est, la performance de ces systèmes serait supérieure à celle des membres de ce groupe prise individuellement, à condition qu'il existe une relation assez forte et continue entre ces derniers. Harris et al. (2014) abondent dans ce sens lorsqu'ils parlent du rappel partagé entre les partenaires d'un couple; c'est l'habileté à communiquer efficacement entre les partenaires et la complémentarité des informations dont ils disposent qui facilitent la récupération des souvenirs partagés. Dans la même veine, la capacité des parents de poser des bonnes questions lorsqu'ils discutent avec leurs jeunes enfants de souvenirs jouent un rôle majeur dans la capacité de ces derniers à récupérer effectivement l'information épisodique et autobiographique et contribue également au développement de cette faculté chez eux (Hayne & Herbert, 2004). Les enfants viendraient à internaliser le style narratif de leurs parents – le genre de récits qu'ils construisent en posant des questions sur les souvenirs – et l'appliqueraient quand viendrait ensuite le temps de récupérer un souvenir.

Enfin, nous avons vu également dans le chapitre précédent comment l'échafaud extrasomatique peut soutenir de différentes façons la récupération de souvenirs (Metzger, 2002; Augustinack et al. 2014). La configuration de l'environnement matériel de l'individu peut indiquer ou provoquer la récupération de l'information en mémoire épisodique indépendamment de sa volonté ou d'une activation de son lobe frontal - dont la tâche serait d'initier la récupération. C'est, comme nous l'avons vu, le cas des stratégies qui visent à implémenter des intentions dans l'environnement immédiat de l'individu. Ces intentions sont rattachées à des objets spécifiques qui assument la fonction d'initier le rappel. Une personne craignant de ne pas se souvenir de remettre sa dissertation pourrait, par exemple, laisser traîner sa copie sur sa table de travail dans l'espoir qu'elle lui rappelle ultérieurement son obligation de la remettre. Dans cette perspective, cette copie remplacerait fonctionnellement les régions du lobe frontal impliqué dans la récupération de l'information.

Bref, nous voyons que chacun des échafauds contribue à sa manière au processus de récupération des souvenirs, que ce soit simplement en participant, en facilitant ou carrément en

assumant la charge computationnelle derrière l'émergence de cette faculté. Cela étant dit, il ne faut pas conclure, comme nous l'avons mentionné précédemment, que les échafauds contribuent tous simultanément et au même degré à l'émergence du phénomène observé. Il se peut fort bien qu'un seul échafaud – jumelé avec le cerveau – soit suffisant, à un moment donné, pour produire la faculté d'intérêt. C'est le cas notamment de certaines *affordances* offertes par un objet du monde extérieur et la récupération de l'information en mémoire. À d'autres moments, il est tout aussi possible que les différents échafauds contribuent de concert pour produire cette même faculté. Ultimement, il est de notre avis que le phénomène observé pourrait être expliqué, à différents moments, par des configurations d'activité variées entre les composantes des échafauds intra-, inter- et extrasomatiques. La **figure 3.6** ci-dessous montre un réseau incorporant les différents échafauds discutés jusqu'à présent. À noter qu'il ne présente pas – encore – de TALoPASS.

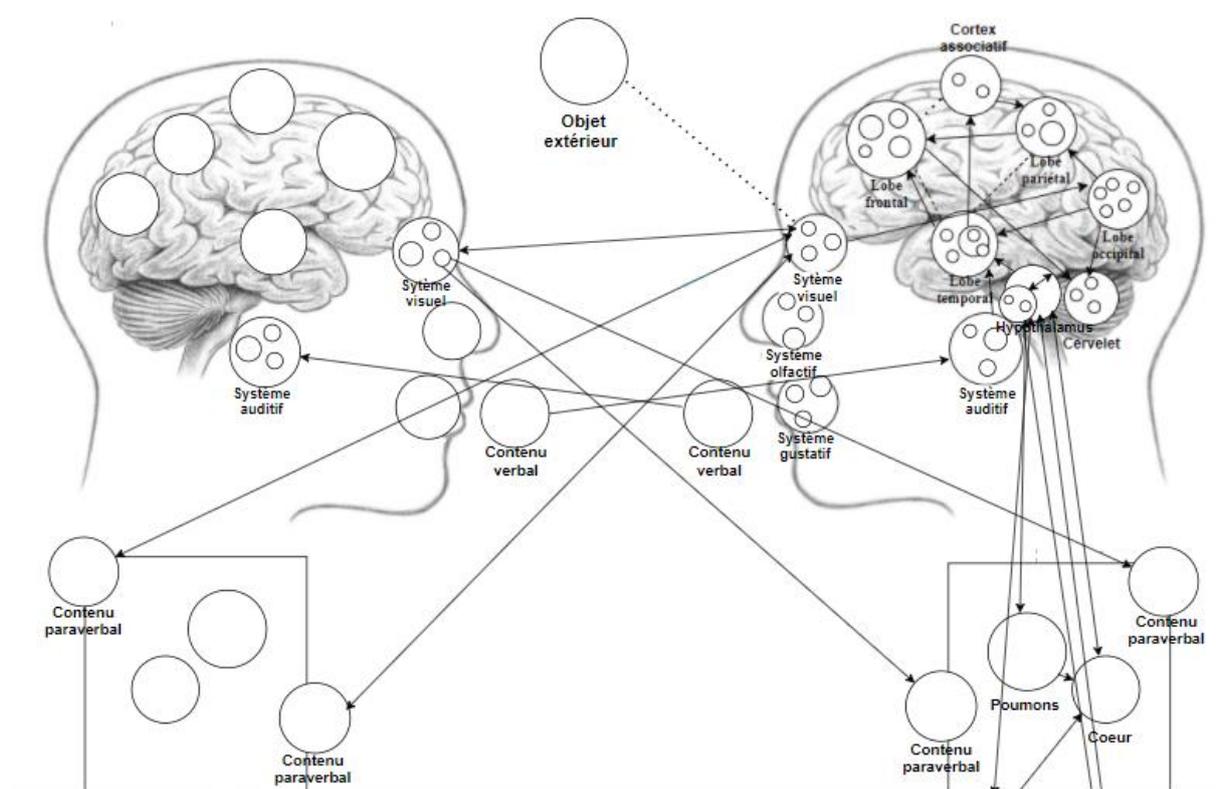


Figure 3.6 - Schéma d'un réseau cognitif qui incorporerait les échafauds intra-, inter- et extrasomatiques. Ce schéma n'est qu'un croquis et ne reflète pas la complexité réelle des interactions possibles entre les différents échafauds et le cerveau. Par exemple, il pourrait y avoir simultanément différents interlocuteurs et/ou différents objets qui stimuleraient l'agent via différents canaux sensoriels. L'important ici est simplement de donner un aperçu concret de l'allure générale d'un réseau de la cognition pour appuyer notre propos.

Cela étant dit, prenons les quelques pages restantes à ce chapitre pour montrer un exemple concret – de la vie de tous les jours – du genre de relation qu'il pourrait y avoir entre les

différentes structures participantes à la récupération de l'information en mémoire épisodique et les transitions dynamiques qu'il pourrait y avoir entre elles. Cela nous permettra de montrer, du même coup, comment il est possible de modéliser ces relations sous la forme de TALoPASS.

Prenons, pour les fins de cet exemple, le cas fictif de Guillaume. Celui-ci parle avec sa conjointe Sophie de leur voyage qui a eu lieu il y a cinq ans. Certaines régions de son lobe frontal s'activent, dont les opercules frontaux bilatéraux et le cortex préfrontal gauche, pour initier la récupération de certaines informations sur un événement cocasse qui se serait déroulé pendant leur séjour à l'hôtel. Échec total, il a bien beau se gratter le crâne et se donner des indices verbaux (ex. en se posant des questions comme « est-ce que c'était dans la chambre? dans le corridor? »), mais il n'arrive tout simplement pas à se souvenir du dit événement. Il se tourne la tête, dans une tentative désespérée (aussi, parce que le regard inquisiteur de Sophie fait pression sur lui), et, soudainement, il voit à la télévision un bâtiment rustique qui partage des similarités architecturales frappantes avec l'hôtel où Sophie et lui étaient hébergées. L'image du bâtiment lui revient quasi-instantanément à l'esprit et aussitôt il se souvient de l'ascenseur qui s'était arrêté entre deux étages pendant une heure (l'événement cocasse).

On l'aura compris, certaines régions du cerveau de Guillaume se sont activées lors de la tentative de la récupération de l'information en mémoire, mais, ultimement, c'est une composante appartenant à l'échafaud extra-somatique qui lui a permis de récupérer l'information au sujet de l'événement cocasse. Ici, le TALoPASS comprendra des interactions entre des éléments du cerveau, comme les aires sensorielles qui ont traité les informations venant de la télévision, mais aussi des régions comme l'hippocampe et le cortex périrhinal, chargées des processus de familiarité et de recollection, et certaines zones du cortex associatif. Le TALoPASS en question comprendra également des interactions avec des composantes de l'échafaud extrasomatique, en l'occurrence, dans cet exemple, la télévision. **La figure 3.7** illustre schématiquement le genre de configuration d'activité qui aurait lieu à ce moment du processus de récupération de l'information.

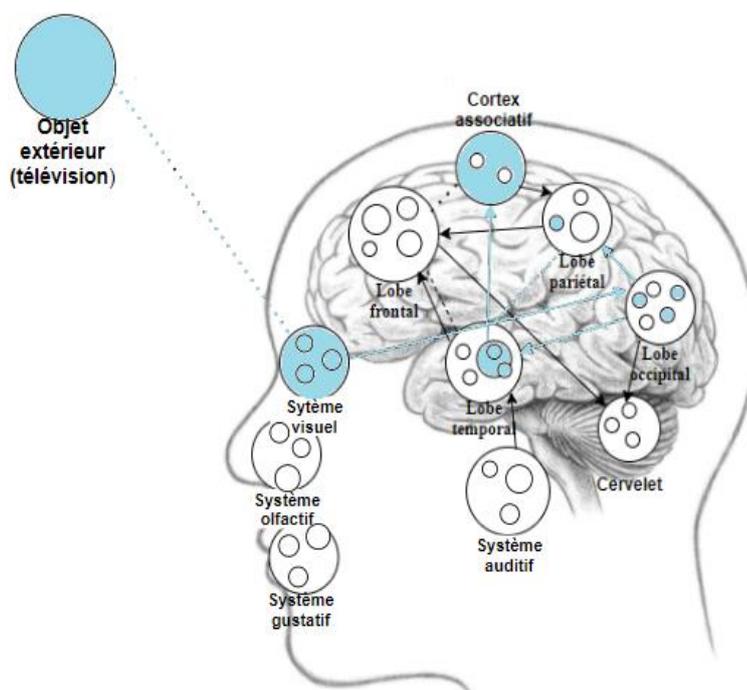


Figure 3.7 - Représentation visuelle schématique d'un TALoPASS qui émergerait de l'interaction entre une composante extra-somatique, la télévision, et du réseau cérébral humain. La plupart des régions cérébrales activées normalement lors de la récupération de l'information en mémoire épisodique sont actives, mis-à-part celles du lobe frontal (les opercules frontaux et le cortex préfrontal gauche), car les images du bâtiment à la télévision se substituerait fonctionnellement à ces régions. Le réseau cérébral comprend maintenant différents nœuds pour représenter les sens par lesquels les objets du monde extérieur directement accessibles *afforderaient* des possibilités motrices et cognitives. Dans l'exemple de la télévision, c'est la vue du bâtiment qui a provoqué le rappel du souvenir chez Guillaume, et c'est donc le réseau de la vision - au niveau de l'œil et du nerf optique - qui est concerné. L'information sera traitée dans différentes aires du cortex visuel et ferait l'objet d'un traitement parallèle dans les voies ventrales et dorsales du système visuelle pour déterminer le « quoi » et la localisation du stimulus dans l'espace - représentée par les flèches bleues en provenance du lobe occipital. À partir de là, l'information désirée sera récupérée normalement dans les aires associatives. Évidemment, un réseau « complet » aurait présenté tous les objets de l'environnement immédiat de Guillaume susceptibles d'indicer ou de provoquer la récupération du souvenir en question, que ce soit par le biais du système visuel, olfactif, auditif ou autre.

En pensant au dit ascenseur, Guillaume éprouve un léger malaise. Il sent son rythme cardiaque et respiratoire augmenter. L'événement lui revient en tête avec une telle vivacité, comme s'il vivait à nouveau la détresse ressentie lorsqu'il fut coincé entre les quatre murs de l'ascenseur. Ce qui n'était qu'une représentation très vague de l'événement devient soudainement plus clair. Il se souvient maintenant des murs verdâtres de l'appareil et du miroir brisé derrière lui. Le décor de cet appareil, se souvient-il, était de mauvais goût et trahissait la qualité plutôt médiocre de l'hôtel où il séjournait avec sa conjointe. On voit ici que les émotions découlant de l'activité physiologique des viscères de Guillaume ont facilité la récupération d'informations spécifiques (Tyng et al. 2017) comme les détails de l'ascenseur qui ont été au moment de l'expérience associés avec des émotions fortes. Le TALoPASS, ici, change d'une configuration qui encapsule des composantes extra-somatiques (ce ne sont pas les objets au moment présent qui

influencent sa récupération) à une configuration qui comprend maintenant des éléments intrasomatiques comme le cœur et les poumons. Même le réseau cérébral change un peu pour inclure, en plus des régions déjà impliquées dans le réseau de la récupération de l'information en mémoire épisodique, des régions comme l'hypothalamus et l'amygdale, parmi d'autres. La **figure 3.8** met en relief cette nouvelle configuration.

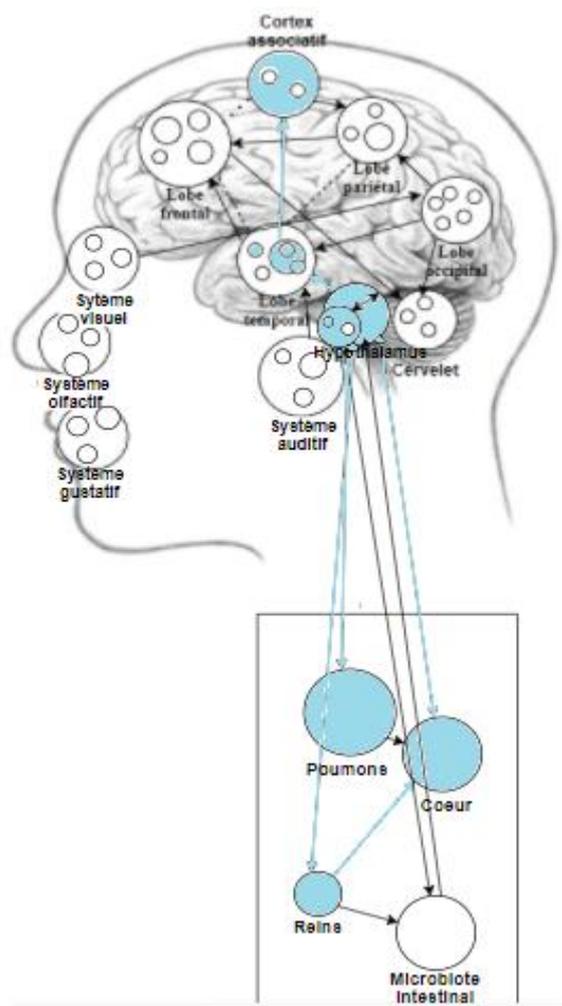


Figure 3.8 - Représentation visuelle du TALoPASS hypothétique présent au moment du rappel d'informations spécifiques associé aux émotions fortes vécues par Guillaume qui incluerait des composantes de l'échafaud intrasomatique. On observe différentes connexions (excitation et rétroinhibition) entre le cœur, les poumons et l'hypothalamus. Celui-ci est bordé par l'hypophyse qui comprend deux sous-composantes: la neurohypophyse à droite et l'adénohypophyse à gauche. Seule cette dernière a été retenue dans le TALoPASS vu son rôle direct dans la sécrétion d'hormones responsables des activités physiologiques nécessaires à certaines émotions (Watts, 2007), comme dans le cas de l'axe hypothalamo-hypophyséo-surrénalien et des émotions négatives. Cet axe est représenté par la flèche bleue qui lie l'hypothalamus aux reins. L'activité coordonnée de ces viscères contribuerait, dans une perspective de boucle corporelle, à produire des hormones et des marqueurs somatiques (voir chapitre 2) susceptibles d'influencer la récupération de l'information en mémoire épisodique via des émotions fortes (Tyng et al. 2017).

Sa conjointe Sophie commente l'événement en riant. « Te souviens-tu de la réaction des secouristes? ». Guillaume ne comprend pas immédiatement pourquoi cette dernière rit ainsi, lui qui ressent encore un certain malaise vis-à-vis dudit événement, mais le commentaire de sa conjointe finit par lui rappeler enfin le caractère cocasse de la situation dans laquelle ils se trouvèrent. Le couple venait tout juste de prendre un bain de minuit avant de regagner l'ascenseur et n'avait sur eux que des serviettes en guise de recouvrement. Les secouristes avaient dévisagé longuement le couple avec des airs incrédules. L'état émotionnel de Guillaume passe du malaise au soulagement. Il rit maintenant aux éclats. Dans ce cas-ci, on peut dire que l'information manquante n'était pas localisée dans le cortex associatif de Guillaume, mais dans celui de Sophie. Toutefois, il ne s'agit pas d'une réduction à l'activité cérébrale seule de deux cerveaux en synchronisation, mais d'une expérience intersubjective d'interaction sociale habile entre deux personnes où plusieurs éléments participent au rappel de l'information: la prosodie particulière de Sophie, les mots précis qu'elle emploie pour décrire la situation, ses mouvements corporels, etc. La **figure 3.9** illustre les nouvelles composantes et interactions qui ont lieu à ce moment-ci du rappel.

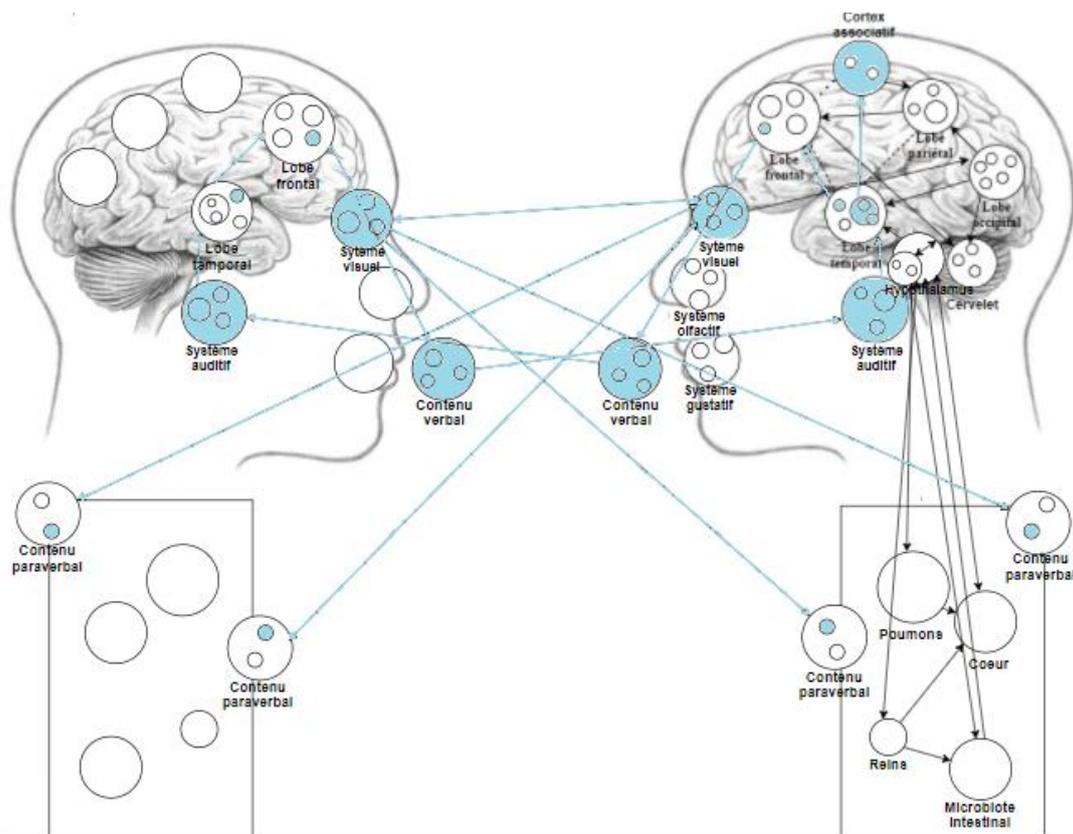


Figure 3.9 - Schéma illustrant un TALoPASS émergent des interactions avec un échafaud intersomatique. Sophie, dans le cas de Guillaume, peut indiquer ou provoquer le rappel d'informations spécifiques via du contenu verbal (e.g. mots spécifiques, débit, prosodie, etc.), du contenu paraverbal (ex. la posture, les mouvements corporels à différents niveaux du corps, soit les épaules ou les bras, parmi d'autres) ou par le biais d'émotions qui peuvent s'exprimer à travers ces deux types de contenu (ex. à travers la prosodie ou la posture). Un réseau plus complet aurait pu inclure différentes régions du cerveau de Sophie, comme les régions qui participent aux processus émotionnels présents (ex. hypothalamus, insula, amygdale, etc.) et aux processus mnésiques (ex. hippocampe, cortex périrhinal, cortex associatif, etc.), mais il est davantage question ici des composantes contribuant directement à l'interaction sociale habile entre les deux (contenu verbal, paraverbal et émotions) que les processus qui sous-tendent ceux-ci.

Qu'est-ce que nous pouvons retenir de cet exemple? Un premier point est qu'il serait faux de dire que la récupération du souvenir, dans le cas de Guillaume, est localisée de manière rigide dans des régions x et y du cerveau ou dans des lieux extérieurs précis. La performance finale, le rappel complet de l'événement, est la résultante complexe de différentes combinaisons d'éléments et d'activités à plusieurs moments du processus de rappel qui sont mis en relief par l'étude des transitions, des différences entre ces différentes configurations. Deuxièmement, et faisant suite à un point que nous avons abordé dans le premier chapitre au sujet de la spécificité des fonctions cognitives, les différences d'activité pourraient traduire en réalité l'apparition de processus différents là où nous croyions initialement qu'un seul processus était à l'œuvre. Le cas de Guillaume illustre bien cette idée avec l'évolution des TALoPASS. Au début, la première configuration comprend certaines régions du cerveau et la télévision présente dans son appartement. Puis, la deuxième, des viscères comme le cœur et les poumons et des nouvelles

régions du cerveau (l'hypothalamus et l'amygdale). Enfin, la dernière configuration montre des nouvelles relations jaillir de l'interaction entre les régions cérébrales de Guillaume et de Sophie (pour les souvenirs complémentaires) qui passent par des activités verbales et paraverbales.

À partir de ces configurations dissemblables, on pourrait postuler l'existence de trois sous-types de récupération de l'information épisodique qui pourraient être considérées comme des facultés séparées, une locale qui concernerait les caractéristiques de l'environnement dans le souvenir, une émotionnelle – les émotions ressenties – et une sociale qui concernerait le caractère interpersonnel du souvenir (quel était la relation entre les protagonistes dans la scène qui fait l'objet du rappel?). Les transitions entre les TALoPASS pourraient ainsi servir de critère pour saisir les relations existantes entre certaines fonctions cognitives et nous permettre de mieux les répertorier en fonction de la participation de certaines composantes ou la présence de certaines activités à la configuration observée. Ce deuxième point renforce l'idée que nous devons étudier les composantes de la cognition – le quoi –, les activités entre elles, et surtout les transitions entre ces activités pour élargir notre compréhension du cognitif et des facultés qui méritent l'appellation de cognitives.

Conclusion

Dans ce chapitre, nous avons vu comment il est possible d'investiguer la cognition scientifiquement sans faire appel à des localisations définitives exactes. Certes, nous pourrions, si l'envie nous prend, identifier des *loci* temporaires pour un phénomène donné, comme les opercules frontaux bilatéraux ou la télévision dont il a été question dans l'exemple de la récupération de l'information en mémoire épisodique, mais, étant donné l'inter-substituabilité de ces éléments et leur dispersion dans l'espace, il est de notre avis que la question de la localisation est inadéquate pour traiter des phénomènes complexes qui donnent lieu à la cognition. La véritable question qui nous semble pertinente, au terme de ce mémoire, est celle d'identifier les composantes clés de la cognition – le quoi – et les processus qui se manifesteraient entre celles-ci – le comment.

La façon la plus appropriée de démystifier le quoi et le comment de la cognition est de modéliser celle-ci sous la forme d'un réseau complexe étendu. La relation entre le quoi et le comment se

calque d'ailleurs assez bien sur la relation entre les nœuds et les liens. Une telle modélisation permet, comme nous l'avons vu, de visualiser aisément les composantes importantes de la cognition et les configurations d'activités qui apparaîtraient entre celles-ci, ce que nous avons nommé les *transiently assembled local physical and social subsystems* (TALoPASS). Les TALoPASS désignent précisément ce genre de configurations, mais ont, en plus, la particularité d'inclure des composantes et des activités qui ne sont pas traditionnellement présentes dans les réseaux neurocentrés qui sont proposés d'ordinaire en neurosciences cognitives.

Il est de notre avis que la notion de TALoPASS présente des opportunités de recherche très prometteuses pour les sciences cognitives, à condition évidemment de bien délimiter le réseau complexe de la cognition. Si ce travail est bien fait, les TALoPASS permettraient potentiellement de résoudre différents problèmes rencontrés actuellement dans les sciences cognitives, comme celui de la localisation, de l'identification des composantes clés, de la granularisation et de la transition entre les états mentaux. Ces opportunités n'ont hélas pas été développées pleinement dans ce chapitre, faute d'un manque de temps et de ressources, mais il est néanmoins possible de concevoir l'apparition éventuelle d'un domaine de recherche interdisciplinaire autour de ces opportunités qui pourrait bénéficier d'un cadre théorique comme celui des TALoPASS, les neurosciences cognitives réseautiques étendues.

Conclusion

Nous avons vu à travers ce mémoire les différentes catégories de preuves qui motivent la réduction épistémologique et ontologique des états mentaux aux états cérébraux en neurosciences cognitives. Que ce soit l'étude des lésions cérébrales ou les données obtenues en neuroimagerie, un promoteur de cette réduction identitaire sont les doubles associations anatomo-fonctionnelles qu'on peut observer entre des fonctions cognitives et des régions du cerveau. Pour une fonction x , une région y est activée et, inversement, lorsque cette région y est inactivée ou éliminée du circuit, la fonction x se voit affectée négativement. Selon cette logique, on pourrait délimiter les éléments appartenant au système cognitif en appliquant un critère qui révèle des doubles associations structures-fonctions. Nous avons proposé, à cet effet, d'utiliser le critère de démarcation de la manipulabilité mutuelle. Selon ce critère, une composante donnée fait partie du mécanisme responsable de l'émergence du phénomène Ψ , si, dans une intervention idéale, une intervention sur cette composante modifie la valeur de Ψ et qu'une manipulation de Ψ provoque un changement d'activité de la composante observée.

Nous avons ensuite appliqué ce critère à différents exemples de structures extraneuronales – les échafauds intra-, inter- et extrasomatiques que Varga présente dans son cadre théorique de l'échafaudage actif de l'esprit – pour défendre l'idée que la cognition peut être réalisée par l'activité de composantes ontologiquement variées et distribuées dans l'espace. Couplées avec l'idée que la cognition est un phénomène dynamique constamment en évolution, la variété ontologique et la dispersion spatiale des composantes constitutives remettent en question l'intérêt de localiser définitivement les structures responsables de l'émergence et la régulation des processus cognitifs, comme ces structures s'étendraient sur l'environnement distal de l'individu et quelles changeraient tout au long du processus et d'occasion en occasion. Ce faisant, nous avons suggéré de remplacer la question de la localisation - la question du où - par d'autres questions plus pertinentes en sciences cognitives, les questions du quoi et du comment. En d'autres mots, nous avons proposé de tourner notre regard sur les composantes qui formeraient le système cognitif, leurs propriétés et les interactions dynamiques entre elles, plutôt que la localisation de ces composantes dans l'espace.

Pour ce faire, nous avons introduit le cadre théorique des *transiently assembled local physical and social subsystems* (TALoPASS) qui est inspiré de suggestions comme celles des TALoNs

d'Anderson (2014) et des TECS de Clark (2008). Les TALoPASS sont des configurations d'activité transitoire entre les différentes composantes neuronales et extraneuronales (intrasomatique, intersomatique et extra somatique) qui formeraient le système cognitif. L'idée centrale aux TALoPASS est que les facultés cognitives sont des processus continus dynamiques dont l'émergence repose sur l'interaction de composantes variées ontologiquement et qui peuvent changer tout au long du processus.

Étant donné la complexité des TALoPASS, nous avons suggéré de visualiser ces configurations d'activité en faisant appel à une modélisation réseautique où les différentes composantes pourraient être représentées à l'aide de nœuds et leurs interactions à l'aide de liens. Une telle modélisation, avons-nous soutenu, nous permettrait de saisir visuellement les TALoPASS à différents moments clés de l'émergence des processus cognitifs, et, surtout, les transitions entre ceux-ci. En effet, il est de notre avis que les transitions entre les TALoPASS sont cruciales à notre compréhension du mécanisme responsable de l'émergence des facultés cognitives. Il est possible, à partir de celles-ci, de dégager des informations importantes sur la contribution respective de certaines composantes au sein du mécanisme (ex. lesquelles sont impliquées plus longuement dans le processus? À quel moment exactement une composante laisse sa place à une autre dans les opérations qui mènent à l'émergence du processus?).

Cela étant dit, nous croyons que le cadre théorique des TALoPASS représente une avenue fertile de recherche en sciences cognitives. Par sa nature flexible et inclusive, ce cadre théorique offre une bonne marge de manœuvre pour tester et confirmer différentes hypothèses sur les composantes et les processus responsables des facultés cognitives, sans pour autant être teintée par des *a priori* sur la définition accordée à la cognition, les structures qui la composent et ses limites spatiales. De plus, les TALoPASS pourraient servir de critère de démarcation du cognitif. En effet, la stabilité des composantes à travers les configurations d'activité transitoire et la force de leurs connexions avec d'autres composantes jugées importantes pourraient éventuellement servir de critère pour déterminer leur appartenance au système cognitif.

Non seulement cela, mais l'évolution de ces configurations à travers le temps pourrait nous aider à comprendre les relations fonctionnelles complexes entre certains processus cognitifs et potentiellement à informer certaines taxonomies existantes au sujet des capacités cognitives. Ces configurations nous permettraient par exemple d'identifier les transitions d'une fonction cognitive à une autre à partir des modifications observées dans les composantes présentes à

chacune des configurations et leurs interactions avec d'autres parties du réseau (ex. si une partie du lobe frontal et les différents objets qui remplacent fonctionnellement cette structure sont inactives dans une configuration donnée à un moment du processus, il se peut que la capacité à laquelle nous sommes confronté ne soit plus de la récupération de l'information en mémoire, mais un autre processus, comme la production orale du langage). De même, si, en analysant deux processus, nous réalisons qu'ils partagent systématiquement plusieurs composantes et interactions en commun, cela pourrait nous amener à réviser les frontières taxonomiques entre ceux-ci.

Qui plus est, les TALoPASS pourraient être utilisés dans des domaines variés, au-delà de l'analyse fonctionnelle des processus cognitifs, comme dans l'étude du développement cognitif et la modélisation des mécanismes régulateurs de certaines psychopathologies. Bien que le présent mémoire n'ait pas investigué ces possibilités en détail, il est concevable, en effet, de modéliser l'émergence et la régulation des symptômes pathologiques présents dans certains troubles mentaux à l'aide d'une succession de TALoPASS. Pour faire écho au début de ce texte, on voit que des psychopathologies comme la schizophrénie sont des entités cliniques complexes qui ne peuvent pas être réduites qu'à un profil d'activité cérébrale dysfonctionnel. La schizophrénie comporte l'expérience phénoménologique de l'individu et ses relations avec le monde (Krueger, 2018), des relations qui contribuent à améliorer ou détériorer ses symptômes à travers le temps.

Il est bien documenté que les personnes schizophrènes tendent à éprouver de la difficulté avec les relations interpersonnelles et qu'elles privilégient des routines répétitives sécurisantes (Seeman, 2007). Ces routines, en retour, diminuent l'accessibilité à des nouvelles opportunités environnementales – le contact avec de nouveaux échafauds sociaux et matériels – qui pourraient améliorer certains symptômes. Dans ce contexte, on pourrait modéliser un des mécanismes régulateurs de la schizophrénie, comme une succession de configurations où l'agent entretient des liens très faibles voire négatifs avec différentes composantes inter- et extrasomatique sur plusieurs configurations (ex. il a vu une amie, mais a échoué à entretenir une discussion authentique avec celle-ci), et corrélérer la force de ces liens avec l'évolution des symptômes dans le temps. Cela dit, il serait tout aussi intéressant, sinon plus, d'identifier des composantes extraneuronales qui pourraient jouer un rôle stable thérapeutique dans cette psychopathologie (ex. des composantes qui préviendraient potentiellement la détérioration de certains symptômes). Ultimement, l'objectif serait d'avoir des configurations où ces

informations seraient combinées pour avoir un portrait plus complet de la situation écologique de l'individu et des facteurs qui contribueraient à améliorer ou détériorer son état.

Malgré ces différents domaines d'application, nous sommes conscients que les TALoPASS peuvent être rebutantes en raison de l'investissement requis pour en tirer profit. Il ne s'agit pas d'un cadre de type « *plug and play* » où il suffirait de prendre des éléments prédéfinis (ex. des composantes, des liens, la directionnalité et le poids de ceux-ci, etc.) pour ensuite l'appliquer à une problématique. Ce cadre, dans le but précisément d'être flexible et inclusif, requiert un effort particulier du modélisateur, car il incombe à ce dernier de sélectionner et justifier les éléments à retenir dans le modèle et à se doter de techniques pour filtrer efficacement l'information (réduire la complexité du modèle). Celui-ci peut faire usage d'un critère de démarcation, comme celui de la manipulabilité mutuelle, pour déterminer les éléments à inclure dans son modèle, tout comme il peut rassembler différentes sources d'influences facilitantes et causales selon ses besoins (ex. une recherche spécifique sur le rôle des systèmes de mémoire transactionnelle dans la récupération de l'information). La quantité d'information à inclure dans le modèle est laissée ultimement à la discrétion du modélisateur, mais il faut retenir que la valeur informative et explicative du modèle, en bout de ligne, est proportionnelle à celle-ci; plus il y a d'informations, plus probable est d'identifier des propriétés et des liens intéressants entre certaines composantes.

Nous ne savons pas quel avenir attend les TALoPASS. Ne seront-ils qu'une simple occurrence isolée dans le réseau immensément vaste des connaissances scientifiques ou le premier pas vers un nouveau domaine de recherches (ex. les neurosciences cognitives réseautiques étendues)? Quoiqu'il en soit, nous croyons que ce cadre est un pas de plus dans la bonne direction. Nous ne prétendons évidemment pas qu'il s'agit d'une réponse absolue aux maux qui affligent les sciences cognitives, mais qu'il réoriente, de pair avec de nombreuses propositions (ex. ICM, TALoNS, TECS, etc.), notre attention vers les éléments qui comptent réellement dans ce domaine: l'identification d'une taxonomie et de phénomènes cognitifs à expliquer ainsi que l'identification de composantes et d'interactions qui peuvent servir d'explication à ces phénomènes. Dans cette longue quête périlleuse à bord du bateau de Neurath, nous avons colmaté certaines brèches laissées par le localisationnisme; reste à savoir combien de temps il pourra tenir en mer avant la prochaine réparation.

Liste des références

Adams, E. L., Master, L., Buxton, O. M. & Savage, J. S. (2022). Sleep parenting practices are associated with infant self-soothing behaviors when measured using actigraphy. *Sleep medicine*, 95, 29-36.

Adams, F. & Aizawa, K. (2008). *The bounds of cognition*. Blackwell, Oxford, 216p.

Adams, F. & Garrison, R. (2013). The mark of the cognitive. *Minds and machines*, 23 (3), 339-352.

Addis, D. R., Barense, M. & Duarte, A. (2015). *The Wiley handbook on the cognitive neuroscience of memory*. Wiley-Blackwell, Oxford, 462p.

Adolphs, R. (2016). Human lesion studies in the 21st century. *Neuron*, 90 (6), 1151-1153.

Aguirre, G. K., Zarahn, E. & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *Neuroimage*, 5, 199–212.

Ahmed, E. & Hashish, A.H. (2006). On modeling the immune system as a complex system. *Theory in biosciences*, 124, 413-418.

Alberini, C. M. (2011). The role of reconsolidation and the dynamic process of long-term memory formation and storage. *Frontiers in behavioral neuroscience*, 5 (12).

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002). Lymphocytes and the cellular basis of adaptive immunity. Dans *Molecular biology of the cell*. Garland science, New York, 712p.

Alexanderson, G. L. (2006). About the cover: Euler and Königsberg's bridges: a historical view. *American mathematical society*, 43 (4), 567-573.

Aljondi, R., Szoeki, C., Steward, C., Yates, P. & Desmond, P. (2019). A decade of changes in brain volume and cognition. *Brain imaging and behavior*, 13 (1).

Allport, G. W. (1968). The historical background of modern social psychology. Dans Lindzey, G. & Aronson, E. *The handbook of social psychology*, vol. 1, 1-80.

Alstott, J., Breakspear, M., Hagmann, P., Cammoun, L. & Sporns, O. (2009). Modeling the impact of lesions in the human brain. *PLoS computational biology*, 5 (6).

Anderson, M. L. (2014). *After phrenology: neural reuse and the interactive brain*. The MIT press, Cambridge, 385p.

Andreasen, N. C. (2010). The lifetime trajectory of schizophrenia and the concept of neurodevelopment. *Dialogues in clinical neuroscience*, 12 (3), 409-415.

Arbula, S. (2020). Focal left prefrontal lesions and cognitive impairment: a multivariate lesion-symptom mapping approach. *Neuropsychologia*, 136.

Arciniegas, D. B., Held, K. & Wagner, P. (2002). Cognitive impairment following traumatic brain injury. *Current treatment options in neurology*, 4 (1), 43-57.

Argote, L. & Ren, Y. (2012) Transactive memory systems: a microfoundation of dynamic capabilities. *Journal of management studies*, 49 (8), 1375-1382.

Asch, S.E. (1956) Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological monograph: general and applied*, 70 (416).

Augustinack, J. C., Van der Kouwe, A. J. W., Salat, D. H., Benner, T., Stevens, A. A., Annese, J., Fischl, B., Frosch, M. P. & Corkin, S. (2014). H.M.'s contributions to neuroscience: a review and autopsy studies. *Hippocampus*, 24 (11), 1267-1286.

Ay, N., Müller, M. & Szkola, A. (2010). Effective complexity and its relation to logical depth. *IEEE transactions on information theory*, 56 (9), 4593-4607.

Baars, B. J. & Gage, N. M. (2010). Chapter 4 - the tools: imaging the living brain. Dans *Cognition, brain, and consciousness* (second edition). Academic press, p.94-125.

Ballard, D., Hayhoe, M. M. & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience* 7, (1) 66–80.

Barabási, A.-L. (2014). *Network science*, [en ligne]: [Chapter 0 – Network Science by Albert-László Barabási \(networksciencebook.com\)](#) (Consulté le 15/06/2021).

Baron-Cohen, S. 2003. *The essential difference: Men, women, and the extreme male brain*. London: Allen Lane.

Baron-Cohen, S. & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, 34 (2), 163-175.

Barrett, D. A. (2013). Multiple realizability, identity theory, and the gradual reorganization principle. *British journal of philosophy of science*, 64, 325-346.

Bechtel, W. (2008). L'épistémologie des données en neurosciences cognitives. Dans Poirier, P. & Faucher, L. (dir.) *Des neurosciences à la philosophie*. Éditions Syllepse, Paris, 528p.

Bechtel, W. (2010). Identity, reduction, and conserved mechanisms: perspectives from circadian rhythm research. Dans Gozzano, S. & Hill, C. S. (dir.) *New perspectives on type identity*. Cambridge university press, 43-65p.

Bechtel, W. & Richardson, R. C. (2010). *Discovering complexity: decomposition and localization as strategies in scientific research*. The MIT press, Cambridge, 286p.

Bell, M. A. & Cuevas, K. (2012). Using EEG to study cognitive development: issues and practices. *Journal of cognition and development*, 13 (3), 281-294.

- Benmelouka, A., Sherif, A. & Ebada, M. A. (2019). A review of the relationship between gut microbiota and memory. [en ligne] [\(7\) A Review of the Relationship Between Gut Microbiota and Memory | Request PDF \(researchgate.net\)](#) (consulté le 20/07/2021).
- Bennett, C. H. (1988). Logical depth and physical complexity. Dans Herken, R. (dir.) *The universal turing - a half-century survey*. Oxford university press, 227-257.
- Berger, A. (2003). Positron emission tomography. *British medical journal*, 326 (7404).
- Bernieri, F. & Rosenthal, R. (1991). Interpersonal coordination: behavior matching and interactional synchrony. Dans Feldman R. & Rime B. (dir.) *Fundamentals of nonverbal behavior*. Cambridge University Press, Cambridge, p. 401–432.
- Bhargava, A., Arnold, A. P., Bangasser, D. A., Denton, K. M., Gupta, A., Krause, L. M. H., Mayer, E. A., McCarthy, M., Miller, W. L., Raznahan, A. & Verma, R. (2021). Considering sex as a biological variable in basic and clinical studies: an endocrine society scientific statement. *Endocrine reviews*, 42 (3), 219-258.
- Bickle, W. J. (2003). *Philosophy and neuroscience: a ruthlessly reductive account*. Springer, 235p.
- Bickle, W. J. *Multiple Realizability*, The Stanford Encyclopedia of Philosophy, [en ligne]: [Multiple Realizability \(Stanford Encyclopedia of Philosophy/Summer 2020 Edition\)](#) (consultée le 19/03/2022).
- Billman, G. E. (2020). Homeostasis: the underappreciated and far too often ignored central organizing principle of physiology. *Frontiers in physiology*, 11, 200.
- Binder, J. R. & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in cognitive sciences*, 15 (11), 527-536.
- Binder, J. R. (2015). The Wernicke area. *Neurology*, 85 (24), 2170-2175.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*. [en ligne] [Fast unfolding of communities in large networks - IOPscience](#) (consulté le 20/01/2022).
- Bloom, B. S. (1956). *Taxonomy of educational objectives: the classification of educational goals*. Longmans, New York.
- Bonaz, B., Lane, R. D., Oshinsky, M. L., Kenny, P. J., Sinha, R., Mayer, E. A. & Critchley, H. D. (2021). Diseases, disorders, and comorbidities of interoception. *Trends in neuroscience*, 44 (1), 39-51.
- Brackenbury, T. & Pye, C. (2005). Semantic deficits in children with language impairments: issues for clinical assessment. *Language, speech, and hearing services in schools*, 36, 5-16.
- Breitenfeld, T., Jurasic, M. J. & Breitenfeld, D. (2014). Hippocrates: the forefather of neurology. *Neurological sciences*, 35, 1349-1352.
- Brewer, W. F. & Pani, J. R. (1996). Reports of mental imagery in retrieval from long-term memory. *Consciousness and cognition*, 5 (3), 265-287.

- Brizendine, L. (2007). *The female brain*. London: Bantam Press.
- Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie. *Bulletin de la société anatomique*, 6, 330-357.
- Bruineberg, J. & Fabry, R. (2022). Extended mind-wandering. *Philosophy and the mind sciences*, 3, 1-20.
- Brunet, A., Saumier, D., Liu, A., Streiner, D. L., Tremblay, J. & Pitman, R. K. (2018). Reduction of PTSD symptoms with pre-reactivation propranolol therapy: A randomized controlled trial. *American Journal of Psychiatry*, 175 (5), 427–433.
- Bula, W., Nikiel, J. & Tymchatyn, E. D. (1994). The Königsberg bridge problem for Peano Continua. *Canadian journal of mathematics*, 46 (6), 1175-1187.
- Bunge, S. A. & Kahn, I. (2009). Cognition: an overview of neuroimaging techniques. *Encyclopedia of neuroscience*, 2, 1063-1067.
- Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles and readers. *Neuron*, 68 (3), 362-385.
- Cacioppo, J. T. & Decety, J. (2009). What are the brain mechanisms on which psychological processes are based? *Perspectives on psychological science*, 4, 10-18.
- Carter, J. A., Clark, A., Kallestrup, J., Palermos, S. O. & Pritchard, D. (dir.) (2018). *Socially extended epistemology*. Oxford university press, Oxford, 318p.
- Celesia, G. G. (2012). Alcmaeon of Croton's observations on health, brain, mind, and soul. *Journal of the history of the neuroscience*, 21, 409-426.
- Chen, G., Chen, G., Xie, C., Ward, D. B., Li, W., Antuono, P. & Li, S.-J. (2012). A method to determine the necessity for global signal regression in resting-state fMRI studies. *Magnetic resonance in medicine*, 68 (6), 1828-1835.
- Chen, H., Li, Y., Liu, Q., Shi, Q., Wang, J., Shen, H., Chen, X., Ma, J., Ai, L. & Zhang, Y. M. (2019). Abnormal interactions of the salience network, central executive network, and default-mode network in patients with different cognitive impairments loads caused by leukoaraiosis. *Frontiers in neural circuits*, 13 (42).
- Cisek, P. (2019). Resynthesizing behavior through phylogenetic refinement. *Attention, perception & psychophysics*, 81 (7), 2265-2287.
- Clark, A. (1989). *Microcognition*. MIT Press, Cambridge, 226p.
- Clark, A. (1997). *Being there: putting brain, body, and world together again*. MIT press, Cambridge, 292p.
- Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, 58 (1), 7-19.
- Clark, A. (2008). *Supersizing the mind: embodiment, action, and cognitive extension*. Oxford university press, New York, 286p.

- Clark, A. (2010). Coupling, constitution, and the cognitive kind: a reply to Adams and Aizawa. Dans *The extended mind*, MIT press, Cambridge, p. 81-99.
- Clark, C. N., Golden, H. L. & Warren, J. D. (2015). Acquired amusia. *Handbook of clinical neurology*, 129, 607-631.
- Coughlan, A. K. & Warrington, E. K. (1981). The impairment of verbal semantic memory: a single case study. *Journal of neurology, neurosurgery and psychiatry*, 44 (12), 1079-1083.
- Cox, G. E. & Shiffrin, R. M. (2012). Criterion setting and the dynamics of recognition memory. *Topics in cognitive science*, 4, 135-150.
- Craver, C. (2007). Constitutive explanatory relevance. *Journal of philosophical research*, 32, 3-20.
- Craver, C. (2007). *Explaining the brain: mechanisms and the mosaic unity of neuroscience*. Oxford university press, 308p.
- Craver, C. (2008). Fonction et mécanismes dans les neurosciences contemporaines. Dans Poirier, P. et Faucher, L. (dir.) *Des neurosciences à la philosophie*. Paris : Syllepse, p. 67-90.
- Da Silva, F. L. (2013). EEG and MEG: relevance to neuroscience. *Neuron*, 80 (5), 1112-1128.
- Dael, N., Mortillaro, M. & Scherer, K. R. (2011). Emotion expression in body action and posture. *Emotion*, 12 (5), 1085-1101.
- Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of prefrontal cortex. *Philosophical transactions: biological sciences*, 351 (1346), 1413-1420.
- Damasio, A. R. (1998). Emotion in the perspective of an integrated nervous system. *Brain research reviews*, 26, 83-86.
- Damasio, H. (2018). Phineas Gage: the brain and the behavior. *Revue neurologique*, 174 (10), 738-739.
- Deloire, M., Salort, E., Bonnet, M., Arimone, Y., Boudineau, M., Amieva, H., Barroso, B., Ouallet, J.-C., Pachai, C., Galliaud, E., Petry, K. G., Dousset, V., Fabrigoule, C. & Brochet, B. (2005). Cognitive impairment as marker of diffuse brain abnormalities in early relapsing remitting multiple sclerosis. *Journal of neurology, neurosurgery & psychiatry*, 76 (4), 519-526.
- De Blasi, B., Caciagli, L., Storti, S. F., Galovic, M., Koepp, M., Menegaz, G., Barnes, A. & Galazzo, I. B. (2020). Noise removal in resting-state and task fMRI: functional connectivity and activation maps. *Journal of neural engineering*, 17 (4).
- Del Pozo, M., Manuel, C., González-Arangüena, E. & Owen, G. (2011). Centrality in directed social networks. A game theoretic approach. *Social networks*, 33, 191-200.
- Dennett, D. C. (1987). *The intentional stance*. MIT press, Cambridge, 388p.

- D'Esposito, M. (2000). Functional neuroimaging of cognition. *Seminars in neurology*, 20 (4), 487-498.
- Dickerson, B. C. & Eichenbaum, H. (2010). The episodic memory system: neurocircuitry and disorders. *Neuropsychopharmacology*, 35, 86-104.
- Di Paolo, E. & De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in human neuroscience*, 6 (163).
- Di Prospero, N. D., Kim, S. & Yassa, M. A. (2022). Chapter 8 - Magnetic resonance imaging biomarkers for cognitive decline in down syndrome. Dans Head, E. & Lott, I. (dir.) *The neurobiology of aging in Alzheimer disease in down syndrom*, Academic press, 149-172.
- Dijkstra, K., Kaschak, M., & Zwaan, R. (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition*, 102 (1), 139-149.
- Diykh, M., Li, Y. & Wen, P. (2016). EEG sleep stages classification based on time domain features and structural graph similarity. *IEEE transactions on neural systems and rehabilitating engineering*, 24 (11), 1159-1168.
- Dominique, J.-F. & Margraf, J. (2008). Glucocorticoids for the treatment of post-traumatic stress disorder and phobias: A novel therapeutic approach. *European Journal of Pharmacology*, 583 (2-3), 365-371.
- Dronkers, N. F., Redfern, B. B. & Knight, R. T. (2000). The neural architecture of language disorders. Dans Gazzaniga, M. (dir.) *The new cognitive neurosciences*, MIT press, Cambridge, 949-958.
- Dürschmid, S., Edwards, E., Reichert, C., Dewar, C., Hinrichs, H., Heinze, H.-J., Kirsch, H. E., Dalal, S. S., Deouell, L. Y. & Knight, R. T. (2016). Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *PNAS*, 113 (24), 6755-6760.
- Edelman, G. M. & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the national academy of sciences of the USA*, 98 (24), 13763-13768.
- Ellfolk, U., Joutsa, J., Rinne, J. O., Parkkola, R., Jokinen, P. & Karrasch, M. (2014). Striatal volume is related to phonemic verbal fluency but not to semantic or alternating verbal fluency in early Parkinson's disease. *Journal of neural transmission*, 121, 33-40.
- Feinstein, J. S. (2013). Lesion studies of human emotion and feeling. *Current opinion in neurobiology*, 23 (3), 304-309.
- Feldman, R. (2003). Infant-mother and infant-father synchrony: the coregulation of positive arousal. *Infant mental health journal*, 24 (1), 1-23.
- Fell, J. & Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature reviews neuroscience*, 12 (2), 105-118.
- Fernández, A. R., Triviño Ibáñez, E. M., Gómez Río, M., Pérez Lázaro, J.J., Fernández Ruiz, I., Ramírez Navarro, A., García Rivero, Y., Córdoba Cañete, E., Romero Fernández, C. & Llamas-Elvira, M. (2019). Development of a positron emission tomography risk map. *Spanish journal of nuclear medicine and molecular imaging*, 38 (1), 38-45.

Figdor, C. (2010). Neuroscience and the multiple realization of cognitive functions. *Philosophy of science*, 77 (3), 419-456.

Fine, C. (2012). Explaining, or sustaining, the status quo? The potentially self-fulfilling effects of "hardwired" accounts of sex differences. *Neuroethics*, 5, 285-294.

Frith, C. (2007). *Making up the mind: how the brain creates our mental world*. Malden, MA: Blackwell Publication. 246p.

Fuchs, Thomas (2011). The brain - a mediating organ. *Journal of consciousness studies*, 18 (7-8), 196-221.

Gall F. J. (1835). *On the functions of the brain and each of its parts: With observations on the possibility of determining the instincts, propensities, and talents, or the moral and intellectual dispositions of men and animals, by the configuration of the brain and head*. Marsh, Capen and Lyon, Lewis, Boston.

Gallagher, S. & Crisafi, A. (2009). Mental institutions. *Topoi*, 28, 45-51.

Ganis, G. & Kosslyn, S. M. (2002). Neuroimaging. Dans *Encyclopedia of the human brain*, Academic press, 493-505.

Ghanbari, M., Soussia, M., Jiang, W., Wei, D., Yap, P.-T., Shen, D. & Zhang, H. (2022). Alterations of dynamic redundancy of functional brain subnetworks in Alzheimer's disease and major depression disorders. *NeuroImage: clinical*, 33.

Gifford, G., McCutcheon, R. & McGuire, P. (2020). Chapter 9 - Neuroimaging studies in people at clinical high risk for psychosis. Dans Thompson, A. & Broome, M. R. (dir.) *Risk factors for psychosis*. Academic press, p.167-182.

Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery clinics of north America*, 22 (2), 133-139.

Goldman, D. (2011). Molecular etiologies of schizophrenia: are we almost there yet? *American journal of psychiatry*, 168 (9), 879-881.

Gollwitzer, P. M. (1999). Implementation intentions: strong effects of simple plan. *American psychologist*, 54 (7), 493-503.

Gordon, R. & Ramani, G. B. (2021). Integrating embodied cognition and information processing: a combined model of the role of gesture in children's mathematical environments. *Frontiers in psychology*, 12.

Gosling, C. J., Caparos, S. & Moutier, S. (2020). The interplay between the importance of a decision and emotion in decision-making. *Cognition & emotion*, 34 (6), 1260-1270.

Gothard, K. M. (2020). Multidimensional processing in the amygdala. *Nature reviews neuroscience*, 21, 565-575.

- Grodzinsky, Y. & Santi, A. (2008). The battle for Broca's region. *Trends in cognitive sciences*, 12 (12), 474-480.
- Gu, S. et al. (2015). Controllability of structural brain networks. *Nature communications*, 8414 (6).
- Gur, R. C., Turetsky, B. I., Matsui, M., Yan, M., Bilker, W., Hughett, P. & Gur, R. E. (1999). Sex differences in brain gray and white matter in healthy young adults: correlations with cognitive performance. *Journal of neuroscience*, 19 (10), 4065-4072.
- Gurian, M. & Stevens, K. (2004). With boys and girls in mind. *Educational Leadership*, 62 (3): 21-26.
- Hameroff, S. & Penrose, R. (2014). Consciousness in the universe: a review of the 'Orch OR' theory. *Physics of life reviews*, 11 (1), 39-78.
- Hänggi, J., Fövényi, L., Liem, F., Meyer, M. & Jäncke, L. (2014). The hypothesis of neuronal interconnectivity as a function of brain size - a general organization principle of the human connectome. *Frontiers in neuroscience*, 8.
- Hanslmayr, S., Staresina, B. P. & Bowman, H. (2016). Oscillations and episodic memory: addressing the synchronization/desynchronization conundrum. *Trends in neurosciences*, 39 (1), 16-25.
- Harbecke, J. (2014). The role of supervenience and constitution in neuroscientific research. *Synthese*, 191, 725-743.
- Harris, C. B., Barnier, A. J., Sutton, J. & Keil, P. G. (2014). Couples as socially distributed cognitive systems: remembering in everyday social and material contexts. *Memory studies*, 7 (3), 285-297.
- Hasegawa, I., Hayashi, T. & Miyashita, Y. (1999). Memory retrieval under the control of the prefrontal cortex. *Annals of medicine*, 31 (6), 380-387.
- Hawking, S. (2013). *La brève histoire de ma vie*, Flammarion, 173p.
- Hayne, H. & Herbert, J. (2004). Verbal cues facilitate memory retrieval during infancy. *Journal of experimental psychology*, 89, 127-139.
- Heyck, M. & Ibarra, A. (2019). Microbiota and memory: a symbiotic therapy to counter cognitive decline? *Brain circulation*, 5 (3), 124-129.
- Hickok, G., Costanzo, M., Capasso, R. & Miceli, G. (2011). The role of Broca's area in speech perception: evidence from aphasia revisited. *Brain language*, 119 (3), 214-220.
- Hoekzema, E., Barba-Müller, E., Pozzobon, C., Picado, M., Lucco, F., Garcia-Garcia, D., Soliva, J. C., Tobeña, A., Desco, M., Crone, E. A., Ballesteros, A., Carmona, S. & Vilarroya, O. (2016). Pregnancy leads to long-lasting changes in human brain structure. *Nature neuroscience*, 20 (2), 287-295.

- Hoffman, G. (2012). What, if anything, can neuroscience tell us about gender differences? Dans Bluhm, R., Jacobson, A. J., Maibom, H. L. (dir.) *Neurofeminism: issues at the intersection of feminist theory and cognitive science*. Palgrave Macmillan, London.
- Hofman, M. A. (2014). Evolution of the human brain: when bigger is better. *Frontiers in neuroanatomy*, 8 (15).
- Huettel, S. A. (2012). Event-related fMRI in cognition. *NeuroImage*, 62 (2), 1152-1156.
- Hutchins, E. (1996). *Cognition in the wild*. The MIT press, Cambridge, 402p.
- Hutto, D. D. & Myin, E. (2013). *Radicalizing enactivism: basic minds meet content*. MIT press, Cambridge, 232p.
- Jablensky, A., McNeil, T. F. & Morgan, V. A. (2017). Barbara Fish and a short history of the neurodevelopmental hypothesis of schizophrenia. *Schizophrenia bulletin*, 43 (6), 1158-1163.
- Jiang, S. et al. (2000). An experimental study on traumatic brain injury for inducing neuronal apoptosis in rats. *Journal of west China University of medical sciences*, 31 (3), 358-361.
- Joffily, M. & Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLOS computational biology*, 9 (6).
- Johnson, E. L., Kam, J. W. Y., Tzovara, A. & Knight, R. T. (2020). Insights into human cognition from intracranial EEG: a review of audition, memory, internal cognition, and causality. *Journal of neural engineering*, 17 (5).
- Kaiser, A. (2012). Re-conceptualizing "sex" and "gender" in the human brain. *Zeitschrift für psychologie*, 220 (2), 130-136.
- Kaplan, S. (2009). Grasping at ontological straws: overcoming reductionism in the Advaita Vedanta - neuroscience dialogue. *Journal of the American academy of religion*, 77 (2), 238-274.
- Kaplan, D. M. (2012). How to demarcate the boundaries of cognition. *Biology & Philosophy*, 27, 545-570.
- Kida, S. (2019). Reconsolidation/destabilization, extinction and forgetting of fear memory as therapeutic targets for PTSD. *Psychopharmacology*, 236 (1), 49-57.
- Kieronska, S. & Sloniewski, P. (2019). The usefulness and limitations of diffusion tensor imaging - a review study. *European journal of translational and clinical medicine*, 2 (2).
- Kim, S. G., Richter, W. & Ugurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic resonance in medicine*, 37 (4), 631-636.
- Kirchhoff, M., Parr, T., Palacios, E., Friston, K. & Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference, and the free energy principle. *Journal of the royal society interface*, 15 (138).
- Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, 73, 31-68.

- Korn, C. W. & Bach, D. R. (2015). Maintaining homeostasis by decision-making. *PLoS computational biology*, 11 (5).
- Krishnan, K. R. R. (2012). Chapter 6 - Structural imaging in psychiatric disorders. *Handbook of clinical neurology*, 106, 89-95.
- Kropotov, J. D. (2016). Chapter 1.4 - Positron emission tomography. Dans *functional neuromarkers for psychiatry*, 27-30p.
- Krueger, J. (2018). Schizophrenia and the scaffolded self. *Topoi*, 39, 597-609.
- Kwok, S. C. (2013). Where neuroimaging and lesion studies meet. *Journal of neuroimaging*, 23, 1-4.
- Lakes-Harlan, R. (2013). Lesion-induced insights in the plasticity of the insect auditory system. *Frontiers in physiology*, 4 (48).
- Leask, S. J. (2018). Environmental influences in schizophrenia: the known and the unknown. *Advances in psychiatric treatment*, 10 (5).
- Lebel, C. & Deoni, S. (2018). The development of brain white matter microstructure. *Neuroimage*, 182, 207-218.
- LeDoux, J. E. (2002). *Synaptic self: how our brains become who we are*. Penguin Books. 406p.
- Lehnertz, K., Rings, T. & Bröhl, T. (2021). Time in brain: how biological rhythms impact on EEG signals and on EEG-derived brain networks. *Frontiers in network physiology*, 27.
- Lewis, K. (2003). Measuring transactive memory systems in the field: scale development and validation. *Journal of Applied Psychology*, 88, 587– 604.
- Lewis, K. and Herndon, B. (2011). Transactive memory systems: current issues and future research directions. *Organization Science*, 22, 1254– 1265.
- Li, J., Ren, H., He, Y., Li, Z., Ma, X., Yuan, L., Ouyang, L., Zhou, J., Wang, D., Li, C., Chen, X., Han, H. & Tang, J. (2020). Anterior cingulate cortex glutamate levels are related to response to initial antipsychotic treatment in drug-naïve first-episode schizophrenia patients. *Frontiers in psychology*, 11, 1-9.
- Liang, D. W., Moreland, R. & Argote, L. (1995). Group versus individual training and group performance: the mediating role of transactive memory. *Personality and social psychology bulletin*, 21, 384-393.
- Liao, W., Ding, J., Marinazzo, D., Xu, Q., Wang, Z., Yuan, C., Zhang, Z., Lu, G. & Chen, H. (2011). Small-world directed networks in the human brain: multivariate Granger causality analysis of resting state fMRI. *Neuroimage*, 54 (4), 2683-2694.
- Liu, C., Wu, J., Liu, W. & Hu, W. (2021). Enhancing graph neural networks by a high-quality aggregation of beneficial information. *Neural networks*, 142, 20-33.

- Luo, J., Hippe, D. S., Rahbar, H., Parsian, S., Rendi, M. H. & Partridge, S. C. (2019). Diffusion tensor imaging for characterizing tumor microstructure and improving diagnostic performance on breast MRI: a prospective observational study. *Breast cancer research*, 21 (1), 102.
- Lv, Y., Wei, N. & Li, K. (2019). Directed connectivity in large-scale brain networks for precision grip force control. 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC).
- Maibom, H. & Bluhm, R. (2014). A situationist account of sex/gender differences: implications for neuroimaging research. Dans Schmitz, S. & Höppner, G. (dir.) *Gendered neurocultures: feminist and queer perspectives on current brain discourses*. GRéTA, Berlin.
- Maiese, M. (2016). Affective scaffolds, expressive arts, and cognition. *Frontiers in psychology*, 359, 7.
- Maiese, M. (2016). *Embodied selves and divided minds*. Oxford university press, Oxford, 295p.
- Mainy, N., Kahane, P., Minotti, L., Hoffman, D., Bertrand, O. & Lachaux, J.-P. (2007). Neural correlates of consolidation in working memory. *Human brain mapping*, 28 (3), 183-193.
- Mäkinen, E. & Siirtola, H. (2005). The barycenter heuristic and the reorderable matrix. *Informatica*, 29 (3), 357-364.
- Mandik, P. (2011). Supervenience and neuroscience. *Synthese*, 180, 443-463.
- Marroquin, B. & Nolen-Hoeksema, S. (2015). Emotion regulation and depressive symptoms: close relationships as social context and influence. *Journal of personality and social psychology*, 109 (5), 836.
- Marshall, J. C. & Fink, G. R. (2003). Cerebral localization. Then and now. *Neuroimage*, 20, 52-57.
- Mather, M., Cacioppo, J. T. & Kanwisher, N. (2013). How fMRI can inform cognitive theories. *Perspectives on psychological science*, 8 (1), 108-113.
- McClelland, J. L. (2001). Cognitive neuroscience. Dans Smelser, N. J. & Baltes, P. B. (dir.) *International encyclopedia of the social & behavioral sciences*. Pergamon, 23185 p.
- McClelland, J. L. (2010). Emergence in cognitive science. *Topics in cognitive science*, 2 (4), 751-770.
- Mcguffin, M. J. (2012). Simple algorithms for network visualization: a tutorial. *Tsinghua science and technology*, 17 (4), 1-16.
- Medina, J. & Fischer-Baum, S. (2017). Single-case cognitive neuropsychology in the age of big data. *Cognitive neuropsychology*, 34 (7-8), 440-448.
- Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology*, 19, 329-344.

Menary, K., Collins, P. F., Porter, J. N., Muetzel, R., Olson, E. A., Kumar, V., Steinbach, M., Lim, K. O. & Luciana, M. (2013). Associations between cortical thickness and general intelligence in children, adolescents, and young adults. *Intelligence*, 41 (5), 597-606.

Menon, V., Ford, J. M., Lim, K. O., Glover, G. H. & Pufferbaum, A. (1997). Combined event-related fMRI and EEG evidence for temporal-parietal cortex activation during target detection. *Neuroreport*, 8 (14), 3029-3037.

Merritt, M. (2013). Instituting impairment: extended cognition and the construction of female sexual dysfunction. *Cognitive systems research*, 25-26, 47-53.

Metzger, M. A. (2002). Context dependent memory: the role of environmental cues. Dans *Forget it? Sources, theories, and mechanisms of alterations in mnemonic function*. Erudition books.

Metzger, M. A. (2017). Dynamical models of cognition. Dans Magnani, L. & Bertolotti, T. *Handbook of model-based science*. Springer.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67, 371-378.

Mitchell, M. (2010). *Complexity: a guided tour*. Oxford university press, 368p.

Molnár, Z. (2004). Thomas Willis (1621-1675), the founder of clinical neuroscience. *Nature reviews neuroscience*, 5, 329-335.

Mortimer, J. et al. (2012). Changes in brain volume and cognition in a randomized trial of exercise and social interaction in a community-based sample of non-demented chinese elders. *Journal of Alzheimer's disease*, 30 (4), 757-766.

Moses, W. W. (2011). Fundamental limits of spatial resolution in PET. *Nuclear instruments and methods in physics research*, 648 (1), 236-240.

Mudrik, L. & Maoz, U. (2014). "Me & my brain": exposing neuroscience's closet dualism. *Journal of cognitive neuroscience*, 27 (2), 211-221.

Mürner-Lavanchy, I., Steinlin, M., Kiefer, C., Weisstanner, C., Ritter, B. C., Perrig, W. & Everts, R. (2014). Delay of cortical thinning in very preterm born children. *Early human development*, 90 (9), 443-450.

Neurath, O. (1921). Anti-Spengler. Dans *Empiricism and sociology*. Vienna circle collection, vol. 1, Reidel, Dordrecht, 158-213.

Newman, M. E. J. (2003). The structure and function of complex networks. *Society for industrial and applied mathematics*, 45 (2), 167-256.

Nicolini, C., Bordier, C. & Bifone, A. (2017). Community detection in weighted brain connectivity networks beyond the resolution limit. *Neuroimage*, 146, 28-39.

Nikolov, N. S. (2015). Sugiyama algorithm. *Encyclopedia of algorithms*, 1-6.

- Noachtar, S. & Rémi, J. (2009). The role of EEG in epilepsy: a critical review. *Epilepsy & behavior*, 15 (1), 22-33.
- Noirhomme, Q. & Laureys, S. (2014). Consciousness and unconsciousness: an EEG perspective. *ENCS*, 45 (1), 4-5.
- O'Donnell, L. J. & Westin, C.-F. (2011). An introduction to diffusion tensor image analysis. *Neurosurgery clinics of North America*, 22 (2), 185.
- Ohman, A. (2005). The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology*, 30 (10), 953-958.
- Opitz, B., Rinne, T., Mecklinger, A., Von Cramon, D. Y. & Schröger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *NeuroImage*, 15 (1), 167-174.
- Orbell, S., Hodgkins, S. & Sheeran, P. (1997). Implementation intentions and the theory of planned behavior. *Personality & social psychology bulletin*, 23 (9), 945-954.
- Pavitt, C. (2003). Why we have to be reductionists about group memory. *Human communication research*, 29 (4), 592-599.
- Peper, E. & Lin, I.-M. (2012). Increase or decrease depression: how body postures influence your energy level. *Biofeedback*, 40 (3), 125-130.
- Perone, S. & Simmering, V. R. (2017). Applications of dynamic systems theory to cognition and development: new frontiers. *Advances in child development and behavior*, 52, 43-80.
- Poirier, P. & Faucher, L. (2020). A new hope: a better ICM to understand human cognitive architectural variability. *Synthese*.
- Poldrack, A., Mumford, J. A. & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge university press, 228p.
- Prestwich, A., Sheeran, P., Webb, T. L. & Gollwitzer, P. M. (2015). Implementation intentions. Dans Conner, M. & Norman, P. (dir.) *Predicting and changing health behavior*. Open university press, p.321-357.
- Price, C. J. & Friston, J. K. (2002). Degeneracy and cognitive anatomy. *Trends in cognitive sciences*, 6 (10), 416-421.
- Putnam, H. (1967). The nature of mental states. Dans Capitan, W. H. & Merrill, D. D. (dir.) *Art, mind, and religion*. Pittsburgh University Press, 223p.
- Ranzenberger, L. R. & Snyder, T. (2022). Diffusion tensor imaging. StatPearls [en ligne]: <https://pubmed.ncbi.nlm.nih.gov/30726046/> (consulté le 10/03/2022).
- Recanatesi, S., Katkov, M., Romani, S. & Tsodyks, M. (2015). Neural network model of memory retrieval. *Frontiers in computational neuroscience*, 9 (149).

Reck, C., Hunt, A., Fuchs, T., Weiss, R., Noon, A., Moehler, E., Downing, G., Tronick, E. Z. & Mundt, C. (2004). Interactive regulation of affect in postpartum depressed mothers and their infants. *Psychopathology*, 3, 272-280.

Ren, Y., Carley, K. M. and Argote, L. (2006). The contingent effects of transactive memory: when is it more beneficial to know what others know? *Management Science*, 52, 671–682.

Ribes-Iñesta, E. (2004). Behavior is abstraction, not ostension: conceptual and historical remarks on the nature of psychology. *Behavior and philosophy*, 32 (1), 55-68.

Riskind, J. H. (1983). Nonverbal expressions and the accessibility of life experience memories: A congruence hypothesis. *Social Cognition*, 2, 62–86.

Robinson, O. J., Vytal, K., Cornwell, B. R. & Grillon, C. (2013). The impact of anxiety upon cognition: perspectives from human threat of shock studies. *Frontiers in human neuroscience*, 7, 203.

Rocca, J. (1997). Galen and the ventricular system. *Journal of the history of the neurosciences*, 6 (3), 227-239.

Rosen, M. L., Sheridan, M. A., Sambrook, K. A., Peverill, M. R., Meltzoff, A. N. & McLaughlin, K. A. (2018). The role of visual association cortex in associative memory formation across development. *Journal of cognitive neuroscience*, 30 (3), 365-380.

Rowe, C. & Healy, S. D. (2014). Measuring variation in cognition. *Behavioral ecology*, 25 (6), 1287-1292.

Rowlands, M. (2010). *The new science of the mind: from extended mind to embodied phenomenology*. MIT press, Cambridge, 264p.

Royer, J. M., Sefkow, S. R. & Kropf, R. B. (1977). Contributions of existing knowledge structure to retroactive inhibition in prose learning. *Contemporary educational psychology*, 2 (1), 31-36.

Ruff, C. C. & Huettel, S. A. (2014). Experimental methods in cognitive neuroscience. Dans Glimcher, P. W. & Fehr, E. (dir.) *Neuroeconomics*. Academic press, 560p.

Rugg, M. D. & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current opinion in neurobiology*, 23 (2), 255-260.

Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *Journal of philosophy*, 101 (8), 389-428.

Salvan, P., Tournier, J. D., Bataille, D., Falconer, S., Chew, A., Kennea, N., Aljabar, P., Dehaene-Lambertz, G., Arichi, T., Edwards, A. D. & Counsell, S. J. (2017). Language ability in preterm children is associated with arcuate fasciculi microstructure at term. *Human brain mapping*, 38 (8), 3836-3847.

Sass, L., Borda, J. P., Madeira, L., Pienkos, E. & Nelson, B. (2018). Varieties of self disorder: a bio-pheno-social model of schizophrenia. *Schizophrenia bulletin*, 44 (4), 720-727.

- Sayood, K. (2008). Information theory and cognition: a review. *Entropy*, 20 (9), 706.
- Schedlbauer, A. M., Copara, M. S., Watrous, A. J. & Ekstrom, A. D. (2014). Multiple interacting brain areas underlie successful spatiotemporal memory retrieval in humans. *Scientific reports*, 6431 (4).
- Schermerhorn, A. C. & Cummings, E. M. (2008). Transactional family dynamics: a new framework for conceptualizing family influence processes. *Advances in child development and behavior*, 36, 187-250.
- Searle, J. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, 3, 417-424.
- Seeman, M. V. (2007). Symptoms of schizophrenia: normal adaptations to inability. *Medical hypotheses*, 69 (2), 253-257.
- Seligman, E., Schubert, T. & Achutha Kiran Kumar, M. V. (2015). Effective FPV for design exercise. Dans *Formal verification*. Elsevier, 408p.
- Sheya, A. & Smith, L. (2019). Development weaves brains, bodies, and environments into cognition. *Language, cognition and neuroscience*, 34 (10), 1266-1273.
- Shuren, J. E., Schefft, B. K., Yeh, H. S., Privitera, M. D., Cahill, W. T. & Houston, W. (1995). Repetition and the arcuate fasciculus. *Journal of neurology*, 242 (9), 596-598.
- Sidharta, S. L., Sajeev, K. J., Nelson, A. J., Cooke, J. C. & Worthley, M. (2013). Stress-induced cardiomyopathy and possible link to cerebral executive function: a case report. *The primary care companion for CNS disorders*, 15 (6).
- Sieniutycz, S. (2020). Properties of complex systems. Dans *Complexity and complex thermo-economic systems*. Elsevier, 415p.
- Siew, C. S. Q., Wulff, D. U., Beckage, N. M. & Kenett, Y. N. (2019). Cognitive network science: a review of research on cognition through the lens of network representations, processes, and dynamics. *Cognitive network science: a new frontier*.
- Skudlarski, P., Jagannathan, K., Calhoun, V. D., Hampson, M., Skudlarska, B. A. & Pearlson, G. (2008). Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations. *Neuroimage*, 43 (3), 554-561.
- Smirnov, A. A. (1973). Conditions for retroactive inhibition. Dans *Problems of the psychology of memory*. Monographs in Psychology. Springer, Boston.
- Smith, A. M., Lewis, B. K., Ruttimann, U. E., Ye, F. Q., Sinnwell, T. M., Yang, Y., Duyn, J. H. & Frank, J. A. (1999). Investigation of low frequency drift in fMRI signals. *Neuroimage*, 9, 526-33.
- Smith, L. B. (2005). Cognition as a dynamic system: principles from embodiment. *Developmental review*, 25, 278-298.

- Smyth, C. (2020). A network model of cognition: results of an experiment on vehicle driving with navigation aids while replying to test questions. [en ligne] [Model of Cognition Control: Results of an Experiment on Vehicle Driving with Navigation Aids while Replying to Test Questions \(researchgate.net\)](#) (consulté le 04/02/2022).
- Smyth, M. J., Cretney, E., Kelley, J. M., Westwood, J. A., Street, S. E. A., Yagita, H., Takeda, K., Van Dommelen, S. L. H., Degli-Esposito, M. A. & Hayakawa, Y. (2005). Activation of NK cell cytotoxicity. *Immunology*, 42 (4), 501-510.
- Soares, J. M., Marques, P., Alves, V. & Sousa, N. (2013). A hitchhiker's guide to diffusion tensor imaging. *Frontiers in neuroscience*, 7 (31).
- Soikkeli, R., Partanen, J., Soininen, H., Pääkönen, A. & Riekkinen, P. (1991). Slowing of EEG in Parkinson's disease. *Electroencephalography and clinical neurophysiology*, 79 (3), 159-165.
- Sotiras, A., Toledo, J. B., Gur, R. E., Gur, R. C., Satterthwaite, T. D. & Davatzikos, C. (2017). Patterns of coordinated cortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion. *Proceedings of the national academy of sciences*, 114 (13), 3527-3532.
- Spivey, M. (2020). *Who you are - the science of connectedness*. The MIT press, Cambridge, 359p.
- Sporns, O., Tononi, G. & Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS computational biology*, 1 (4).
- Sprevak, M. (2009). Extended cognition and functionalism. *Journal of philosophy*, 106 (9), 503-527.
- Staempfli, P., Reischauer, C., Jaermann, T., Valavanis, A., Kollias, S. & Boesiger, P. (2008). Combining fMRI and DTI: a framework for exploring the limits of fMRI-guided DTI fiber tracking and for verifying DTI-based fiber tractography results. *Neuroimage*, 1 (39), 119-126.
- Stein, D. G. (2017). Brain injury: functional recovery after. Dans *Reference module in neuroscience and biobehavioral psychology*. Elsevier. [en ligne] [\(7\) Brain Injury: Functional Recovery After \(researchgate.net\)](#) (Consulté le 17/12/2021).
- Stefanou, M. I. (2018). The footprints of neuroscience in Alexandria during the 3rd-century BC: Herophilus and Erasistratus. *Journal of medical biography*, 0, 1-9.
- Sterelny, K. (2010). Minds: extended or scaffolded? *Phenomenology and the cognitive sciences*, 9 (4), 465-481.
- Stevenson, I. H. & Kording, K. P. (2012). How advances in neural recording affect data analysis. *Nature neuroscience*, 14 (2), 139-142.
- Stigler, K. A. & McDougle, C. J. (2013). Chapter 3.1 - Structural and functional MRI studies of autism spectrum disorders. Dans Buxbaum, J. D. & Hof, P. R.(dir.) *The neuroscience of autism spectrum disorders*, 251-266.

- Stuart, G., Spruston, N., Sakmann, B. & Häusser, M. (1997). Action potential initiation and backpropagation in neurons of the mammalian CNS. *Trends in neurosciences*, 20 (3), 125-131.
- St-Jacques, P. L., Kragel, P. A. & Rubin, D. C. (2012). Dynamic neural networks supporting memory retrieval. *Neuroimage*, 57 (2), 608-616.
- Sun, L., Fan, Z., Ding, X., Huang, Y. & Paisley, J. (2019). Region-of-interest undersampled MRI reconstruction: a deep convolutional neural network approach. *Magnetic resonance imaging*, 63, 185-192.
- Sutton, P., Doyle, L. E. & Nolan, K. E. (2006, june). A reconfigurable platform for cognitive networks. *Proceedings of IEEE*, Mykonos island, Greece.
- Takahata, R. & Moghaddam, B. (2003). Activation of glutamate neurotransmission in the prefrontal cortex sustains the motoric and dopaminergic effects of phencyclidine. *Neuropsychopharmacology*, 28 (6), 1117-1124.
- Tatum, W. O. (2014). *Handbook of EEG interpretation*. Demos medical publishing, 361p.
- Theiner, G. (2013). Transactive memory systems: a mechanistic analysis of emergent group memory. *Review of philosophy and psychology*, 4, 65-89.
- Thomas, J. B., Brier, M. R., Ortega, M., Benzinger, T. L. & Ances, B. M. (2015). Weighted brain networks in disease: centrality and entropy in HIV and aging. *Neurobiology of aging*, 36 (1), 401-412.
- Thomas, R. W. & DaSilva, L. A. (2009). Cognitive networking. Dans Fette, B. A. (dir.) *Cognitive radio technology*. Academic press, 828p.
- Tremblay, P. & Dick, A. S. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain & language*, 162, 60-71.
- Turner, R. (2016). Uses, misuses, new uses, and fundamental limitations of magnetic resonance imaging in cognitive science. *Philosophical transactions of the Royal society B: biological sciences*, 371 (1705).
- Tymofiyeva, O., Connolly, C. G., Ho, T. C., Sacchet, M. D., Blom, E. H., LeWinn, K. Z., Xu, D. & Yang, T. T. (2017). DTI-based connectome analysis of adolescents with major depressive disorder reveals hypoconnectivity of the right caudate. *Journal of affective disorders*, 207, 18-25.
- Tyng, C. M., Amin, H. U., Saad, M. N. M. & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in psychology*, 8 (1454).
- Uspensky, V. A. (2001). Why Kolmogorov complexity? Dans Goles, E. & Martinez, S. *Nonlinear phenomena, and complex systems*, 6, Springer, Dordrecht.
- Utevsky, A. V., Smith, D. V. & Huettel, S. A. (2014). Precuneus is a functional core of the default-mode network. *Journal of neuroscience*, 34 (3), 932-940.

Uttal, W. R. (2001). *The new phrenology – the limits of localizing cognitive processes in the brain*. MIT press, 255p.

Vaidya, A. R., Pujara, M. S., Petrides, M., Murray, E. A. & Fellows, L. K. (2019). Lesion studies in contemporary neuroscience. *Trends in cognitive science*, 23 (8), 653-671.

Van der Schyff, D., Schiavio, A., Walton, A., Velardo, V. & Chemero, A. (2018). Musical creativity and the embodied mind: exploring the possibilities of 4E cognition and dynamical systems theory. *Music & Science*, 1.

Varga, S. (2018). *Scaffolded minds: integration and disintegration*. Cambridge: The MIT press.

Vassal, F., Schneider, F., Boutet, C., Jean, B., Sontheimer, A. & Lemaire, J.-J. (2016). Combined DTI tractography and functional MRI study of the language connectome in healthy volunteers: extensive mapping of white matter fascicles and cortical activations. *PLoS One*, 11 (3).

Wade, J. & Heydari, B. (2014). Complexity: definition and reduction techniques, some simple thoughts on complex systems. [en ligne] <https://www.semanticscholar.org/paper/Complexity%3A-Definition-and-Reduction-Techniques%2C-on-Wade-Heydari/0c44bbd9ef9ddec8e268009b50243159af7793d7> (consulté le 05/01/2022).

Watts, A. G. (2007). Anatomy of the HPA axis. Dans Fink, G. *Encyclopedia of stress*, Academic press, 166-170.

Wegner, D. M. (1986). Transactive memory: a contemporary analysis of the group mind. Dans Mullen, B. & Goethals, G. R. *Theories of group behavior*, Springer, New York, 185-208.

Wendelken, C., O'Hare, E. D., Whitaker, K. J., Ferrer, E. & Bunge, S. A. (2011). Increased functional selectivity over development in rostral lateral prefrontal cortex. *Journal of neuroscience*, 31 (47), 17260-17268.

Wilkinson, D. & Halligan, P. W. (2004). Opinion - the relevance of behavioral measures for functional-imaging studies of cognition. *Nature reviews neuroscience*, 5 (1), 67-73.

Wilson, V. & Peper, E. (2004). The effects of upright and slumped postures on the recall of positive and negative thoughts. *Applied psychophysiology and biofeedback*, 29 (3), 189-195.

Wimsatt, W. C. (1986). Forms of aggregativity. Dans Grene, M. G. et al. *Human nature and natural knowledge*, Dordrecht, Reidel, 259-291.

Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology & philosophy*, 25 (3), 287-318.

Woodward, J. (2015). Interventionism and causal exclusion. *Philosophy and phenomenological research*, 91 (2), 303-347.

Yoo, J., Kwon, J. & Choe, Y. (2014). Predictable internal brain dynamics in EEG and its relation to conscious states. *Frontiers in neurobotics*, 8 (18).

- Yuan, B., Fang, Y., Han, Z., Song, L., He, Y. & Bi, Y. (2017). Brain hubs in lesion models: predicting functional network topology with lesion patterns in patients. *Scientific reports*, 7.
- Yuh, E. L. (2017). Structural imaging of traumatic brain injury. Dans Winn, H. R. (dir.) *Youmans and Winn neurological surgery*, 344, 2819-2836.
- Zarahn, E., Aguirre, G.K. & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage*, 5, 179-97.
- Zelano, C., Jiang, H., Zhou, G., Arora, N., Schuele, S., Rosenow, J. & Gottfried, J. A. (2016). Nasal respiration entrains human limbic oscillations and modulates cognitive function. *The journal of neuroscience*, 36 (49), 12448-12467.
- Zenil, H. & Tegner, J. (2014). Information theory, Kolmogorov complexity and algorithmic probability in network biology. [en ligne] https://www.researchgate.net/publication/259742981_Information_Theory_Kolmogorov_Complexity_and_Algorithmic_Probability_in_Network_Biology (consulté le 28/08/2021).
- Zhang, L., Tang, X., Wang, C., Ding, D., Zhu, J., Zhou, Y., Diao, S., Kong, Y., Cai, X., Li, C., Yao, Y. & Fang, Q. (2021). Predictive model of dysphagia and brain lesion-symptom mapping in acute ischemic stroke. *Frontiers in aging neuroscience*, 13.
- Zhang, Q., Van Vugt, M., Borst, J. P. & Anderson, J. R. (2018). Mapping working memory retrieval in space and in time: a combined electroencephalography and electrocorticography approach. *NeuroImage*, 174, 472-484.
- Zhao, C., Kaulakis, R., Morgan, J. H., Hiam, J. H., Sanford, J. P. & Ritter, F. E. (2012). Socio-cognitive networks: modeling the effects of space and memory on generative social structures. 21st Behavior Representation in Modeling & Simulation (BRIMS), conference, Amelia island.
- Zhou, J., Neale, J. H., Pomper, M. G. & Kozikowski, A. P. (2005). NAAG peptidase inhibitors and their potential for diagnosis and therapy. *Nature reviews drug discovery*, 4, 1015-1026.
- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C. & Sun, M. (2020). Graph neural networks: a review of methods and applications. *AI open*, 1, 57-81.