UNIVERSITÉ DU QUÉBEC À MONTRÉAL

UNE ALTERNATIVE ÉNACTIVE ET OBJECTIVE À LA CONCEPTION DU TROUBLE MENTAL DES RESEARCH DOMAIN CRITERIA

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN PHILOSOPHIE

PAR

SIMON GOYER

AOÛT 2022

UNIVERSITÉ DU QUÉBEC À MONTRÉAL Service des bibliothèques

Avertissement

La diffusion de cette thèse 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

Je tiens particulièrement à témoigner ma reconnaissance à Mona Gupta, professeure de psychiatrie à l'UdeM, et à Luc Faucher, professeur de philosophie à l'UQAM, qui ont codirigé ma thèse de doctorat. Mona m'a fait connaitre le monde de la recherche (cueillette de données, analyse de données, conférences, organisation d'une journée scientifique, rédaction d'un article avec elle et Nancy Nyquist Potter, rencontres avec des chercheuses et chercheurs, enseignement à des résidents en psychiatrie, etc.). En travaillant avec elle au CRCHUM, j'ai pris davantage confiance en moi. Je la remercie également pour toutes les discussions philosophiques et scientifiques sur la psychiatrie que j'ai eues avec elle. Celles-ci m'ont beaucoup aidé dans la rédaction de ma thèse. Quant à Luc, je le remercie pour les critiques constructives relatives au contenu de ma thèse. J'ai aussi rédigé avec lui quatre articles/chapitres de philosophie. Je lui en suis grandement reconnaissant. Je suis aussi heureux d'avoir été son auxiliaire de recherche. J'ai pu, à cette occasion, participer à l'élaboration du contenu d'un séminaire en philosophie de la psychiatrie. J'ai aussi pu donner quelques capsules dans un cours qu'il donnait avec Pierrre Poirier. Je remercie d'ailleurs ce dernier et Christophe Malaterre, tous les deux professeurs de philosophie à l'UQAM, pour leurs critiques constructives du contenu de ma thèse.

Je suis aussi redevable envers Mauro Rossi, lui aussi professeur de philosophie à l'UQAM, qui m'a fait part de deux conseils — l'un concernant la rédaction des textes philosophiques et l'autre relatif à un texte que j'avais écrit pour une de mes demandes de bourse — qui m'ont été fort utiles.

Je remercie aussi les membres du *Laboratoire étudiant interuniversitaire en philosophie des sciences* (LEIPS) pour toutes les discussions que j'ai eues avec elles et avec eux lors de nos réunions et de nos rencontres informelles. J'en garde de très beaux souvenirs!

De plus, je remercie les membres du *Café L'Étincelle* où j'ai écrit une partie de ma thèse. Leur bonne humeur est contagieuse!

Je remercie, de plus, le *Conseil de recherches en sciences humaines du Canada* (CRSH) pour la Bourse doctorale d'études supérieures du Canada Joseph-Armand-Bombardier dont j'ai pu bénéficier de 2015 à 2018. Cette bourse d'études m'a grandement aidé financièrement lors de la rédaction de ma thèse. J'ai également pu coécrire quatre articles et un chapitre de livre pendant cette période de financement :

- Nyquist Potter, N., Gupta, M. et Goyer, S. (2019). Diagnostic reasoning in psychiatry: Acknowledging an explicit role for intersubjective knowing. *Philosophy, Psychiatry, & Psychology*, 26(1), p. 49-64.
- Faucher, L. et Goyer, S. (2017). RDoC's special kind of reductionism and its possible impact on clinical psychiatry. Dans M. Johnson, L. Syd et S. Rommelfanger, K. (dir.), *The Routledge handbook of neuroethics* (p. 412-428). New York, NY: Routledge.
- Faucher, L. et Goyer, S. (2017). Research Domain Criteria. Dans A. Wenzel (dir.), *The SAGE encyclopedia of abnormal and clinical psychology* (p. 2862-2863). Thousand Oaks, CA: SAGE Publications.
- Faucher, L. et Goyer, S. (2016). Le Research Domain Criteria (RDoC), le réductionnisme et la psychiatrie clinique. *Revue de Synthèse*, 137(1-2), p. 117-149.

• Faucher, L. et Goyer, S. (2015). RDoC: Thinking outside the DSM box without falling into a reductionist trap. Dans Demazeux, S. et Singy, P. (dir.), *The DSM-5 in perspective: Philosophical reflections on the psychiatric babel* (p. 199-224). New York, NY: Springer (Série « History, Philosophy and Theory of the Life Sciences »).

Je remercie aussi grandement la Faculté des sciences humaines de l'UQAM pour la Bourse doctorale d'excellence (volet recrutement) que j'ai gagnée en 2014, de même que le Centre interuniversitaire de recherche sur la science et la technologie (CIRST), le Groupe de recherche interuniversitaire sur la normativité (GRIN) et le Centre de recherche en éthique (CRÉ) de qui j'ai aussi obtenu des bourses d'études supérieures et où j'ai eu l'occasion de rencontrer des personnes passionnées par le monde des connaissances et la réflexion.

Enfin, je dis merci du fond du cœur à tous ceux qui étaient là dans ma vie quotidienne lors de la rédaction de cette thèse. Merci spécialement à mes parents Suzanne et Normand d'être présents et de me soutenir sur les chemins de l'existence. Vous êtes très importants pour moi. Je vous aime. J'espère que nous aurons encore beaucoup de soupers ensemble! Merci à mon frère Matthieu pour ses blagues et son amitié. Je t'aime frérot! Merci à Marlène pour tout son amour, son soutien émotif et son soutien financier. Je remercie également Marlène pour les innombrables discussions sur les notions abordées dans ma thèse que nous avons eues tous les deux. Marlène, je suis tellement heureux que nous soyons ensemble dans cette aventure qu'est la vie! Je t'aime! Pour finir, je salue chaleureusement Bastien que Marlène et moi accueillons avec tout notre amour dans ce monde!



TABLE DES MATIÈRES

LIS	TE DES	S ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES	ζi
List	e des sy	mboles et des unités	ΚV
RÉS	SUMÉ .	X	V
ABS	STRAC	Txv.	111
INT	RODU	CTION	. 1
		E I Research Domain Criteria (RDoC) et interprétation philosophique de on du trouble mental	
1.1	Le N	ational Institute of Mental Health (NIMH), son plan stratégique et les	
1.2	Qu'e	st-ce qui motive le NIMH à mettre sur pied les RDoC?	11
1.3	Que s	sont les RDoC?	16
1.4	Quels	s types de recherches sont menées au sein des RDoC?2	20
	1.4.1 1.4.2	Cartographie et manipulation des circuits neuraux responsables des dysfonctions émotives et cognitives dans l'anxiété et la dépression	
1.5	Interp	orétation philosophique de la conception du trouble mental des RDoC2	
	1.5.1 1.5.2 1.5.3	Conception du trouble mental et matrice des RDoC	34 la 50
	1.5.41.5.5	Le principe épistémoquantitatif (et la variante quantitative du réalisme des entités basé sur la robustesse d'Eronen)	63
1.6	Conc	lusion du premier chapitre	81

CH	APITRI	E II Critique de la conception du trouble mental des RDoC	83
2.1 rech		pothèse de l'identité psychoneurale de Bunge est problématique et nuit psychiatrique	
	2.1.1 2.1.2 2.1.3 2.1.4	Incrustation environnementale des mécanismes mentaux	94 104
		incipe épistémoquantitatif sous-jacent aux RDoC nuit à la recherche ue	117
	2.2.1	Ce principe épistémoquantitatif empêche de faire des études	
	2.2.2	phénoménologiques	es
2.3	La TI	BS a de la difficulté à saisir la normativité dans le monde du vivant	144
2.4 (DC	•	rsfonction préjudiciable (DP) et la dysfonction correctement dévalorisées alternatives à la TBS?	
	2.4.1 2.4.2 2.4.3	La DP de Wakefield	157
2.5	Conc	lusion du deuxième chapitre	163
(CÉ	EOST):	E III Une conception énactive et objective de la santé et du trouble une solution à quatre problèmes de la conception du trouble mental des	
3.1	La co	onception énactive autopoïétique de la cognition (CÉAC)	165
	3.1.1	La CÉAC résout les trois premiers problèmes de la conception du troumental des RDoC	
		ÉOST (basée sur la CÉAC) et la solution au quatrième problème de la du trouble mental des RDoC	
	3.2.1 3.2.2	Présentation de la CÉOST La CÉOST résout le quatrième problème de la conception du trouble mental des RDoC	
3.3	Conc	lusion du troisième chapitre	233
CO	NCLUS	SION	234

	Origine du concept d'institution cognitive (et de la théorie de l'espr tendu)	
ANNEXE B	Sur la théorie des systèmes dynamiques	. 262
ANNEXE C	Émergence et énactivisme autopoïétique	. 265
	Quelques brèves réflexions sur les modèles énactifs de de Haar Ward et de Fuchs	-
RÉFÉRENCE	SS	. 286

LISTE DES FIGURES

Figu	Page
1.1	Matrice des RDoC contenant les aspects environnemental et développemental (NIMH, 2021a)
1.2	Modèle mécaniste du phénomène de la transmission synaptique (Moyes et Schulte, 2006, chapitre 5)
1.3	Les trois stades métaboliques de la respiration cellulaire (Campbell et Reece, 2012, p. 188)
1.4	Loi normale (Ingrand, 2017)60
1.5	Modèle neurobiocognitif de la dépression majeure basé sur le modèle cognitif de Beck (Disner <i>et al.</i> , 2011, p. 475)
2.1	Outil informatique d'auto-évaluation topographique emBODY (Nummenmaa et al., 2014, p. 647)96
2.2	Topographie corporelle des émotions basiques (en haut) et non basiques (en bas) (Nummenmaa <i>et al.</i> , 2014, p. 647)
3.1	Mécanisme autopoïétique nécessaire à l'existence d'un système vivant (Varela 1997, p. 75)
3.2	Cellules de Bénard vues de côté (Cellules de Bénard, 2021)
3.3	Cellules de Bénard vues d'en haut (Deacon et Koutroufinis, 2014, p. 414) 195
A.1	Modèle en sandwich de l'esprit (inspiré de Hurley, 2001, p. 9)

A.2	Exemple de rotation mentale faite sur une forme en trois dimensions (Shepard et	t
	Metzler, 1971, p. 702)247	1

LISTE DES TABLEAUX

Tab	leau Page
1.1	Matrice des RDoC (adaptée et traduite de NIMH, 2021b)30
2.1	Sondage fictif sur la perception qu'ont les patients de la qualité des soins qu'ils reçoivent de la part du personnel soignant travaillant dans une unité psychiatrique (inspiré de Saldana, 2011, p. 11)

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

ADN Acide désoxyribonucléique

AOF Approche organisationnelle des fonctions (de Mossio, Saborido et

Moreno)

AOFm Approche organisationnelle des fonctions modifiée (approche de

Mossio, Saborido et Moreno modifiée)

APA American Psychiatric Association

BCSC BioCognitifsSocioCulturel (cet acronyme est associé au terme

« système » ou au terme « mécanisme »)

CAMH Centre de toxicomanie et de santé mentale

CBT Cognitive behavioral therapy

CÉAC Conception énactive autopoïétique de la cognition

CÉOS Conception énactive et objective de la santé

CÉOST Conception énactive et objective de la santé et du trouble

CÉOT Conception énactive et objective du trouble

COVID-19 Maladie à coronavirus

DCD Dysfonction correctement dévalorisée (de Powell et Scarffe)

DM Dépression majeure

DP Dysfonction préjudiciable (de Jerome Wakefield)

DSM Manuel diagnostique et statistique des troubles mentaux

EEG Électroencéphalogramme

ERC Essai randomisé contrôlé

ESI Externalizing Spectrum Inventory

ÉSR Épicerie solidaire de Rosemont

IRMf Imagerie par résonance magnétique fonctionnelle

HHS Axe hypothalamo-hypophyso-surrénalien

HHS U.S. Department of Health and Human Services

LGBTQ Lesbiennes, gais, bisexuelles, transgenres et queers

NIH National Institutes of Health

NIMH National Institute of Mental Health

NTN Norme téléologique naturelle

OS Objectifs stratégiques

RDoC Research Domain Criteria

RE Réalisme des entités [entity realism]

RER Réalisme des entités basé sur la robustesse [robustness-based entity]

realism]

SLA Sclérose latérale amyotrophique

SPT Supportive Psychotherapy

spTMS Stimulation magnétique transcrânienne non invasive

SSPS Syndrome de stress post-traumatique

TBS Théorie biostatistique de la santé (de Christopher Boorse)

TCC Thérapie cognitivocomportementale

TDAH Trouble déficitaire de l'attention avec ou sans hyperactivité

TEM The Embodied Mind: Cognitive Science and Human Experience (livre

de Varela, Thompson et Rosch paru en 1991)

TGP Théorie générale des processus (de Johanna Seibt)

TOC Trouble obsessionnel-compulsif

TPL Trouble de la personnalité limite

TSD Théorie des systèmes dynamiques

TSPT Trouble de stress post-traumatique

VIH-1 Virus de l'immunodéficience humaine

LISTE DES SYMBOLES ET DES UNITÉS

 $\forall x$ Pour tout x

Ca²⁺ Ion de calcium

= Signe « égal »

RÉSUMÉ

En 2009, le *National Institute of Mental Health* (NIMH), le principal institut public responsable du financement de la recherche sur les désordres mentaux aux États-Unis, a mis sur pied les *Research Domain Criteria* (RDoC), un cadre de recherche pour la psychiatrie. En soutenant des études empiriques dans ce dernier, le NIMH espère améliorer notre compréhension des troubles mentaux et contribuer au développement de traitements nouveaux et efficaces.

Les RDoC sont-ils un bon cadre pour faire de la recherche en psychiatrie? Je soutiens que non. Plus précisément, dans ma thèse, je soutiens deux choses. La première est que le plus gros problème des RDoC est sa conception du trouble mental. La deuxième chose est que cette dernière devrait être remplacée par une conception énactive et objective de la santé et du trouble (CÉOST). J'articule cette double thèse tout au long des trois chapitres qui composent ce travail doctoral.

Le chapitre 1 me sert à présenter les RDoC (sections 1.1-1.4) et à mettre en lumière ce qui me semble être, d'une part, la conception du trouble mental qui leur est inhérente (section 1.5) et, d'autre part, les hypothèses, théories et principes philosophiques qui forment la base de cette dernière, à savoir une certaine version à la fois pluraliste et réductionniste du néomécanisme de Bechtel (section 1.5.2), une interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Boorse (section 1.5.3), le principe épistémoquantitatif (qui suppose ce que j'ai appelé la « variante quantitative » du réalisme des entités basé sur la robustesse d'Eronen) (section 1.5.4) et l'hypothèse de l'identité psychoneurale formulée par Bunge (section 1.5.5).

Le chapitre 2 contient ma critique de la conception du trouble mental des RDoC. Selon mon analyse, principalement quatre problèmes affectent cette dernière et nuisent à la recherche psychiatrique. Le premier problème est que l'hypothèse de l'identité psychoneurale sous-jacente à la conception du trouble mental des RDoC est inadéquate (section 2.1). En effet, en plus d'être généralement insensible au fait que les mécanismes mentaux sont fortement incrustés dans l'environnement (physique, social et culturel) (section 2.1.1), elle est incapable de saisir le fait que la cognition soit

constituée de mécanismes corporels (section 2.1.2), multipersonnels (2.1.3) et cognitivo-institutionnels (section 2.1.4).

Le deuxième problème est que l'hypothèse de l'identité psychoneurale nuit à la recherche psychiatrique (section 2.1). En effet, si cette hypothèse ne saisit pas l'incrustation de la cognition dans l'environnement (physique, social et culturel) et sa composition extraneurale, alors (a) elle ne peut pas favoriser des recherches sérieuses sur les mécanismes physiques, sociaux et culturels propices au développement des troubles mentaux (section 2.1.1) et (b) elle ne peut pas favoriser des recherches sur les mécanismes corporels (section 2.1.2), multipersonnels (2.1.3) et cognitivo-institutionnels (section 2.1.4) constituant les troubles mentaux. Cela est regrettable parce que ces recherches ont un intérêt épistémique et pragmatique incontestable.

Le troisième problème est dû au principe épistémoquantitatif que je pense être sousjacent aux RDoC (section 2.2). En effet, comme ce principe suppose que les phénomènes non quantifiables n'ont pas d'intérêt épistémique, alors il empêche que soient menées, au sein des RDoC, des recherches phénoménologiques (section 2.2.1) et qualitatives (section 2.2.2). Cela est malheureux parce que ces recherches ont une pertinence épistémique et des retombées pratiques prometteuses.

Le quatrième problème de la conception du trouble mental des RDoC est dû au fait que la TBS qui lui est sous-jacente (interprétée de manière adaptationniste ou non) est incapable de saisir la normativité dans les systèmes vivants et, conséquemment, dans les personnes (section 2.3). J'examine à la section 2.4 si l'utilisation de la dysfonction préjudiciable (DP) de Wakefield (section 2.4.1) ou la dysfonction correctement dévalorisée (DCD) de Powell et Scarffe (section 2.4.2) pourrait régler ce problème que rencontre la TBS et montre que ce n'est pas le cas (section 2.4.3).

Le chapitre 3 est l'endroit où je m'engage dans la partie constructive de mon argumentaire. J'y propose, en effet, une solution aux quatre problèmes susmentionnés (sections 3.1.1 et 3.2.2). Je crois que ces problèmes sont tous résolus par une conception énactive et objective de la santé et du trouble (CÉOST) (section 3.2.1), laquelle a pour base la conception énactive autopoïétique de la cognition (CÉAC) (section 3.1).

Mots-clés: Research Domain Criteria, néomécanisme, théorie biostatistique de la santé, principe épistémoquantitatif, réalisme des entités basé sur la robustesse, hypothèse de l'identité psychoneurale, composantes extraneurales de l'esprit, recherche neurocognitive, recherche phénoménologique, recherche qualitative, normativité, dysfonction préjudiciable, dysfonction correctement dévalorisée, énactivisme autopoïétique, conception énactive et objective de la santé et du trouble

ABSTRACT

In 2009, the *National Institute of Mental Health* (NIMH), the principal public body responsible for funding research on mental disorders in the United States, established the *Research Domain Criteria* (RDoC), a framework for psychiatric research. By supporting empirical studies conducted within this framework, the NIMH hopes to improve our understanding of mental disorders and to contribute to the development of new and effective treatments.

Is RDoC a good framework for doing psychiatric research? I argue that it is not. Specifically, in my thesis, I make two related arguments. First, RDoC's conception of mental disorder is fundamentally flawed. Second, RDoC should be replaced by an enactive and objective conception of health and disorder (EOCHD). I articulate this dual thesis throughout the three chapters that comprise this doctoral work.

Chapter 1 serves to introduce RDoC (sections 1.1-1.4) and to highlight on the one hand its conception of mental disorder (section 1.5) and on the other, the philosophical assumptions, theories, and principles that form the basis of this conception. These include a certain version of Bechtel's neomechanism that is both pluralist and reductionist (section 1.5.2), an adaptationist interpretation of Boorse's biostatistical theory (BST) of health (section 1.5.3), the epistemo-quantitative principle (which assumes what I have called the "quantitative variant" of Eronen's robustness-based entity realism) (section 1.5.4), and Bunge's formulation of the psychoneural identity hypothesis (section 1.5.5).

Chapter 2 contains my critique of RDoC's conception of mental disorder. According to my analysis, there are four principal problems with the RDoC concept of mental disorder that are detrimental to psychiatric research. The first problem is that the psychoneural identity assumption underlying the RDoC mental disorder construct is inadequate (section 2.1). Indeed, in addition to being generally insensitive to the fact that mental mechanisms are strongly embedded in the environment (physical, social and cultural) (section 2.1.1), it is unable to capture the fact that cognition is made up of corporeal (section 2.1.2), multipersonal (2.1.3) and cognitive-institutional (section 2.1.4) mechanisms.

The second problem is that the psychoneural identity hypothesis undermines psychiatric research (section 2.1). Indeed, if this hypothesis does not capture the embedding of cognition in the environment (physical, social and cultural) and its extraneural composition, then (a) it cannot foster serious research on the physical, social and cultural mechanisms that contribute to the development of mental disorders (section 2.1.1) and (b) it cannot foster research on the corporeal (section 2.1.2), multipersonal (2.1.3) and cognitive-institutional (section 2.1.4) mechanisms that constitute mental disorders. This is unfortunate because this research has undeniable epistemic and pragmatic value.

The third problem is due to the epistemo-quantitative principle that I believe underlies RDoC (section 2.2). Indeed, since this principle assumes that non-quantifiable phenomena are not of epistemic interest, then it prevents phenomenological (section 2.2.1) and qualitative (section 2.2.2) research from being conducted within RDoC. This is unfortunate because such research has epistemic relevance and promising practical implications.

The fourth problem with RDoC's conception of mental disorder is that the underlying BST (whether interpreted from an adaptationist perspective or not) is unable to capture normativity in living systems and, consequently, in people (section 2.3). I examine in section 2.4 whether the use of Wakefield's harmful dysfunction (HD) (section 2.4.1) or Powell and Scarffe's properly devalued dysfunction (PDD) (section 2.4.2) could address this problem faced by the BST and show that they do not (section 2.4.3).

Chapter 3 is where I engage in the constructive part of my argument. In it, I propose a solution to the four problems mentioned above (sections 3.1.1 and 3.2.2). I believe that these problems are all solved by an enactive and objective conception of health and disorder (EOCHD) (section 3.2.1), which is based on the enactive autopoietic conception of cognition (CEAC) (section 3.1).

Keywords: Research Domain Criteria, neomechanism, Biostatistical Theory of Health, epistemoquantitative principle, robustness-based entity realism, psychoneural identity hypothesis, extraneural components of the mind, neurocognitive research, phenomenological research, qualitative research, normativity, harmful dysfunction, properly devalued dysfunction, autopoietic enactivism, enactive and objective conception of health and disorder

INTRODUCTION

En 2009, le *National Institute of Mental Health* (NIMH), le principal institut public responsable du financement de la recherche sur les désordres mentaux aux États-Unis, a mis sur pied les *Research Domain Criteria* (RDoC), un cadre de recherche pour la psychiatrie. En soutenant des études empiriques menées dans ce dernier, le NIMH espère améliorer notre compréhension des troubles mentaux et contribuer au développement de traitements nouveaux et efficaces.

Au cœur des RDoC se trouve une conception des troubles mentaux. Thomas R. Insel, l'ancien directeur du NIMH, la formule de la manière suivante : « Les troubles mentaux sont des troubles biologiques mettant en jeu des circuits cérébraux responsables de domaines spécifiques de la cognition, des émotions ou du comportement. » (Insel, 2013b)¹

Sur le site Internet du NIMH, on retrouvait en 2019 (mais on ne retrouve plus en 2022) une formulation légèrement différente de celle d'Insel qui allait comme suit : « les maladies mentales sont des troubles cérébraux exprimés sous forme de syndromes

¹ Sauf avis contraire, toutes les traductions de l'anglais vers le français sont de moi. Pour m'aider dans cette tâche, j'utilise des outils comme *Google Tradution* ou *Deepl*. Mais la traduction finale est la mienne.

complexes cognitifs, émotionnels et comportementaux/sociaux.» (NIMH, 2019a) Bref, au NIMH les troubles mentaux sont conçus comme des troubles du cerveau².

Les RDoC sont-ils un bon cadre pour faire de la recherche en psychiatrie ? Je soutiens que non. Plus précisément, dans ma thèse, je soutiens deux choses. La première est que le plus gros problème des RDoC est sa conception du trouble mental. La deuxième chose est que cette dernière devrait être remplacée par la conception énactive et objective de la santé et du trouble (CÉOST). J'articule cette double thèse tout au long des trois chapitres qui composent ce travail doctoral.

Le chapitre 1 me sert à présenter les RDoC (sections 1.1-1.4) et à en faire une analyse philosophique (section 1.5). À la section 1.1, je présente brièvement le NIMH, sa mission, son plan stratégique pour la recherche et le principal rôle qu'y jouent les RDoC, à savoir celui de développer une classification des troubles mentaux utile pour la recherche (plus précisément utile, comme je le dis précédemment, pour mieux comprendre les troubles mentaux et développer de nouveaux traitements). À la section 1.2, j'explique ce qui motive le NIMH à mettre sur pied les RDoC. Trois raisons le poussent dans cette entreprise. La première est qu'il souhaite libérer la recherche du carcan que sont les catégories du *Manuel diagnostique et statistique des troubles mentaux* (DSM) qu'il juge (comme plusieurs) insuffisamment valides (Cuthbert, 2015, p. 91-92; à ce sujet, voir aussi Cuthbert et Insel, 2013; Cuthbert et Kozak, 2013; Hyman, 2007, 2010; Sanislow *et al.*, 2019, p. 779; Simmons et Quinn, 2014). La seconde raison

² Kirk, Gomory et Cohen (2015, p. 1) affirment que l'*American Psychiatric Association* (APA) soutient aussi l'idée selon laquelle les troubles mentaux sont des troubles du cerveau. C'est d'ailleurs ce que soutenait Paul Appelbaum, un ancien directeur de l'APA, dont on peut lire les propos dans le *LA Times*: « Nos cerveaux sont des organes biologiques de par leur nature même » [...] « Tout désordre est par essence un processus biologique. » (Davis, 2003) Cette similarité entre les conceptions de l'APA et du NIMH implique que ma thèse pourrait contribuer plus largement à une réforme qui pourrait toucher une plus grande partie de la psychiatrie. Cependant, j'ai choisi de traiter seulement de la conception que l'on retrouve dans les RDoC.

est qu'il veut développer une psychiatrie clinique de précision (Insel, 2014, p. 396; voir aussi Insel et Cuthbert, 2015; Sanislow et al., 2019, p. 781). La troisième raison est qu'il y a, en ce moment, une crise de la santé mentale aux États-Unis et que le NIMH pense que les RDoC pourraient aider à l'atténuer. À la section 1.3, je présente les RDoC de manière détaillée. J'y explique que ceux-ci sont un cadre conceptuel composé de cinq principes organisés en système. Ces principes permettent d'apprendre ce que sont les troubles mentaux, la manière dont il faut les expliquer et la façon dont on doit les étudier. À la section 1.4, je présente les recherches plutôt neuro-centrées et quantitatives qui sont menées dans le cadre des RDoC. Je donne deux exemples de projets de recherche de ce type (sections 1.4.1 et 1.4.2). La section 1.5 est centrale dans mon argumentaire. En effet, c'est dans cette dernière que je commence mon analyse de ce que je pense être la conception du trouble mental inhérente aux RDoC. J'explique que, si les troubles mentaux y sont concus comme des troubles du cerveau, il serait plus précis de dire qu'ils sont appréhendés comme des dysfonctions affectant un ou des domaines ou sous-domaines de fonctionnement de l'esprit/cerveau que l'on retrouve dans la matrice des RDoC que je présente à la même occasion (section 1.5.1). Après avoir fait cette précision, et à la lumière des écrits des défenseurs des RDoC et de ceux de leurs critiques, je poursuis cette présentation de la conception du trouble mental des RDoC en exposant ce qui m'apparait être les hypothèses, conceptions ou théories philosophiques qui lui sont sous-jacentes. À la section 1.5.2, je présente le néomécanisme de William Bechtel (Bechtel, 2008; Bechtel et Abrahamsen, 2005; Wright et Bechtel, 2007). En effet, les tenants des RDoC appréhendent les troubles mentaux comme des mécanismes psychoneuraux dysfonctionnels. C'est aussi souvent en termes mécanistes que les chercheuses et les chercheurs œuvrant au sein des RDoC conçoivent leurs hypothèses et leurs protocoles expérimentaux. Cela dit, il est toutefois important de mentionner que les partisans des RDoC adoptent une version à la fois pluraliste et réductionniste du néomécanisme de Bechtel. Cette dernière ressemble au réductionnisme fragmentaire [patchy reductionism] de Kenneth F. Schaffner (2006,

2008, 2011, 2013, 2016, chapitre 5) (j'expliquerai pourquoi cela est le cas à la fin de la section 1.5.2). À la section 1.5.3, j'expose une interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Christopher Boorse (1975, 1976a, 1977, 1997, 2014). Je soutiens, citations à l'appui, que le concept de dysfonction de cette interprétation de la TBS est sous-jacent à la conception du trouble mental des RDoC. À la section 1.5.4, je fais connaître le principe epistémoquantitatif sous-jacent à la conception du trouble mental des RDoC selon lequel seulement ce qui peut être mesuré quantitativement a une pertinence épistémique. Je montre également que la conception du progrès en science psychiatrique inhérente à ce principe suppose ce que j'ai appelé la « variante quantitative » du réalisme des entités basé sur la robustesse [robustness-based entity realism] (RER) développé par Markus Eronen (2015, 2019a, 2019b). Enfin, à la section 1.5.5, je présente l'hypothèse de l'identité psychoneurale formulée par Mario Bunge (2010, p. 161). En effet, comme le suppose cette hypothèse, les tenants des RDoC identifient les mécanismes mentaux dysfonctionnels à des mécanismes cérébraux dysfonctionnels.

Le chapitre 2 contient ma critique de la conception du trouble mental des RDoC. Selon mon analyse, principalement quatre problèmes affectent cette dernière et nuisent à la recherche psychiatrique. Le premier problème est que l'hypothèse de l'identité psychoneurale sous-jacente à la conception du trouble mental des RDoC est inadéquate (section 2.1). En effet, en plus d'être généralement insensible au fait que les mécanismes mentaux sont fortement incrustés dans l'environnement (physique, social et culturel) (section 2.1.1), elle est incapable de saisir le fait que la cognition soit constituée de mécanismes corporels (section 2.1.2), multipersonnels (2.1.3) et cognitivo-institutionnels (section 2.1.4).

Le deuxième problème est que l'hypothèse de l'identité psychoneurale nuit à la recherche psychiatrique (section 2.1). En effet, si cette hypothèse ne saisit pas

l'incrustation de la cognition dans l'environnement (physique, social et culturel) et sa composition extraneurale, alors (a) elle ne peut pas favoriser des recherches sérieuses sur les mécanismes environnementaux (physiques, sociaux et culturels) propices au développement des troubles mentaux (section 2.1.1) et (b) elle ne peut pas favoriser des recherches sur les mécanismes corporels (section 2.1.2), multipersonnels (2.1.3) et cognitivo-institutionnels (section 2.1.4) *constituant* les troubles mentaux. Cet appauvrissement de la diversité des recherches qu'implique l'adoption de la conception du trouble mental des RDoC est malheureux parce que, non seulement il nuit à notre compréhension des troubles mentaux, mais aussi à notre capacité d'intervenir sur eux.

Le troisième problème est dû au principe épistémoquantitatif que je pense être sousjacent aux RDoC (section 2.2)³. En effet, comme ce principe suppose que les phénomènes non quantifiables n'ont pas de valeur épistémique, alors il empêche que soient menées au sein des RDoC des recherches phénoménologiques (section 2.2.1) et qualitatives (section 2.2.2). Il est regrettable qu'il en soit ainsi parce que ce type d'enquêtes, comme je le montrerai, enrichit notre compréhension des troubles mentaux et permet de concevoir des manières autres que pharmacologiques d'intervenir sur eux.

Le quatrième problème de la conception du trouble mental des RDoC est dû au fait que la TBS de Boorse qui lui est sous-jacente est incapable de saisir la normativité dans les systèmes vivants et, conséquemment, dans les personnes (section 2.3). À la section 2.4, j'examine si l'utilisation de la dysfonction préjudiciable (DP) de Wakefield (section 2.4.1) ou de la dysfonction correctement dévalorisée (DCD) de Powell et Scarffe (section 2.4.2) pourrait régler ce problème que rencontre la TBS et montre que

³ Ce troisième problème est aussi dû au fait que les tenants des RDoC adoptent ce qui me semble être une interprétation quantitative du RER d'Eronen (2015, 2019a, 2019b). Cette version du RER est inhérente à la conception du progrès scientifique supposée par le principe épistémoquantitatif qu'appliquent les défenseurs des RDoC. Je reviendrai sur cela à la section 1.5.4.

ce n'est pas le cas (section 2.4.3). La DP et la DCD sont, en effet, comme la TBS, incapables de saisir la normativité dans les systèmes vivants. Parce qu'il faut qu'une conception du trouble mental soit capable de distinguer le normal du pathologique, et qu'elle ne peut pas le faire si elle n'est pas en mesure de saisir la normativité dans les systèmes vivant, je soutiens que la conception du trouble mental des RDoC est inadéquate, et ce, peu importe qu'elle ait pour base la TBS (comme c'est à mon avis le cas en ce moment), la DP ou la DCD.

Le chapitre 3 est l'endroit où je m'engage dans la partie constructive de mon argumentaire. J'y propose, en effet, une solution aux quatre problèmes susmentionnés. Je crois que ces problèmes sont tous résolus par la conception énactive et objective de la santé et du trouble (CÉOST). La CÉOST met à profit des idées issues des travaux en biologie théorique qu'ont menés Alvaro Moreno, Matteo Mossio et Cristian Saborido (Mossio et al., 2009, 2010; Mossio et Moreno, 2015; Saborido et Moreno, 2015; Saborido et al., 2016). La CÉOST suppose également la conception énactive autopoïétique de la cognition (CÉAC). Je présente cette dernière à la section 3.1. Ensuite, à la section 3.1.1, je montre pourquoi la CÉAC (et, par conséquent, la CÉOST qui repose sur elle) résout les trois premiers problèmes susmentionnés. Le premier écueil que rencontre la conception du trouble mental des RDoC est résolu parce que la CÉAC intègre le fait que la cognition soit incrustée dans l'environnement et le fait qu'elle soit constituée de mécanismes extraneuraux. En conséquence, la CÉAC résout aussi le deuxième problème. En effet, si la CÉAC intègre les faits susmentionnés, alors elle permet et même commande que soient menées des recherches sérieuses sur (a) les mécanismes environnementaux (physiques, sociaux et culturels) favorisant le développement des troubles mentaux et sur (b) les composantes corporelles, multipersonnelles et cognitivo-institutionnelles de ces derniers. Le troisième problème est résolu parce que la CÉAC favorise les recherches phénoménologiques et qualitatives. Il en est ainsi parce qu'elle contient la thèse énactive autopoïétique de la

continuité vie-esprit qui tient pour centraux notre conscience phénoménale et le monde symbolique et affectif des humains. Quant au quatrième problème, j'en propose une solution à la section 3.2.2 après avoir au préalable, à la section 3.2.1, présenté en détail la CÉOST. Je crois que la solution que je présente est plausible parce que la CÉOST contient une structure conceptuelle qui, contrairement à celle de la TBS (et, conséquemment, à celle de la conception du trouble mental des RDoC), est capable de saisir la normativité dans les systèmes vivants et qui, en conséquence, est en mesure de distinguer un système en santé d'un système qui présente un trouble (mental ou physique).

En somme, l'argumentaire que je déploie dans cette thèse vise à montrer que si l'on veut, à l'instar du NIMH, enrichir nos connaissances à propos des troubles mentaux et contribuer à l'élaboration de traitements novateurs et efficaces, ce n'est pas la conception du trouble mental des RDoC qu'il faut adopter en recherche psychiatrique, mais la CÉOST. En effet, la conception du trouble mental des RDoC nuit à cette entreprise épistémique et curative que poursuit le NIMH, non seulement parce qu'elle ne saisit pas la normativité dans le monde du vivant, mais aussi parce qu'elle réduit les troubles mentaux à des mesures quantitatives de dysfonctions psychobiologiques. De fait, les troubles mentaux sont plus que cela : ils comportent des aspects qualitatifs et phénoménaux et sont potentiellement constitués de composantes extraneurales. Adopter la CÉOST qui saisit ces propriétés et constituants de même que la normativité des systèmes vivants favoriserait donc la réussite de ce projet du NIMH qui vise ultimement à améliorer la qualité de vie de millions de personnes⁴.

⁴ Avant de poursuivre, je tiens à mentionner qu'ont vu le jour les travaux en philosophie de la psychiatrie de Sanneke de Haan (2020b, 2020a) et de Kristopher Nielsen et Tony Ward (2018, 2020a) alors que j'étais avancé dans la rédaction de ma thèse. C'est aussi à ce moment que j'ai pris plus sérieusement connaissance des recherches de Thomas Fuchs (2018, voir particulièrement le chapitre 7). Dans l'annexe D, je présente brièvement leurs travaux et montre en quoi les produits de nos recherches se ressemblent, mais sont aussi différents.

CHAPITRE I

RESEARCH DOMAIN CRITERIA (RDOC) ET INTERPRÉTATION PHILOSOPHIQUE DE SA CONCEPTION DU TROUBLE MENTAL

Dans ce chapitre initial, j'ai deux objectifs. Le premier est de présenter d'où proviennent les RDoC (section 1.1), pourquoi le NIMH les a développés (section 1.2), ce qu'ils sont (section 1.3) et le type de recherches qu'ils motivent (section 1.4). Le deuxième objectif est de proposer une interprétation philosophique de la conception du trouble mental des RDoC (section 1.5). Une telle interprétation peut être comprise comme une explicitation des bases philosophiques sur lesquelles repose cette conception des désordres mentaux (sections 1.5.2-1.5.5).

1.1 Le National Institute of Mental Health (NIMH), son plan stratégique et les RDoC

Aux États-Unis, le NIMH est l'agence fédérale principale responsable de la recherche fondamentale et clinique sur les désordres mentaux. Ce dernier est un des 27 Instituts et Centres dont est composé le *National Institutes of Health* (NIH), l'agence nationale responsable de la recherche médicale dans ce pays. Le NIH fait partie du U.S. *Department of Health and Human Services* (HHS) (voir NIMH, 2017, section « about us »).

La mission première du NIMH est de transformer « la compréhension et le traitement des maladies mentales par l'intermédiaire de la recherche fondamentale et clinique

[pour ouvrir] la voie à la prévention, au rétablissement et à la guérison » (NIMH, 2020a, p. 3).

Dans ses communications, le NIMH insiste sur le fait que sa mission est on ne peut plus urgente en raison de l'ampleur et de l'impact des problèmes de santé mentale aux États-Unis. En effet, on peut lire sur son site Internet que

[s]elon des estimations récentes, les maladies mentales représentent 21,3 pour cent de toutes les années vécues avec un handicap aux États-Unis. On estime [également] que 9,6 millions d'adultes américains souffrent d'une maladie mentale grave (MMG) [, laquelle est à l'origine, chez eux, d'une réduction considérable de leur capacité de fonctionner dans leur vie quotidienne]. En outre, plus de 41 149 Américains meurent chaque année par suicide, ce qui est plus du double de la mortalité annuelle due aux homicides ou au sida. Au-delà de la morbidité et de la mortalité, une estimation prudente situe les coûts financiers directs et indirects associés aux maladies mentales aux États-Unis à plus de 300 milliards \$ par an. La maladie mentale se classe au troisième rang des maladies médicales les plus coûteuses, en ce qui a trait aux dépenses globales en santé, après les problèmes cardiaques et les blessures traumatiques. (NIMH, 2019a)

Pour mener à bien sa mission qui, en dernière instance, est d'aider les personnes atteintes de maladies mentales et de promouvoir à la fois la prévention et la guérison de celles-ci (NIMH, 2020a, p. 3), le NIMH adopte les quatre objectifs stratégiques (OS) suivants qui sont énoncés dans son *Plan stratégique pour la recherche* (NIMH, 2020a, p. 4):

- (1) Définir les mécanismes cérébraux sous-jacents aux comportements complexes (OS1).
- (2) Examiner les trajectoires de la maladie mentale tout au long de la vie (OS2).
- (3) S'efforcer de prévenir et de guérir (OS3).
- (4) Renforcer l'impact de la recherche soutenue par le NIMH sur la santé publique (OS4).

En 2008, le NIMH a développé un plan stratégique dont les OS sont très semblables à ceux mentionnés précédemment (voir NIMH, 2008, p. v)⁵. Le projet des RDoC est né lors de l'élaboration de ce plan en 2008. En effet, les RDoC furent conçus spécifiquement pour mener à bien la stratégie 1.4 de l'OS1 que l'on retrouve dans ce plan stratégique élaboré il y a 14 ans (NIMH, 2015a, p. 12). L'OS1, en 2008, stipule que le NIMH a pour objectif de « promouvoir la découverte, dans les sciences du cerveau et du comportement, afin d'alimenter la recherche portant sur les causes des troubles de santé mentale » (NIMH, 2008, p. v). Le point 1.4, quant à lui, dit que le NIMH veut « [d]évelopper, à des fins de recherche, de nouvelles manières de classifier les troubles mentaux basées sur les *dimensions* de comportements observables et de mesures neurobiologiques » (NIMH, 2008, p. 9, mes *italiques*)⁶⁷.

En somme, l'idée générale qui émerge de la stratégie 1.4 de l'OS1 est qu'en 2008 les RDoC sont conçus pour développer une nosologie des troubles mentaux utile pour la *recherche* et explicitement basée sur une étiologie formulée selon des construits

_

Que penser de ces changements ? Se pourrait-il que le NIMH assume encore plus qu'avant son approche neurocentrée des troubles de santé mentale ?

⁵ Le plan stratégique de 2015 contient lui aussi quatre objectifs stratégiques qui sont très semblables à ceux de 2008. Il est toutefois intéressant de noter que l'OS1 du plan de 2008 est plus semblable à celui de 2020 qu'à celui de 2015. En effet, selon l'OS1 de 2008 et l'OS1 de 2020, il est clair que la recherche doit porter sur le cerveau. Cependant, dans l'OS1 de 2015, ce n'est pas dit explicitement. Voici l'OS1 des plans stratégiques de 2008, 2015 et 2020 :

OS1 (2008): Promouvoir la découverte, dans les sciences du *cerveau* et du comportement, afin d'alimenter la recherche portant sur les causes des troubles de santé mentale (NIMH, 2008, p. v, mes *italiques*)

OS1 (2015): Définir les mécanismes des comportements complexes (NIMH, 2015a, p. 7)

OS1 (2020): Définir les mécanismes cérébraux sous-jacents aux comportements complexes (NIMH, 2020a, p. 4, mes italiques)

⁶ La traduction française de ce point 1.4 est de Demazeux et Pidoux (2015, p. 792-793).

⁷ Ces comportements observables dont il est ici question peuvent être autant de nature cognitive (par exemple, la capacité mnémonique d'un individu suivant un test psychologique quelconque) que de nature émotive (par exemple, la réaction d'un individu à des visages présentant une gamme d'émotions).

dimensionnels (donc mesurables) appartenant aux neurosciences moléculaires, cognitives et comportementales.

1.2 Qu'est-ce qui motive le NIMH à mettre sur pied les RDoC?

Au moins trois raisons me semblent motiver le lancement des RDoC. La première est que le NIMH souhaite libérer la recherche du carcan que sont les catégories du *Manuel diagnostique et statistique des troubles mentaux* (DSM)⁸ qu'il juge (comme plusieurs) insuffisamment valides (Cuthbert, 2015, p. 91-92; à ce sujet, voir aussi Cuthbert et Insel, 2013; Cuthbert et Kozak, 2013; Hyman, 2007, 2010; Sanislow *et al.*, 2019, p. 779; Simmons et Quinn, 2014). On dit que les catégories du DSM manquent de validité principalement pour trois raisons. La première est que plusieurs d'entre elles regroupent des individus présentant des symptômes très différents (c'est le problème

⁸ Le DSM est publié par l'American Psychiatric Association (APA). Cette classification des troubles mentaux a des fonctions en clinique, pour les compagnies d'assurance, en éducation médicale, à la Cour et, finalement, en recherche. En effet, cet ouvrage est utilisé par les psychiatres et les médecins dans leur pratique clinique. En ce moment, c'est selon ces catégories que les cliniciennes et les cliniciens font leurs diagnostics. Cette classification sert aussi aux compagnies d'assurance. Par exemple, celles-ci vont accorder à un individu une assurance emploi dans la mesure où ce dernier présente un diagnostic du DSM invalidant pour le travail et dans la mesure où ce dernier suit une thérapie appropriée pour sa condition (pour un trouble de l'adaptation au travail qui devient une dépression majeure, par exemple, l'individu s'engagera à suivre une thérapie et à prendre des antidépresseurs). Cette classification sert, en outre, à l'éducation des futurs intervenants en santé mentale. Elle influe ainsi sur la manière dont les futurs cliniciennes et cliniciens concoivent les différents troubles mentaux. Elle a également des fonctions légales. En effet, un individu que l'on évalue, selon certaines catégories du DSM, comme étant incapable de discerner le bien du mal lors de son crime peut être jugé non responsable de ses actes lors d'un procès. Enfin, depuis la parution du révolutionnaire DSM-III (1980), cette nosologie sert de cadre pour la recherche en psychiatrie. Le DSM-III fut qualifié de révolutionnaire entre autres parce qu'il ne contenait plus d'étiologie psychanalytique comme le DSM-I (1952) et le DSM-II (1968), mais était athéorique et contenait une classification purement empirique fondée sur des signes et symptômes issus d'observations cliniques. Le DSM-5 (2013), à l'instar du DSM-IV (1994) et du DSM IV-TR (2000), est conçu dans le même esprit que le DSM-III. Cependant, il faut noter que, dans le DSM-5, quelques catégories, comme celle référant au trouble neurocognitif dû à la maladie de Huntington, ne sont pas essentiellement de nature phénoménologique et font référence à des processus physiopathologiques (American Psychiatric Association, 2013, p. 638-641).

de l'hétérogénéité). Par exemple, 636 120 combinaisons de critères diagnostiques peuvent être associées à la catégorie du syndrome de stress post-traumatique (SSPT) (Galatzer-Levy et Bryant, 2013). Dans le même ordre d'idées, Hyman (2010, p. 160) fait remarquer que deux individus présentant un diagnostic de dépression majeure peuvent en principe ne partager qu'un seul symptôme des neuf faisant partie de cette catégorie.

La deuxième raison est qu'un grand nombre des catégories du DSM distinguent des ensembles d'individus qui devraient fort probablement être regroupés (c'est le problème de la comorbidité). En effet, fréquemment, au moins deux troubles mentaux (comme ils sont conçus dans le DSM) coexistent chez un même patient (la dépression majeure et le trouble anxieux, par exemple) (Hyman, 2010, p. 167-169). Si on fait l'hypothèse que ces patients ne présentent pas deux troubles mentaux, mais un seul, il serait alors préférable, lors des recherches empiriques, de les regrouper sous la même catégorie.

La troisième raison qui fait croire que les catégories du DSM manquent de validité est qu'elles sont trop spécifiques (c'est le problème de la surspécification) (Cuthbert, 2014a, p. 3 : 45-3:54; Hyman, 2010, p. 166-167). Le fait qu'elles présentent cette propriété est problématique en recherche parce que certaines personnes qui devraient être incluses dans des études empiriques ne le sont pas et que l'on perd, pour cette raison, des données qui auraient pu être utiles pour mieux comprendre les troubles mentaux.

Ces trois problèmes mettant à jour le manque de validité des catégories du DSM font qu'il est très difficile (voire impossible) de trouver, lors de recherches empiriques, des profils génétique, neural et psychologique semblables chez les individus qui composent ces catégories.

Le NIMH pense régler ce problème de la validité des catégories des troubles mentaux du DSM en élaborant, dans le cadre des RDoC, de *nouvelles catégories* pour la recherche (dimensionnelles cette fois-ci et non catégorielles comme celles du DSM). Ces catégories constituées des construits des RDoC saisiraient des individus présentant des profils semblables sur les plans génétique, neural et comportemental. Elles mettraient également à jour des correspondances entre des mesures d'activité de systèmes biopsychologiques dysfonctionnels et des symptômes psychiatriques (voir section 1.3, principe 3). Ainsi, les nouvelles catégories diagnostiques issues des RDoC se distingueraient de celles du DSM qui ne sont pas fondées sur une étiologie formulée selon des construits dimensionnels appartenant aux neurosciences moléculaires, cognitives et comportementales.

L'élaboration de ces nouvelles catégories des troubles mentaux au sein des RDoC se fait suivant *une approche transdiagnostique* en recherche psychiatrique. Selon cette approche, les construits des RDoC qui servent à développer les nouvelles catégories sont étudiés et, éventuellement, validés suivant des échantillons dont les membres sont regroupés selon plusieurs catégories diagnostiques du DSM (et même selon d'autres critères comme « toutes personnes se présentant à l'urgence psychiatrique »). Il est essentiel de faire cette démarche de validation des construits des RDoC si l'on veut, dans l'avenir, avoir des catégories diagnostiques valides.

La seconde raison qui motive la mise sur pied des RDoC est que le NIMH souhaite développer une psychiatrie clinique de précision (Insel, 2014, p. 396; voir aussi Insel et Cuthbert, 2015; Sanislow *et al.*, 2019, p. 781). Le propre de la psychiatrie de précision est d'offrir des traitements en fonction du profil psychobiologique de la personne qui est diagnostiquée. Le rôle que les RDoC jouent dans l'élaboration de ce type de pratiques cliniques est de contribuer à la découverte de marqueurs psychobiologiques quantitativement mesurables. Trouver ces marqueurs est pertinent

pour au moins trois raisons. D'abord, ceux-ci devraient aider à poser un diagnostic plus précis (ex.: à l'aide de techniques d'imagerie cérébrale ou avec des examens génétiques) ⁹. Ensuite, ceux-ci contribueraient à relancer la recherche sur les médicaments psychiatriques, laquelle stagne en ce moment (Insel et Sahakian, 2012). En effet, en raison du manque actuel de nouvelles cibles psychobiologiques, les compagnies pharmaceutiques ont réduit massivement leur investissement dans la recherche sur les médicaments ¹⁰. Enfin, ces marqueurs psychobiologiques sont intéressants parce qu'ils pourront être la cible d'efficaces nouveaux traitements pharmacologiques, mais aussi de traitements physiques invasifs (comme la stimulation cérébrale profonde), de traitements physiques non invasifs (comme la stimulation magnétique transcrânienne) et de traitements psychologiques (comme une thérapie cognitive comportementale) qui seront adaptés aux spécificités (génétiques, neurales et psychologiques) des patients.

Je suis venu ici [au NIMH] avec le préjugé que nous pourrions régler les problèmes diagnostiques avec la génomique et l'imagerie. J'avais tort. Nous avons dépensé beaucoup d'argent sur ces deux sujets et, en fin de compte, il n'y avait pas vraiment de preuves, peut-être un peu de preuves, mais ce n'était pas la technologie transformatrice que nous aurions espérée, comme cela avait été le cas pour les maladies cardiaques, certaines maladies neurologiques et bien d'autres aspects de la médecine. (Insel, 2018)

À l'heure actuelle, si Insel pense toujours que le cerveau est la seule base physique d'où émergent les troubles de santé mentale, il croit que l'intelligence artificielle et la science des données, jointes à ce que disent les neurosciences et la clinique sur les troubles mentaux, sont les disciplines qui peuvent faire progresser la psychiatrie. Ayant cela en tête, en 2017, il a cofondé l'entreprise *Mindstrong*.

⁹ À ce sujet, il est important de noter que l'ancien directeur du NIMH, Thomas R. Insel, adopte maintenant une attitude critique relativement aux convictions qu'il avait, il y a quelques années, sur ce qu'était de la bonne recherche en santé mentale et sur ce qu'étaient les outils idéaux pour faire des diagnostics psychiatriques. De fait, il dit:

¹⁰ Étant donné que « l'efficacité des médicaments dont nous disposons pour traiter les troubles psychiatriques est limitée » (Krystal, 2016, 3:12-3:32), les sympathisants des RDoC, comme l'ancien directeur du NIMH Thomas R. Insel (2011), pensent qu'il est d'autant plus nécessaire de redémarrer cette entreprise.

À titre d'exemples, les recherches menées par les équipes d'Etkin et de Sukhodolsky, qui sont présentées aux sections 1.4.1 et 1.4.2, prennent part au développement de cette psychiatrie de précision. En effet, comme on le verra, l'équipe d'Etkin, en élaborant, à l'aide de la stimulation magnétique transcrânienne non invasive (spTMS) et de l'imagerie par résonance magnétique fonctionnelle (IRMf), des cartes causales des circuits neuraux actifs chez les gens anxieux et dépressifs, contribue au développement de diagnostics précis et de traitements physiques non invasifs très ciblés (voir section 1.4.1). Quant à l'équipe de Sukhodolsky, elle participe au développement de connaissances neurocognitives qui permettront de mieux choisir le type de thérapies qui conviendrait à un enfant présentant des comportements agressifs (voir section 1.4.2).

La troisième raison qui pousse le NIMH à développer les RDoC est qu'il y a, en ce moment, une crise de la santé mentale aux États-Unis (voir section 1.1)¹¹. Le NIMH croit qu'il sera en mesure d'intervenir efficacement sur cette situation alarmante grâce à une meilleure compréhension des troubles mentaux et en raison du développement d'outils diagnostiques et de traitements performants que permettront les RDoC. C'est à l'aide de ce cadre de recherche, et de la psychiatrie de précision que ce dernier permet de développer, que le NIMH croit pouvoir concrétiser sa vision de l'avenir, à savoir celle de construire un monde où l'on serait capable de prévenir et de guérir les troubles mentaux (NIMH, 2015a, p. 18).

¹¹ À noter également qu'il y a une crise *mondiale* de la santé mentale. En effet, on peut lire sur le site Internet du *Centre de toxicomanie et de santé mentale* (CAMH) que, « [s]elon l'Organisation mondiale de la santé, environ 450 millions de personnes sont actuellement atteintes d'une maladie mentale, ce qui en fait la première cause d'invalidité dans le monde » (CAMH, 2019).

1.3 Que sont les RDoC?

À la lumière d'une analyse que je fais de la littérature écrite à leur sujet par leurs représentants et leurs adhérents ¹², les RDoC peuvent être compris comme cinq principes ¹³ qui forment un système où le premier d'entre eux est éclairé par tous les autres. De ces cinq principes, on peut tirer une conception de ce que sont les troubles mentaux. On peut également inférer de ces principes la manière d'expliquer ces désordres et la façon de les étudier. Voici ces cinq principes :

- (1) Les troubles mentaux sont des dysfonctions des circuits du cerveau responsables de la cognition, des émotions et des comportements.
- (2) Cette conception du trouble mental est organisée autour de construits.

Ces construits, constitués eux-mêmes de sous-construits, décrivent les grands domaines de fonctionnement du cerveau impliqués dans les troubles mentaux.

¹² On trouve parmi eux Joshua A. Gordon (l'actuel directeur du NIMH), Bruce Cuthbert (directeur de l'Unité des RDoC), Sarah Morris (directrice adjointe de l'Unité des RDoC) et Uma Vaidyanathan (Gestionnaire de programme scientifique à l'Unité des RDoC). Charles Sanislow est également un défenseur des RDoC. Ces gens étaient, en effet, impliqués dans l'élaboration du cadre et l'ont défendu, au nom du NIMH, dans de nombreuses publications. On peut aussi dire que Thomas R. Insel est un défenseur des RDoC. Celui-ci, en effet, en plus de les avoir lui aussi défendus dans plusieurs écrits, était directeur du NIMH lors de leur mise sur pied (pour avoir accès à une liste complète des gens impliqués directement dans le développement des RDoC, j'invite la lectrice ou le lecteur à consulter la section « RDoC Unit and Workgroup Members » de la section « Research Domain Criteria [RDoC] » sur le site Internet de la NIMH [2019b]). Par ailleurs, je considère que les personnes qui ont coécrit avec des membres de l'Unité des RDoC des articles pour promouvoir les RDoC sont aussi des tenants ou défenseurs de ce programme de recherche. Michael J. Kozak, entre autres, fait partie des gens que je classe dans cette catégorie (voir, par exemple, Cuthbert et Kozak, 2013; Kozak et Cuthbert, 2016). Enfin, toutes les personnes faisant de la recherche dans le cadre des RDoC peuvent être considérées, du moins en principe, comme des adhérentes de ce cadre conceptuel. Par exemple, Lisa M. McTeague, Madeleine S. Goodkind et Amit Etkin (2016) font partie de cette catégorie.

¹³ Je prends cette idée que les RDoC sont un ensemble de principes de Vaidyanathan qui en parle dans un Webinar (2016, p. 1:15-5:35).

Ces domaines sont les systèmes de valence négative (ex. : peur, anxiété), les systèmes de valence positive (ex. : apprentissage de comportements menant à une récompense), les systèmes cognitifs (ex. : mémoire de travail, attention), les systèmes responsables des processus sociaux (ex. : identification des expressions faciales, représentation de soi), les systèmes d'activation et de modulation (ex. : activité cérébrale en état de repos) et les systèmes sensorimoteurs (ex. : agentivité et titularité). Ces domaines (ou construits) se trouvent dans la matrice des RDoC qui est présentée à la section 1.5.1.

Ces construits doivent respecter trois critères (Sanislow, 2016, p. 1160): (a) des données soutiennent leur validité en tant qu'unité fonctionnelle de comportements [(on pourrait aussi simplement dire qu'ils sont des construits utilisés en sciences cognitives)] (b) il est possible de les associer à des symptômes [psychiatriques] cliniques et (c) des données montrent qu'il est possible de les associer à l'activité de circuits ou de systèmes neuraux. Considérant ce troisième critère, comme l'explique Sanislow selon MacCorquodale et Meelh, ces construits devraient être conçus comme des construits hypothétiques [hypothetical construct] parce qu'ils « sont liés à une "entité, un processus ou un événement" (MacCorquodale et Meehl, 1948, p. 95-96) qui correspond ici à un mécanisme interne » (Sanislow, 2016, p. 1160).

Ces domaines et les construits qui les composent ne sont pas conçus comme indépendants les uns des autres. Au contraire, on pense qu'ils sont reliés entre eux et, conséquemment, qu'ils s'affectent mutuellement. Par exemple, les sous-construits des systèmes de valence négative ou positive sont reliés au construit « contrôle cognitif ». On peut d'ailleurs développer, à partir de ces derniers, le construit de haut-niveau « dérégulation des émotions » qui ne se trouve pas dans la matrice (Sanislow, 2016, p. 1160).

Ces construits sont des entités théoriques *quantitativement mesurables* et intègrent autant des éléments psychologiques que biologiques. Ils sont mesurés, dans le cadre de paradigmes expérimentaux divers, selon les sept types d'unité d'analyse suivant : « gènes », « molécules », « cellules », « circuits neuraux », « physiologie » (ex. : niveau de cortisol), « comportement » (ex. : mesure du niveau de performance de la mémoire de travail) et « auto-évaluation » [self-report] (ex. : mesure des comportements liés à la peur de la peur ou mesure des sensations associées à l'expérience de l'anxiété avec l'*Anxiety Sensitivity Index*).

(3) Il faut analyser ces construits selon *au moins* deux unités d'analyse où l'une d'entre elles *doit* être de nature biologique (l'expression « nature biologique » réfère ici aux domaines génétique, moléculaire, cellulaire, physiologique et des circuits neuraux).

Ce troisième principe est, en effet, sous-jacent à cette affirmation de Cuthbert et Kozak (2016, p. 292) : « [l]es construits des RDoC doivent nécessairement comprendre des processus biologiques. » ¹⁴ Ce troisième principe est aussi cohérent avec l'OS1 du plan stratégique du NIMH selon lequel il faut « [d]éfinir les mécanismes *cérébraux* sous-jacents aux comportements complexes » (NIMH, 2020a, p. 4, mes *italiques*) (voir section 1.1).

¹⁴ Il semble exister des exceptions à cette règle. Par exemple, dans l'étude de Bedwell, Gooding, Chang et Trachik (2014), aucune donnée biologique n'est mentionnée. En effet, ce sont essentiellement des mesures d'auto-évaluation et des mesures comportementales qui sont mises en relation, à savoir le niveau d'anhédonie sociale et le niveau de performance de la mémoire déclarative. Il existe peut-être des exceptions à cette règle que formulent Cuthbert et Kozak parce que les représentants des RDoC définissent maintenant les RDoC comme un ensemble de *desiderata* plutôt que comme un ensemble de *normes* à respecter (Stoyanov *et al.*, 2019, p. 60).

Il faut préciser que les construits biologiques privilégiés par le NIMH sont ceux qui réfèrent aux *circuits* neuraux. En effet, le NIMH a décidé, pour l'instant, de faire des circuits neuraux *l'unité centrale* autour de laquelle s'articuleront les modèles explicatifs des troubles mentaux (voir, par exemple, Insel, 2010; Sanislow, 2016, p. 1160; Sanislow *et al.*, 2010, p. 633). On peut penser qu'il a fait ce choix, entre autres, pour des raisons de faisabilité. En effet, étant donné l'état des connaissances et des techniques d'investigation actuelles, il est envisageable de faire correspondre l'activité de certains circuits neuraux à certains comportements et d'associer ce mécanisme neurocomportemental à un syndrome psychiatrique. Cependant, il est beaucoup moins concevable que l'on puisse faire correspondre des mécanismes biomoléculaires à ces mêmes comportements et, en conséquence, que l'on puisse associer ce mécanisme « moléculocomportemental » à un syndrome psychiatrique.

En faisant ce choix pragmatique, le NIMH pense qu'il a plus de chance d'atteindre un de ses objectifs qui est d'établir une correspondance entre les données des neurosciences et les aspects des comportements ou de la cognition des gens qui reçoivent des diagnostics cliniques (Sanislow, 2016, p. 1160). S'il réussit à atteindre ce but, alors le NIMH produira des catégories diagnostiques basées sur des mesures de fonctionnement des systèmes psychobiologiques. Ces catégories de troubles mentaux seront, en cela, différentes de celles du DSM dont ce fondement est manquant.

(4) Le fonctionnement *normal* et le fonctionnement *anormal* (ou le dysfonctionnement) de l'esprit/cerveau sont décrits avec des mesures *quantitatives* des construits des RDoC. Dans cette perspective, le normal et le pathologique se trouvent sur un même continuum et sont associés à des *niveaux* de fonctionnement des systèmes psychobiologiques.

(5) Des *facteurs développementaux* et *environnementaux* ont un impact sur le niveau de fonctionnement de l'esprit/cerveau.

Les architectes des RDoC soutiennent qu'il faut absolument comprendre la manière dont le cerveau se développe dans le temps parce que de plus en plus d'études montrent que les troubles de santé mentale sont des troubles développementaux (NIMH, 2015a, p. 31). On pense, en effet, que plusieurs troubles mentaux sont latents durant l'enfance et se développent jusqu'à l'adolescence ou jusqu'au début de l'âge adulte, moment où les symptômes apparaissent plus clairement. Un des buts du NIMH serait que l'on soit éventuellement capable d'intervenir de manière préventive et précise sur le cerveau avant que les symptômes n'apparaissent.

L'aspect « interactions avec l'environnement » a aussi été retenu par les concepteurs des RDoC puisque c'est un fait incontestable que les troubles de santé mentale sont souvent causés, à certains moments clés de l'histoire de vie de l'individu et selon une importance variable, par des facteurs environnementaux physiques (ex. : substance toxique) et/ou sociaux (ex. : la pauvreté). En effet, les facteurs environnementaux délétères pour la santé causent des changements dans l'expression génétique et cela entraîne des changements dans les circuits neuraux qui deviennent éventuellement dysfonctionnels.

1.4 Quels types de recherches sont menées au sein des RDoC?

Les RDoC favorisent grandement la poursuite de recherches portant sur la neurobiologie des fonctions et dysfonctions (émotives, cognitives et comportementales) de l'esprit/cerveau. Ces recherches, soutient-on au NIHM, favoriseront éventuellement

la découverte de traitements efficaces pour les troubles mentaux. Cette orientation est en harmonie avec les OS1 et OS3 du *Plan stratégique pour la recherche* du NIMH selon lesquels il faut «[d]éfinir les mécanismes *cérébraux* sous-jacents aux comportements complexes » (OS1) et « [s]'efforcer de prévenir et de guérir » (OS3) les troubles mentaux (NIMH, 2020a, p. 4, mes *italiques*) (voir section 1.1). Cette orientation est également compatible avec une des conclusions principales que l'on peut tirer de l'étude de Carcone et Ruocco (2017) qui font une revue systématique du type de recherches qui ont été menées dans le cadre des RDoC. En effet, leur travail montre qu'un très grand nombre des études menées au sein des RDoC visent à « décrire les relations entre les comportements ou les caractéristiques de haut niveau et leurs fondements biologiques de bas niveau » (Carcone et Ruocco, 2017, p. 3)¹⁵. J'ajouterais à ces propos de Carcone et Ruocco que ces études des bases biologiques des

¹⁵ Carcone et Ruocco (2017, p. 2) expliquent, dans leur article, la manière dont ils ont procédé pour déterminer le corpus d'études qui représentent bien le type de recherches menées dans le cadre des RDoC. D'abord, ils ont utilisé le moteur de recherche ProQuest pour avoir accès aux études menées dans le cadre des RDoC indexées dans les bases de données PsycINFO et Medline. Les expressions qu'ils ont utilisées, lors de leur enquête, ont été « Research Domain Criteria » et/ou l'acronyme « RDoC ». Les résultats générés ont été filtrés et se sont limités aux travaux évalués par les pairs publiés dans une revue scientifique après le 4 avril 2013. Leur requête a produit 330 résultats. Les résumés de ces 330 résultats ont été ensuite filtrés afin que soient exclus du corpus les recensions, les commentaires et les résultats que la recherche n'aurait pas dû produire (comme une étude portant sur le carbone organique dissous réfractaire, lequel est un paramètre utilisé dans l'évaluation de la chimie de l'eau). Ce premier filtrage a laissé passer 61 articles sur les 330 (269 d'entre eux ont donc été exclus). Ces 61 articles ont été lus et analysés par Carcone et Ruocco qui leur ont fait subir un second filtrage. Lors de celui-ci, pour être retenus, les 61 articles devaient satisfaire à des critères d'inclusion et ne pas satisfaire à des critères d'exclusion. Les critères d'inclusion étaient les suivants : (1) l'article examine empiriquement un ou plus d'un domaine, construit ou élément des RDoC et (2) l'article fait directement référence au cadre des RDoC pour justifier l'étude, pour interpréter les résultats ou quand il est question des implications et des orientations futures. Par ailleurs, les critères d'exclusion qui ne devaient pas être satisfaits étaient les suivants : (1) l'article est une proposition de recherche dans le cadre des RDoC qui n'a pas encore été menée (2) l'article porte sur la validation empirique d'une mesure psychométrique liée à un construit des RDoC plutôt que sur l'examen du concept lui-même et (3) l'article est centré sur une étude de cas. Ce sont 48 articles des 61 articles qui ont été retenus par Carcone et Ruocco lors de ce second filtrage. Ce sont ces 48 articles que ces derniers ont analysés pour faire leur revue systématique.

comportements ont aussi pour but de contribuer au développement de traitements psychiatriques.

Je présente aux sections 1.4.1 et 1.4.2 deux projets de recherche qui exemplifient bien le type d'études scientifiques qui se font dans le cadre des RDoC.

1.4.1 Cartographie et manipulation des circuits neuraux responsables des dysfonctions émotives et cognitives dans l'anxiété et la dépression

Amit Etkin, professeur associé au *Department of Psychiatry and Behavioral Sciences* de l'université Stanford et membre du *Stanford Neurosciences Institute*, a reçu une subvention du NIMH pour son projet de recherche intitulé «Mapping and Manipulating Circuits for Emotion and Cognition in Anxiety and Depression»¹⁶. En harmonie avec les objectifs des RDoC, cette recherche a pour objectif d'élaborer, suivant diverses interventions et mesures (voir plus bas), une carte causale des circuits cérébraux impliqués dans les dysfonctionnements de la régulation émotionnelle et des fonctions exécutives présentes dans les troubles de l'anxiété et la dépression (Etkin, 2014b). Selon Etkin, ce projet de recherche est novateur parce qu'il n'y a pas, à l'heure actuelle, de telles cartes causales. En effet, dit-il,

les études classiques de neuroimagerie fournissent des informations sur les corrélations entre le cerveau et le comportement, mais pas sur la manière dont l'activité dans une région du cerveau entraîne directement l'activation ou l'inhibition d'une autre [région]. (Etkin, 2014b)

Etkin ajoute que

[p]our comprendre les relations causales [qui existent] entre les différentes régions [cérébrales] ou réseaux cérébraux, il est nécessaire d'exercer un

_

¹⁶ Le numéro du projet de recherche est « 5R01MH103324-03 ».

contrôle expérimental sur des régions spécifiques du cerveau et, simultanément, [d'observer] les conséquences [de cette intervention expérimentale] sur d'autres régions [cérébrales] ou réseaux cérébraux. (Etkin, 2014b)

Afin d'exercer un contrôle sur des régions spécifiques du cerveau, l'équipe d'Etkin utilise conjointement la stimulation magnétique transcrânienne non invasive (spTMS) et l'imagerie par résonance magnétique fonctionnelle (IRMf). Plus précisément, cette équipe de recherche (a) intervient avec la spTMS sur différents circuits neuraux sousjacents à des émotions et à des processus cognitifs et (b) observe avec l'IRMf quels effets cause cette intervention.

Construire cette carte causale des circuits neuraux permettra à l'équipe d'Etkin de voir « lesquelles des voies cérébrales sont intactes et lesquelles sont anormales » (Etkin, 2014b) chez les gens présentant des troubles anxieux et/ou dépressifs. L'élaboration de cette carte causale des circuits neuraux sera éventuellement pertinente pour les cliniciennes et cliniciens puisqu'elle les aidera à déterminer les circuits neuraux sur lesquels intervenir et la manière de le faire. Par exemple, pour réduire l'activité d'un dysfonctionnement neural, on pourra déterminer « si la stimulation doit être faite sur des voies intactes ou sur des voies anormales » (Etkin, 2014b). De plus, si la voie neurale sur laquelle on doit intervenir est hors de portée d'une intervention directe, cette carte causale des circuits neuraux permettra de déterminer le circuit sur lequel on doit intervenir pour indirectement atteindre le circuit que l'on cible et sur lequel on veut ultimement agir.

Conformément au désir du NIMH que les chercheuses et chercheurs adoptent une approche transdiagnostique dans leur recherche (voir section 1.2), l'échantillon utilisé dans cette étude empirique est composé d'individus qui ont des symptômes de dépression, d'anxiété, de stress post-traumatique ou qui n'expérimentent aucun de ces symptômes (Etkin, 2014a).

1.4.2 Étude des circuits neuraux à l'aide de la thérapie cognitivocomportementale chez les enfants agressifs

Denis Sukhodolsky, professeur associé au Yale Child Study Center, a reçu une subvention du NIMH pour son projet de recherche intitulé « Using CBT [(Cognitive behavioral therapy)] to examine circuitry of frustrative non-reward in aggressive children »¹⁷. L'équipe de Sukhodolsky veut comprendre la relation qui existe, chez les enfants, entre l'activité des circuits neuraux et les comportements que contient le construit « non-récompense frustrante » [frustrative Nonreward] (ce construit se trouve dans la matrice des RDoC à la section 1.5.1). En mettant en lumière cette relation, l'équipe de Sukhodolsky veut contribuer à la validation de ce construit (pour plus de détails sur cette notion de validation, voir section 1.5.4). Ce construit se trouve dans le domaine de fonctionnement « systèmes de valence négative » que l'on retrouve dans la matrice des RDoC (voir section 1.5.1). Comme on peut le lire dans le résumé du projet de recherche, « une non-récompense frustrante est définie comme une réaction déclenchée en réponse au retrait [d'une récompense] ou [...] [en réponse] [au fait qu'on empêche l'accès] à une récompense » (Sukhodolsky, 2014). Cette réponse est une agression réactive, laquelle doit être distinguée des comportements d'agression instrumentale et d'agression manifeste (Sukhodolsky et al., 2016, p. 39). Les mécanismes neuraux associés à l'agression réactive sont une activité exacerbée de l'amygdale et, en même temps, une hypoactivité des régions préfrontales, lesquelles ont pour fonction de réguler les émotions (Sukhodolsky, 2014).

Pour comprendre la relation qui existe, chez les enfants, entre l'activité des circuits neuraux et les comportements que contient le construit « non-récompense frustrante », l'équipe de Sukhodolsky utilise un essai randomisé contrôlé (ERC). Cet ERC est conçu

_

 $^{^{17}}$ Le numéro du projet de recherche est « 5R01MH101514-04 ».

comme une intervention (à mon avis au sens de Woodward, 2003) qui a pour but de moduler l'activité de ces circuits neuraux sous-jacents à ces comportements.

Les 80 enfants qui participent à cet ERC seront assignés de manière aléatoire à suivre soit 12 séances de thérapie cognitivocomportementale (TCC), soit 12 séances de psychothérapie aidante [supportive psychotherapy] (SPT) (Sukhodolsky, 2016, p. 38). La TCC a ceci de spécifique qu'elle enseigne aux enfants, contrairement à la SPT, des techniques pour réguler leurs émotions et résoudre des problèmes (Sukhodolsky et al., 2016, p. 43). Les enfants qui recevront la SPT formeront le groupe contrôle et ceux qui recevront la TCC formeront le groupe expérimental. Dans cette étude empirique, la variable indépendante est le traitement du cerveau/esprit des enfants via la TCC. Quant à elle, la variable dépendante est le niveau de l'activité neural et la quantité de comportements agressifs.

Le premier objectif de cette recherche est de vérifier si une TCC réussie par un enfant entraînera une activité moins grande de son amygdale et une activité plus grande des régions cérébrales impliquées dans le contrôle des émotions (Sukhodolsky, 2014; Sukhodolsky *et al.*, 2016, p. 39). Si ce changement de ces activités cérébrales advient après une TCC et si ce changement est associé à une réduction des comportements d'agression réactive chez l'enfant, cela fournira des données soutenant la validité du construit « non-récompense frustrante » et, par le fait même, une meilleure connaissance des circuits du cerveau responsables de l'agression et de la régulation des émotions.

Le deuxième objectif est de comprendre pourquoi, chez les enfants agressifs présentant des traits d'insensibilité et de pauvreté émotionnelle [callous-unemotional traits], les biomarqueurs reliés à une perception socioaffective dysfonctionnelle et à une plus faible sensibilité à la récompense sont s'associés à une moins bonne réponse à la TCC

(Sukhodolsky, 2014). Autrement dit, l'équipe de Sukhodolsky veut expliquer pourquoi les enfants présentant ce type de profil psychobiologique répondent moins bien à ce genre d'intervention thérapeutique.

Comme on peut le lire sur la page du projet de recherche (Sukhodolsky, 2014), les résultats cliniques seront obtenus à l'aide de différents types de mesures. D'abord, les parents comptabiliseront les comportements agressifs de leur enfant au début (donc avant l'administration de la TCC), au milieu et à la fin de l'étude (donc après les 12 séances de la TCC). En outre, un évaluateur indépendant [blinded rater] attribuera à l'enfant une note d'amélioration que permet de calculer l'Échelle d'impression globale clinique [Clinical Global Impression Scale]¹⁸ (cela suppose que l'enfant sera évalué avec cet outil au début et à la fin de l'étude). Les enfants auront également à effectuer, au début et à la fin de l'expérience, une tâche Go-NoGo (laquelle induit une frustration chez l'enfant) et une tâche de perception émotionnelle de visages. Selon les chercheuses et les chercheurs, il est pertinent de faire passer la tâche Go-NoGo parce que celle-ci active chez l'enfant les circuits neuronaux associés à la régulation des émotions (Sukhodolsky, 2014). Quant à la tâche de la perception émotionnelle de visages, l'équipe de recherche a fait l'hypothèse que les circuits neuraux associés aux systèmes responsables des processus sociaux seraient stimulés lors de celle-ci (NIMH, 2015b, 15: min. 23 sec. -15 min 34 sec.). C'est pendant chacune de ces tâches que l'activité cérébrale des enfants sera enregistrée avec un électroencéphalogramme (EEG) et scannée à l'aide d'une technique d'IRMf. En somme, toutes ces mesures quantitatives permettront de voir de quelle manière la TCC (la variable indépendante) influe sur la quantité de comportements agressifs (la variable dépendante) que produisent les enfants participant à l'étude.

¹⁸ Pour plus de détails sur cette échelle, voir Busner et Targum (2007).

Ces enfants doivent présenter un certain nombre de caractéristiques. D'abord, ils doivent être âgés de 8 à 16 ans. Ils doivent, en outre, présenter des comportements perturbateurs que l'on retrouve dans plusieurs catégories diagnostiques du DSM (explosions de colère, agressivité, défiance ou quérulence)¹⁹. Par ailleurs, les enfants de cette étude doivent être non médicamentés ou doivent prendre depuis au moins six semaines (et continuer de prendre, en principe, jusqu'à la fin de l'étude) une dose stable de médicaments psychiatriques utilisés pour traiter les comportements agressifs, le trouble déficitaire de l'attention avec ou sans hyperactivité (TDAH), l'anxiété ou la dépression (Sukhodolsky *et al.*, 2016, p. 41). Enfin, puisque, comme on l'a vu précédemment, l'activité de certaines de leurs aires cérébrales sera mesurée (avant, pendant et après l'étude), ces enfants doivent être capables de compléter un EEG et de passer un scanneur d'IRMf.

En résumé, cette recherche mettant à profit un ERC devrait contribuer à l'atteinte de deux objectifs. Premièrement, elle devrait aider à mieux comprendre le construit « non-récompense frustrante » de la matrice des RDoC. Cette meilleure compréhension de ce construit sera acquise en le validant par des mesures de l'activité de circuits neuraux, par des mesures physiologiques et par des mesures de comportements. Deuxièmement, elle devrait participer au développement d'une classification basée sur des connaissances neuroscientifiques des troubles mentaux, laquelle aidera à améliorer notre capacité à traiter efficacement les patients (Sukhodolsky, 2014). En effet, sachant le type de profil cérébral qui répond bien à la TCC et le type qui répond moins bien, il

¹⁹ Il est intéressant de noter que cette caractéristique satisfait l'approche transdiagnostique que favorisent les tenants des RDoC (section 1.2) pour l'étude et, éventuellement, pour la validation des construits de la matrice (voir section 1.5.1).

sera possible d'adapter les interventions thérapeutiques aux différents profils des patients.

1.5 Interprétation philosophique de la conception du trouble mental des RDoC

Maintenant que l'on connait les RDoC (section 1.3), les raisons qui ont motivé leur développement (section 1.2) et le type de recherches qu'on y fait (section 1.4), on peut aborder les fondements philosophiques de la conception du trouble mental qui s'y trouve. Pour ce faire, je reviens d'abord sur la conception explicite du trouble mental que contient ce cadre de recherche et la précise en montrant qu'elle est inhérente à ce que l'on appelle la matrice des RDoC (section 1.5.1). Ensuite, dans les autres sections (1.5.2-1.5.5), j'expose les fondements philosophiques qui me semblent sous-jacents aux RDoC.

1.5.1 Conception du trouble mental et matrice des RDoC

Le premier principe des RDoC stipule que les troubles mentaux sont des dysfonctions des circuits du cerveau responsables de la cognition, des émotions et des comportements (voir section 1.3, principe 1). Cette affirmation rend bien compte de la conception du trouble mental des RDoC, mais il faut la préciser. En effet, il est plus juste de dire que, dans ce cadre de recherche, les troubles mentaux sont conçus comme des dysfonctions quantitativement mesurables (section 1.3, principe 4) qui affectent un ou des domaines (ou sous-domaines) de fonctionnement de l'esprit/cerveau. Ces domaines ont été choisis par les tenants des RDoC lors de conférences de consensus (Cuthbert et Insel, 2013, p. 4). Ils sont des construits théoriques que l'on peut mesurer

(section 1.3, principe 2) selon diverses unités d'analyse²⁰ (section 1.3, principes 2 et 3). Ces six domaines de fonctionnement composés eux-mêmes de sous-construits sont inscrits dans la matrice des RDoC²¹ (voir Tableau 1.1).

 20 Il est important de mentionner que, depuis mai 2017, les tenants des RDoC affirment ceci concernant l'unité d'analyse « Gènes » :

Depuis mai 2017, nous avons supprimé les références à des gènes spécifiques de la colonne « Gènes » de la matrice des RDoC. Nous reconnaissons la pertinence évidente d'étudier les aspects génomiques des construits et des domaines des RDoC. Toutefois, l'état actuel des connaissances, dans ce domaine, met en évidence la nécessité de disposer de preuves solides d'association [entre un profil génétique et un trouble mental] résultant généralement d'études d'association suffisamment puissantes à l'échelle du génome et non d'approches par gènes candidats. Au fur et à mesure que nous réévaluons activement les informations à inclure dans cette rubrique, la matrice des RDoC sera mise à jour en conséquence. (NIMH, 2020b)

²¹ À noter que plusieurs sous-construits sont aussi constitués de sous-construits. En raison de l'espace limité, je ne liste pas les sous-construits des sous-construits dans la matrice.

Tableau 1.1 Matrice des RDoC (adaptée et traduite de NIMH, 2021b)

Domaines/construits	Unités d'analyse									
	Gènes	Molécules	Cellules	Circuits neuraux	Physiologie	Comportement	Auto- évaluation	Paradigmes		
Systèmes de valence négative										
Menace aiguë (« peur »)										
Menace potentielle (« anxiété »)										
Menace soutenue										
Perte										
Non-récompense frustrante										
Systèmes de valence positive										
Réactivité aux récompenses										
Apprentissage par les récompenses										
Évaluation des récompenses										
Systèmes cognitifs										
Attention										
Mémoire déclarative										
Langage										
Contrôle cognitif										
Mémoire de travail										
Processus sociaux										
Affiliation et attachement										
Communication sociale										
Perception et compréhension de soi										

Perception et compréhension des autres				
Systèmes d'activation et de modulation				
Éveil				
Rythmes circadiens				
Réveil-sommeil				
Systèmes sensorimoteurs				
Actions motrices				
Agentivité et titularité				
Habitude				
Patrons moteurs innés				

C'est cette représentation de la matrice que l'on rencontre le plus souvent dans les articles portant sur les RDoC (voir, par exemple, Kozak et Cuthbert, 2016, p. 290), mais elle est incomplète. Il y manque l'aspect développemental et l'aspect environnemental vus à la section 1.3 (principe 5). Cette autre représentation de la matrice intègre ces deux aspects :

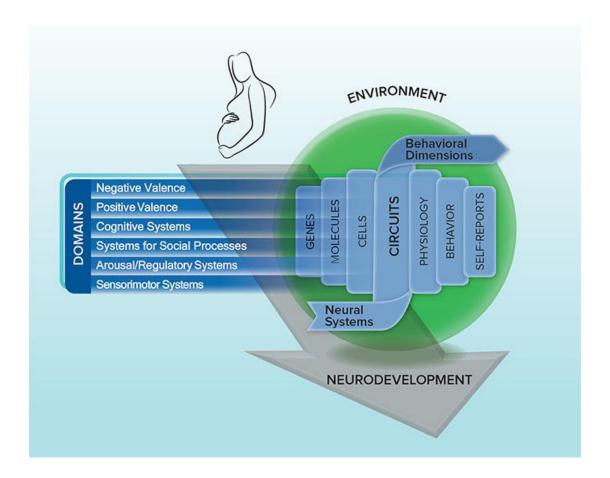


Figure 1.1 Matrice des RDoC contenant les aspects environnemental et développemental (NIMH, 2021a)

La conception du trouble mental des RDoC exprimée succinctement dans le principe 1 et explicitée par les principes 2, 3, 4 et 5 (section 1.3) est sous-jacente à cette matrice. Cette conception du trouble mental peut être mise en lumière simultanément à l'aide

d'hypothèses, de conceptions ou de théories philosophiques qui sont, à mon avis, au fondement des RDoC (pas toujours de façon explicite, mais on peut dire que son modèle du trouble mental repose implicitement et parfois explicitement sur ces hypothèses, conceptions ou théories). Ces dernières sont une version à la fois pluraliste et réductionniste du néomécanisme de Bechtel (section 1.5.2), une interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Boorse (section 1.5.3), le principe épistémoquantitatif (qui suppose ce que j'ai appelé la « variante quantitative » du réalisme des entités basé sur la robustesse d'Eronen) (section 1.5.4) et l'hypothèse de l'identité psychoneurale formulée par Bunge (section 1.5.5).

Il est important d'expliciter cet arrière-plan philosophique pour au moins trois raisons. Premièrement, le connaître enrichit notre compréhension de la conception du trouble mental des RDoC exprimée dans les cinq principes. Deuxièmement, ce cadre philosophique influe sur le contenu de ces cinq principes, lesquels permettent à leur tour un certain type de matrice. Par conséquent, ce cadre influe sur le type de recherches financées par le NIMH. Il détermine également le type de données qui est épistémiquement valables. Troisièmement, cet arrière-plan motive une certaine manière de faire de la psychiatrie clinique²². Bref, ces présupposés philosophiques «[...] ont une influence certaine et vérifiable sur le travail [que les scientifiques] produi[sent] » (Waddington, 1969, p. 72, cité dans Nicholson et Dupré, 2018, p. 38) et, éventuellement, sur la pratique psychiatrique.

²² À ce propos, Faucher et Goyer (2016, p. 420-424) ont décrit certaines conséquences néfastes que les RDoC pourraient avoir sur la clinique. Par exemple, si une clinicienne ou un clinicien adopte une conception biologique du trouble mental, son empathie envers ses patientes ou ses patients baisse (Lebowitz et Ahn, 2014). Cela est regrettable parce que, lorsqu'une clinicienne ou un clinicien est empathique, la relation clinique est meilleure et la patiente ou le patient a plus de chance de guérir (Potter, 2013, p. 299-300).

1.5.2 Une version pluraliste et réductionniste du néomécanisme de Bechtel

Les représentants des RDoC appréhendent les troubles mentaux dans le cadre de ce que l'on nomme le néomécanisme (lequel doit être distingué des philosophies mécanistes et atomistes de l'antiquité grecque et de la philosophie mécanique du XVII siècle avec lesquelles il entretient quelques liens historiques et conceptuels. Voir Glennan, 2017, p. 6-7). La conception du néomécanisme varie un peu selon les philosophes qui l'ont développée (Bechtel, 2007; Bunge, 1967, p. 28-49, 1997; Craver, 2007; Craver et Darden, 2013; Glennan, 1996, 2002, 2017; Machamer et al., 2000; Thagard, 1999; Wright et Bechtel, 2007). Les partisans des RDoC semblent adopter une version pluraliste et réductionniste du cadre néomécaniste élaboré par Bechtel qui, comme je l'explique à la fin de la section, partage un air de famille avec le réductionnisme fragmentaire [patchy reductionism] de Schaffner (2006, 2008, 2011, 2013, 2016, chapitre 5) (les tenants des RDoC font référence à Bechtel dans Cuthbert, 2014b, p. 31; Cuthbert et Kozak, 2013, p. 931 et 934; Kozak et Cuthbert, 2016, p. 292 et 293). Pour cette raison, je présente ici seulement le cadre néomécaniste développé par ce dernier et je fais l'économie des débats qu'ont entre eux les partisans de cette branche en philosophie des sciences.

De leur côté, si les philosophes et autres scientifiques ne mentionnent pas explicitement le néomécanisme de Bechtel quand ils font une analyse critique des RDoC, ils s'entendent tout de même sur l'idée générale selon laquelle l'investigation des mécanismes (dysfonctionnels) sous-jacents aux troubles mentaux est au cœur de ce programme de recherche. Par exemple, les philosophes Zachar et Hoffman affirment que « [les RDoC] cherchent à comprendre la psychopathologie en étudiant les fonctions psychologiques et comportementales de base [...] et leurs mécanismes biologiques » (Hoffman et Zachar, 2017, p. 59-60; voir Sullivan, 2016, p. 1 pour un avis semblable).

Mais comment définir exactement ce que sont ces mécanismes sous-jacents aux diverses activités cognitives et comportementales que les tenants des RDoC veulent découvrir? Bechtel et Abrahamsen conçoivent un mécanisme comme «[...] une structure remplissant une fonction grâce à ses composants, aux opérations de ses composants et [à] leur organisation. [Aussi, ajoutent-ils], le fonctionnement orchestré du mécanisme est responsable d'un ou plusieurs phénomènes » (Bechtel et Abrahamsen, 2005).

Dans cette définition, au moins trois éléments doivent être explicités (voir Bechtel, 2008, p. 14-17). Premièrement, il y a l'idée selon laquelle un mécanisme est responsable d'un phénomène. Bechtel (2008, p. 14) dit que l'on doit comprendre un phénomène, suivant Woodward et Bogen (1988), comme une occurrence dans le monde qui est parfois singulière (comme le Big Bang), mais qui est souvent répétable (comme la ventilation pulmonaire ou les mécanismes attentionnels chez l'humain). Dans les recherches menées dans le cadre des RDoC, les phénomènes étudiés sont des phénomènes répétables. On trouve parmi eux les processus perceptuels, les processus mnémoniques (ex. : la mémoire de travail) et la régulation des émotions (qui tous peuvent être dysfonctionnels dans les troubles de santé mentale).

Deuxièmement, il y a l'idée selon laquelle les mécanismes sont constitués de parties et d'opérations (2008, p. 14). Bechtel dit que les *parties* sont les *composants structurels* du mécanisme et que les *opérations* réfèrent aux processus ou aux changements impliquant ces composants. Bechtel spécifie qu'il utilise le terme « opération » pour insister sur le fait qu'un composant fait une opération *sur* un autre composant. Le terme « opération » implique donc minimalement deux composants. En d'autres mots, le concept « opération » sous-entend qu'un composant ne peut pas opérer sur rien. Par exemple, l'eau dissout le sel ou le cœur qui se contracte pompe le sang.

Troisièmement, il y a l'idée selon laquelle les composants et les opérations d'un mécanisme sont organisés de manière appropriée (Bechtel, 2008, p. 17). Il en est ainsi parce que le bon fonctionnement du mécanisme requiert que ces différentes opérations soient coordonnées les unes avec les autres. Découvrir cette organisation est nécessaire pour comprendre le phénomène produit par le mécanisme. Il existe différents types d'organisation. Dans les mécanismes biologiques, très souvent, ces organisations ne sont pas linéaires (comme dans une chaîne de montage), mais complexes (Bechtel, 2008, p. 17). Par exemple, elles comprennent des mécanismes de rétroaction comme on en observe dans la régulation de la glycémie dans le corps humain. Enfin, dans cette organisation, les relations coordonnées entre les composants du mécanisme ont lieu sur divers niveaux. Par exemple,

[1]e mécanisme de la mémoire spatiale a plusieurs niveaux, dont certains comprennent des organes tels que l'hippocampe qui génère une carte spatiale, d'autres impliquent les interactions cellulaires qui sous-tendent la génération de cette carte, et d'autres encore impliquent les mécanismes moléculaires qui sous-tendent ces interactions cellulaires. (Craver et Tabery, 2015, section 2.4.5)

L'adoption du néomécanisme a au moins deux implications. La première est *ontologique*. Elle dit qu'au sein des RDoC on conçoit les troubles mentaux comme des phénomènes résultant des opérations de *mécanismes* cérébraux *dysfonctionnels*²³. En d'autres termes, lorsqu'il y a présence d'un trouble mental, les entités qui composent les mécanismes n'opèrent pas comme elles le font quand ces mécanismes fonctionnent normalement.

²³ On verra ce que signifie exactement « dysfonctionnel » quand j'aborderai mon interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Boorse à la section 1.5.3. Pour l'instant, une compréhension intuitive de ce que signifie ce terme suffira pour comprendre mon propos.

Plusieurs mécanismes dysfonctionnels peuvent être sous-jacents à un trouble mental.

Par exemple, pour la dépression majeure, on pourrait citer une *dysfonction*

de l'axe hypothalamo-hypophyso-surrénalien (HHS), des circuits cérébraux responsables de la recherche de récompense, des circuits neuraux responsables de la régulation des émotions, des systèmes qui modulent la neurotransmission, des systèmes responsables de la cognition et [une dysfonction] dans les marques épigénétiques. (Cuthbert et Insel, 2013, p. 3)

Adopter le cadre néomécaniste a aussi des implications épistémologiques. En effet, au sein des RDoC, c'est souvent en termes mécanistes que les chercheuses et les chercheurs conçoivent leurs hypothèses et leurs protocoles expérimentaux. Par exemple, Sukhodolsky et son équipe font l'hypothèse qu'une thérapie cognitivocomportementale (TCC) suivie par un enfant présentant des comportements perturbateurs entraînera, chez ce dernier, une activité moins grande des mécanismes responsables du fonctionnement de son amygdale, une activité plus grande des mécanismes responsables de l'activité des régions cérébrales impliquées dans le contrôle des émotions et, conséquemment, une réduction des comportements d'agression réactive (Sukhodolsky, 2014; Sukhodolsky et al., 2016, p. 39) (voir section 1.4.2).

C'est aussi dans ce cadre néomécaniste que les chercheuses et les chercheurs formulent des explications que Bechtel qualifie de mécanistes.

L'explication mécaniste consiste à représenter (verbalement ou sous forme de diagrammes) ²⁴ les parties, les opérations et l'organisation [du

²⁴ Le type de représentations énumérées dans cette parenthèse est incomplet. En effet, Bechtel et Abrahamsen disent ailleurs que « ces représentations peuvent être des représentations mentales internes, mais [qu'] elles peuvent aussi prendre la forme de représentations externes à l'agent cognitif : diagrammes, descriptions linguistiques, équations mathématiques, modèles physiques, etc.» (Bechtel et

mécanisme], et à montrer (souvent par l'intermédiaire de simulations mentales ou informatiques) comment un tel mécanisme produit le phénomène en question. (Bechtel, 2008, p. 48)

En d'autres termes, expliquer un phénomène consiste à faire comprendre, par l'intermédiaire d'une représentation ou, plus spécifiquement, d'un *modèle* du mécanisme ²⁵, *pourquoi* ce phénomène advient en montrant *comment* les parties organisées entre elles dans ce mécanisme opèrent et le produisent.

Le modèle suivant est une explication mécaniste du phénomène de la transmission synaptique :

Abrahamsen, 2005, p. 425). Suivant cela, les modèles peuvent autant être des descriptions verbales, des animations graphiques ou des équations mathématiques (comme les équations de Hodgkin et Huxley qui représentent l'excitabilité cellulaire).

²⁵ Plutôt que d'utiliser le terme «représentation», j'utiliserai maintenant le terme «modèle», lequel par définition est une entité qui «*représente* une partie ou un aspect choisi du monde» (Frigg et Hartmann, 2020, section 1; mes *italiques*). Ce choix de terminologie est cohérent avec la position de Bechtel et Abrahamsen qui affirment que «[l]es explications dans les sciences de la vie consistent souvent à présenter un *modèle* du mécanisme considéré comme responsable d'un phénomène donné» (Bechtel et Abrahamsen, 2005, p. 421, mes *italiques*).

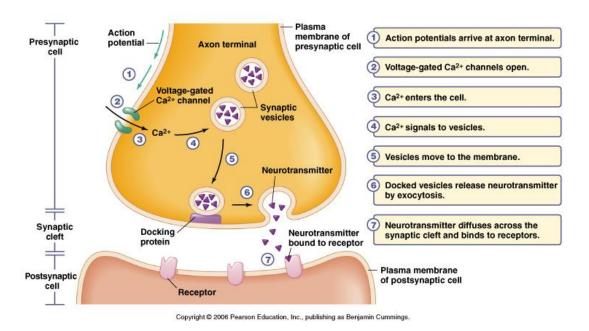


Figure 1.2 Modèle mécaniste du phénomène de la transmission synaptique (Moyes et Schulte, 2006, chapitre 5)

Comme l'affirment Bechtel et Abrahamsen (2005, p. 422), si un modèle mécaniste « explique *pourquoi* en expliquant *comment* », alors le modèle ci-dessus explique *pourquoi* a lieu la transmission synaptique en montrant en sept étapes (il aurait pu y en avoir plus) *comment* ses parties organisées entre elles opèrent et la produisent : (1) le potentiel d'action atteint la terminaison axonale, (2) les canaux à ions de calcium (Ca²⁺) s'ouvrent, (3) le Ca²⁺ entre dans la cellule, (4) le Ca²⁺ « signale » sa présence aux vésicules, (5) les vésicules remplies de neurotransmetteur se déplacent vers la membrane plasmique, (6) les vésicules contenant les neurotransmetteurs se fusionnent avec l'extrémité de la membrane plasmique de l'axone et les neurotransmetteurs sont libérés (il s'agit de l'exocytose des vésicules) et (7) les neurotransmetteurs sont diffusés dans l'espace synaptique et se fixent aux récepteurs de la cellule postsynaptique.

Trois grandes idées de Bechtel (2008, p. 17-22) sur l'explication mécaniste s'appliquent aux RDoC et me semblent importantes à garder en tête. La première est

qu'il faut adopter une conception épistémique de l'explication (que Bechtel préfère à une conception ontique. Pour plus de détails à ce sujet, voir Bechtel, 2008, section 1.3; voir aussi la section 5.2 dans Wright et Bechtel, 2007)²⁶. L'idée cardinale sous-jacente à cette conception est qu'une explication réfère avant tout à une forme d'activité cognitive chez les êtres pensants. Plus spécifiquement, en recherche, cette conception réfère au travail mental qu'opèrent les scientifiques (Bechtel, 2008, p. 18). Cette activité mentale peut prendre plusieurs formes. Par exemple, elle peut être un raisonnement inférentiel à partir de propositions. Elle peut également être une

Les explications objectives ne sont pas des textes; elles sont des choses à part entière. Ce sont des faits, pas des représentations. Elles sont des types de choses qui sont découvertes et décrites. Il n'est pas question d'explications objectives étant « exactes » ou « inexactes », ou « bonnes » ou « mauvaises ». Elles sont simplement là [dans la structure causale du monde]. (Craver, 2007, p. 27)

Craver pense que c'est à partir des explications objectives que l'on doit déterminer les critères normatifs d'une bonne explication. De fait, il dit que «[l]es explications objectives, [c'est-à-dire] les causes et les mécanismes dans le monde, sont le bon point de départ pour réfléchir aux critères d'évaluation des textes explicatifs en neurosciences » (Craver, 2007, p. 27).

Bechtel s'oppose à la conception ontique. De fait, il affirme :

Le problème avec cette conception ontique est que les mécanismes ne s'expliquent pas euxmêmes. Ils sont opérants dans le monde qu'il y ait ou qu'il n'y ait pas des scientifiques investis à leur donner une explication. Expliquer est une activité de scientifiques qui doivent faire un travail mental pour offrir une explication. (Bechtel, 2008, p. 17-18)

Ce que dit ici Bechtel est que, dans l'explication d'un mécanisme, les éléments non ontiques comme les raisonnements des scientifiques (sur la manière dont les opérations des parties d'un mécanisme produisent le phénomène, par exemple) ont *préséance* sur les éléments ontiques de l'explication pris en eux-mêmes comme ceux de parties, d'opérations des parties et d'organisation du mécanisme.

²⁶ Au bénéfice de la lectrice ou du lecteur, il vaut la peine de s'attarder brièvement sur le débat concernant ce que devrait être la nature de l'explication mécaniste. Selon les tenants de la conception ontique, la propriété principale d'une explication n'est pas d'être une activité cognitive d'un être humain (comme le soutiennent les tenants de la conception épistémique), mais d'être un processus causal dans le monde. Cette affirmation de Craver illustre bien cette idée :

simulation mentale. Quoi qu'il en soit, cette activité mentale aboutit ultimement en la production d'un modèle du mécanisme qui intéresse les chercheuses et chercheurs.

Selon la conception épistémique de l'explication, un modèle est un outil heuristique qui influence le type de processus cognitifs que feront ou pourront faire les scientifiques (et, peut-on penser, les cliniciennes et les cliniciens en santé mentale qui raisonneraient selon de tels modèles). Par exemple, le type de modèle que produisent les scientifiques pour expliquer un phénomène influence le type d'hypothèses que développeront celles-ci et ceux-ci. Dans le cadre des RDoC, plusieurs hypothèses, protocoles et inférences sont influencés par le cadre mécaniste. Par exemple, Sharp et ses collègues (2014) ont étudié l'hypothèse selon laquelle une faible activité du striatum ventral — lequel est une structure sous-corticale que l'on peut concevoir comme un *mécanisme* biologique impliqué dans plusieurs fonctions comme celle qui consiste à traiter les stimuli affectifs tels que les récompenses — est un endophénotype²⁷ prédicteur de la dépression majeure (DM) (voir section 2.1 pour avoir plus de détails sur cette étude).

La deuxième idée est que les explications mécanistes (qui prennent la forme d'un modèle) ne contiennent pas de lois du type de celles que l'on retrouve en physique (Bechtel, 2008, p. 48). En cela, les explications mécanistes se distinguent des explications ayant la forme que commande le modèle déductif-nomologique développé par les empiristes logiques (Hempel et Oppenheim, 1948). Selon ce dernier, en effet, une explication a la forme d'un argument composé de propositions et le phénomène

²⁷ Les endophénotypes sont définis par Gottesman et Gould (2003, p. 636) comme des « composants mesurables invisibles à l'œil nu [se situant] le long de la voie entre la maladie [(que l'on peut définir comme les symptômes apparents à l'œil nu que l'on appelle aussi le phénotype)] et le génotype distal ». Aussi, ajoutent-ils (2003, p. 636), « [u]n endophénotype peut être de nature neurophysiologique, biochimique, endocrinologique, neuroanatomique, cognitive ou neuropsychologique (y compris les données d'autoévaluation [...]).

que l'on veut expliquer est déduit de conditions initiales et de lois non probabilistes (comme les lois du mouvement de Newton) ou probabilistes (comme en mécanique quantique). Ces dernières sont universelles au sens où elles ne sont pas locales et s'appliquent à tous les mondes possibles. Dans le cadre des RDoC, plutôt que d'utiliser ce type de lois pour rendre compte des troubles mentaux, les chercheuses et les chercheurs produisent et découvrent des généralisations invariantes (au sens de Woodward, 2003) qui prennent la forme de réseaux nomologiques (voir section 1.5.4 pour plus de détails à ce sujet).

La troisième idée est qu'une explication (et un modèle) mécaniste est nécessairement multiniveaux (Bechtel, 2008, p. 22). En effet, « au cœur de la perspective mécaniste se trouve l'hypothèse que différents types de relations causales sont à trouver entre les composantes d'un mécanisme et entre le mécanisme et les choses dans son environnement » (Bechtel, 2008, p. 21-22).

Il est important de mentionner cette idée parce qu'on pourrait croire à première vue que l'explication mécaniste est réductionniste, c'est-à-dire qu'elle vise ultimement à représenter les parties opérantes du mécanisme qui se situent au niveau fondamental de la réalité²⁸ et qui produisent le phénomène. Mais ce n'est pas le cas. En effet, comme cela est mentionné dans la citation précédente, l'environnement, quel qu'il soit (biologique, abiotique, social, etc.), fait *toujours* partie de l'organisation dans laquelle se trouve le mécanisme à expliquer. Cela implique qu'une explication mécaniste, pour être *complète*, doit *nécessairement* tenir compte de celui-ci. On doit montrer, dans ce type d'explication, la manière dont l'environnement — lui-même un mécanisme occupant un niveau plus haut que le mécanisme que l'on étudie — contraint l'activité des parties du mécanisme à l'origine du phénomène. Suivant les principes sous-jacents

²⁸ À noter que certains pensent qu'il n'existe pas un tel niveau (voir, par exemple, Schaffer, 2003).

aux RDoC (voir section 1.3), les explications produites à la suite des recherches menées en son cadre devraient être multiniveaux. Mais ce ne semble pas être tout à fait le cas. En effet, si les explications formulées dans le cadre des RDoC citent des mécanismes psychologiques et biologiques, elles ne citent pas vraiment de mécanismes environnementaux (physiques, sociaux et culturels) (voir la section 2.1 pour plus de détails à ce sujet; voir aussi la section 2.1.1 pour comprendre à quoi pourrait ressembler un mécanisme physique ou social dans une explication mécaniste d'un trouble mental).

Il est intéressant de poursuivre cet exposé du néomécanisme de Bechtel en expliquant ce qui distingue, selon lui, les mécanismes biologiques (ex. : le pompage du sang) des mécanismes mentaux (ex. : perceptuels, attentionnels et mnémoniques). Ce qui caractérise les mécanismes mentaux est qu'ils traitent de l'information alors que les mécanismes biologiques (à l'exception peut-être des mécanismes génétiques) impliquent seulement des transformations physiques de substances matérielles (Bechtel, 2008, section 1.4).

Pour illustrer ce qu'est un mécanisme strictement biologique, on peut développer un peu le cas du mécanisme de la respiration cellulaire mentionné rapidement par Bechtel (2008, p. 22-23). Essentiellement, celle-ci consiste en la transformation du glucose par les cellules en molécules d'adénosine triphosphate (ATP), lesquelles contiennent de l'énergie. L'ATP est nécessaire à tous les processus vitaux (comme la division cellulaire) ayant lieu dans les organismes vivants. La respiration cellulaire (aérobie)²⁹ peut être représentée par un modèle mécaniste où figurent trois stades métaboliques où ont seulement lieu des transformations physiques de substances matérielles, à savoir

²⁹ La respiration aérobie doit être distinguée de la respiration anaérobie. En effet, la première nécessite du dioxygène pour produire de l'ATP et la deuxième n'en utilise pas.

- (1) la glycolyse, (2) le cycle de l'acide citrique et (3) la phosphorylation oxydative (voir figure 1.2) :
 - (1) La *glycolyse* a lieu quand une molécule de glucose contenu dans le plasma sanguin entre dans le cytosol de la cellule et y est dégradée en deux molécules de pyruvate.
 - (2) Le pyruvate pénètre dans les mitochondries où il est oxydé. Ce processus d'oxydation transforme le pyruvate en acétyl-CoA (acétyl-coenzyme A). L'acétyl-CoA entre dans le *cycle de l'acide citrique* où prend éventuellement fin la dégradation du glucose en dioxyde de carbone.
 - (3) Enfin, comme l'expliquent Campbell et Reece (2012, p. 188, mes *italiques*), « le NADH et une coenzyme similaire appelée FADH₂ transfèrent les électrons provenant du glucose [qui vient d'être oxydé] à des chaînes de transport d'électrons qui sont insérées dans la membrane mitochondriale interne. Durant la *phosphorylation oxydative*, les chaînes de transport d'électrons convertissent l'énergie chimique en une forme d'énergie qui sert à la synthèse de l'ATP au cours d'un processus appelé chimiosmose ».

Pour une molécule de glucose, on note que le rendement énergétique de la glycolyse est de deux molécules d'ATP. Le cycle de l'acide citrique produit lui aussi deux molécules d'ATP. La phosphorylation oxydative, quant à elle, est la plus rentable énergétiquement. En effet, elle produit de vingt-six à vingt-huit molécules d'ATP. Donc, la respiration cellulaire est un mécanisme de fabrication d'énergie très efficace. En effet, pour une molécule de glucose sont produites de trente à trente-deux molécules d'ATP (Campbell et Reece, 2012, p. 197).

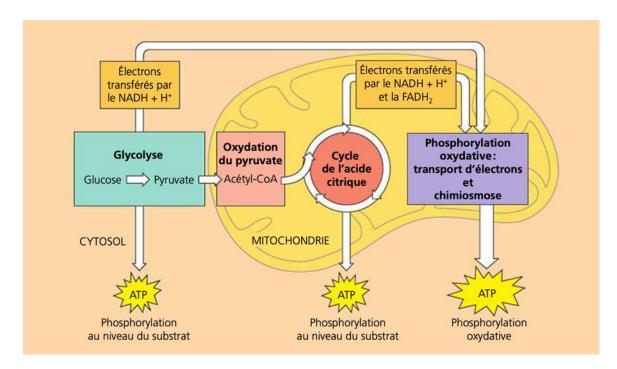


Figure 1.3 Les trois stades métaboliques de la respiration cellulaire (Campbell et Reece, 2012, p. 188)

Maintenant qu'à été exemplifié ce qu'est un mécanisme purement biologique, on peut illustrer ce qu'est un mécanisme mental avec les fonctions perceptuelles. Celles-ci, bien qu'elles impliquent des transformations matérielles (des échanges ioniques, l'activité de certains circuits cérébraux, etc.), sont définies comme « mentales » parce qu'elles impliquent du traitement d'informations. Bechtel, suivant Dretske (1981), conçoit l'information en termes causaux. Plus spécifiquement, il l'appréhende comme l'« effet régulier d'une cause qui permet de déduire les caractéristiques de cette cause à partir des caractéristiques de l'effet » (Bechtel, 2008, p. 24). Par exemple, la perception visuelle que j'ai d'un arbre (l'effet) dans mon esprit est causée par l'arbre dans le monde (la cause) devant moi. Et c'est par l'entremise des caractéristiques de l'effet (celles de ma perception de l'arbre) que je peux déduire les caractéristiques de la cause (l'arbre dans le monde). Comme on peut le constater, suivant cette définition de l'information, l'effet représente la cause (Bechtel, 2008, p. 24) ou, autrement dit, la

présente à nouveau dans l'esprit. C'est pourquoi l'on dit de cet effet qu'il est une *représentation* (Bechtel, 2008, p. 24). Bref, suivant ce qui précède, on peut dire qu'un mécanisme mental est un mécanisme qui traite des représentations qui transportent « des informations sur les objets, les événements et les circonstances rencontrés dans le moment présent ou précédemment » (Bechtel, 2009, p. 553)³⁰.

Dans le cadre des RDoC, ce sont surtout des mécanismes mentaux normaux et anormaux (comme l'attention, la mémoire ou la perception et la compréhension de soi et des autres) qui sont étudiés et représentés à l'aide de modèles. Les mécanismes mentaux que ces modèles mettent en scène sont conçus selon les différentes unités d'analyse que l'on trouve dans la matrice des RDoC (voir section 1.3, principe 2). Selon Bechtel (2008, p. 21-22), aucune de ces unités d'analyse ne devrait avoir de préséance épistémique sur les autres unités. C'est aussi ce que soutiennent les tenants des RDoC qui disent avoir choisi d'utiliser le terme « unité » d'analyse plutôt que le terme « niveau » d'analyse parce que ce dernier pourrait laisser sous-entendre à tort qu'il existe des niveaux d'analyse épistémiquement plus fondamentaux que d'autres (Kozak et Cuthbert, 2016, p. 293). Une conséquence de cela est que, lorsqu'on tente de valider un construit (par exemple, celui de la menace soutenue), les mesures formulées selon l'unité d'analyse « auto-évaluation » des patients ne devraient pas avoir de priorité épistémique sur les mesures formulées selon les autres unités d'analyse (sur la validation, voir la section 1.5.4). Les tenants des RDoC s'inscrivent donc en faux relativement à la tradition subjectiviste suivant laquelle ce qu'expérimente un individu a une priorité épistémique dans l'explication de son expérience. En fait, selon Kozak

³⁰ La notion de représentation est une notion centrale dans la conception computationnelle de la cognition. Bechtel adopte cette conception. Puisque les tenants des RDoC adoptent le néomécanisme de ce dernier, on pourrait penser qu'ils adoptent comme lui une conception computationnelle de la cognition. Bien que je ne développe pas cette idée dans cette thèse, je présente les grandes lignes de cette conception computationnelle dans l'annexe A.

et Cuthbert (2016, p. 292), « les expériences rapportées ont le statut logique d'hypothèses faillibles à propos du fonctionnement de l'énonciateur (Kozak et Miller, 1982; Miller et Kozac, 1993) ». Dans cette perspective, les expériences subjectives des patientes et patients doivent être étudiées et évaluées essentiellement selon la méthode scientifique. Les tenants des RDoC affirment à ce sujet qu'ils adoptent une méthode semblable à ce que le philosophe Daniel Dennett (1991, p. 72-81) nomme l'hétérophénoménologie (Kozak et Cuthbert, 2016, p. 292), laquelle « n'est rien d'autre que la bonne vieille méthode scientifique à la troisième personne appliquée aux phénomènes particuliers de la conscience humaine (et animale) » (Dennett, 2001).

Ces affirmations sur l'égalité de l'intérêt épistémique des unités d'analyse des RDoC peuvent être contestées parce que ces derniers contiennent un principe qui exige que soient *toujours* citées, dans les modèles des troubles mentaux (ou de certains symptômes de ces troubles), les composantes biologiques des construits théoriques (comme l'attention ou la perception) qui sont à l'étude (Kozak et Cuthbert, 2016, p. 292; voir aussi le principe 3 à la section 1.3). On peut donc penser que les tenants des RDoC accordent une importance particulière aux mécanismes qui sont conçus selon les unités biologiques d'analyse (« gènes », « molécules », « cellules », « circuits neuraux » et « physiologie »). On peut aussi penser que l'explication mécaniste est *complète*, selon eux, dans la mesure où ces mécanismes biologiques y sont cités. En cela, les tenants des RDoC adoptent une posture épistémique qui s'apparente à celle adoptée typiquement en neurosciences psychiatriques (voir Bickle et Kostko, 2017).

Cette posture ambigüe, parce qu'à la fois pluraliste et réductionniste, fait croire à Faucher et Goyer (2015, 2016, 2017) que les tenants des RDoC n'adoptent pas le *réductionnisme traditionnel* (Nagel, 1961, chapitre 11) selon lequel la réduction est l'explication totale d'une théorie (un ensemble de lois) ou d'une science de haut niveau par une théorie plus fondamentale de bas niveau (par exemple, la réduction de la

biologie par la chimie) (Schaffner, 2016, p. 122). Pour cette raison, Faucher et Goyer suggèrent que les tenants des RDoC n'adoptent pas l'idée (inhérente à ce type de réductionnisme) selon laquelle la théorie de bas niveau explique tout (et même mieux et davantage) ce qu'explique la théorie de haut niveau. Par exemple, ces derniers n'ont pas pour objectif de réduire une théorie cognitive de la dépression majeure (TcogDM) à l'aide d'une théorie biomoléculaire de la dépression majeure (TbioDM). Les tenants des RDoC ne pensent pas non plus que tout ce qu'explique ou que peut prédire la TcogDM peut l'être par la TbioDM.

Mais si les tenants des RDoC n'adoptent pas le réductionnisme traditionnel qu'adoptent-ils? Faucher et Goyer proposent que, suivant Sanislow et ses collègues (2010, p. 633) qui citent Kendler (2005), la posture épistémique à la fois néomécaniste, pluraliste et réductionniste que promeuvent les tenants des RDoC soit saisie par le réductionnisme fragmentaire [patchy reductionism] développé par Schaffner (2006, 2008, 2011, 2013, 2016, chapitre 5).

Je termine cette section en faisant connaître cinq propriétés des réductions que suppose ce type de réductionnisme qui (a) s'harmonise bien avec cette version pluraliste *et* réductionniste du cadre néomécaniste élaboré par Bechtel et qui (b) selon Faucher et Goyer, est différent du réductionnisme traditionnel :

(1) Elles impliquent des modèles « mécanistes » : contrairement au modèle traditionnel de la réduction qui implique des lois représentées sous forme d'axiome, le modèle de Schaffner implique des modèles mécanistes [tel qu'ils sont décrits précédemment]. On trouve, par exemple, ce genre d'explication en neuroscience moléculaire lorsqu'un groupe de [chercheuses et] de chercheurs tente d'expliquer un phénomène comme l'attention ou la mémoire. Ces [...] [modèles] sont souvent représenté[s] de façon graphique dans les figures que l'on retrouve dans leurs articles.

- (2) Elles sont interniveaux: les mécanismes étiologiques qui sont décrits afin d'expliquer une maladie sont typiquement inter-[unités d'analyse 31] (c'est-à-dire qu'ils impliquent des facteurs situés à différent[e]s [unités d'analyse]). [C'est dans ce sens que l'on peut dire que ces explications réductrices sont pluralistes.] Par exemple, un modèle de la dépression invoquera des changements structurels dans les cellules nerveuses de l'axe hypothalamo-hypophyso-surrénalien ([...][unité d'analyse] physiologique) causés par des événements stressants, comme l'abus physique ou psychologique (niveau social32), qui ont un effet sur le développement d'un style attributif particulier et/ou une image de soi négative ([unité d'analyse auto-évaluation33]).
- (3) Elles sont incomplètes: le plus souvent elles n'expliquent pas toutes les instances d'un phénomène. Par exemple, des facteurs génétiques qui expliquent certains cas de schizophrénie ou de dépression majeure n'expliquent pas d'autres cas des mêmes maladies. De la même façon, certains modèles cognitifs expliqueront une partie des symptômes de patients atteints d'une maladie, mais pas d'autres symptômes propres à cette maladie. [Comme je l'ai dit précédemment,] [c]es réductions sont incomplètes également en ce qu'elles ne sont pas la réduction d'une théorie complète de haut niveau à une autre théorie complète de bas niveau. Schaffner dira que ces réductions sont « à la pièce » ou « locales » (Schaffner, 2011, p. 144).
- (4) Elles sautent des niveaux : c'est une autre façon de dire qu'elles sont incomplètes. Par exemple, une liaison entre une mutation génétique et un comportement ou un désordre peut être établie, mais quelques-uns des mécanismes biomoléculaires [par l'intermédiaire desquels] [...] le gène influe sur la condition sont inconnus. Dans ce type de cas, la représentation de la voie qui mène à la maladie est « lacunaire » [gappy] et l'explication est une esquisse d'explication (Schaffner, 2008, p. 75).
- (5) Elles sont « touffues » [bushy]: typiquement, l'explication d'un phénomène n'implique pas une relation directe entre un gène et un comportement, mais plutôt une espèce de buisson causal où plusieurs

-

³¹ Dans la citation originale, la notion de niveau est utilisée. Je lui préfère la notion d'unité parce que, comme on l'a vu précédemment dans cette section, les tenants des RDoC utilisent cette notion plutôt que la notion de niveau.

³² Ce niveau social réfère, dans la matrice, à l'aspect « environnement » (voir section 1.5.1, tableau 1.2).

³³ Dans la citation originale, il est écrit « niveau psychologique ».

gènes agissent de concert (mais peut-être pas tous en même temps) dans un contexte particulier impliquant des boucles de rétroaction causale (y compris des boucles à l'intérieur et à l'extérieur de l'individu) [et des processus épigénétiques]. Le résultat des réductions dans les disciplines biologiques (dont la psychiatrie) n'est donc pas une théorie axiomatique. (Faucher et Goyer, 2016, p. 133, 2017, p. 417)

1.5.3 Une interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Boorse

On peut soutenir qu'une interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Boorse (1975, 1976a, 1977, 1997, 2014) est sous-jacente à la conception du trouble mental des RDoC, et ce, même si les tenants de ce cadre de recherche ne l'adoptent pas explicitement (voir Porter, 2019, p. 10-12 qui soutient une thèse semblable). Dans ce qui suit, d'abord, je présente les notions de la TBS de fonction normale, de classe de référence et de design fonctionnel et la manière dont ces dernières sont reliées les unes avec les autres. J'explique ensuite comment et pourquoi on peut interpréter la TBS de manière adaptationniste. Je pense que les propos des défenseurs des RDoC et ceux de leurs critiques pointent vers cette interprétation de la TBS et, plus spécifiquement, de sa notion de fonction normale. Après, j'expose la notion de santé qui découle de cette dernière. Puis, je présente le concept de dysfonction de la TBS et montre qu'il saisit le contenu des propos sur les troubles mentaux que formulent les tenants des RDoC. Enfin, j'expose brièvement l'idée de ces derniers que suppose la TBS et selon laquelle les troubles sont des espèces naturelles.

On peut maintenant commencer l'exposition de la TBS. Selon cette dernière, une *fonction normale* d'une partie (par exemple, d'un organe) d'un système vivant est une contribution causale statistiquement typique à la survie et à la reproduction individuelle

au moment présent³⁴ (et, plus précisément, dans une situation spécifique, comme on le verra plus loin) (Boorse, 1977, p. 555-556). Le groupe de systèmes vivants relativement auquel cette contribution causale de cette partie est jugée comme étant statistiquement typique est ce que Boorse appelle une *classe de référence*. Celle-ci est, dans le monde réel, un groupe de systèmes vivants d'âge et de sexe identiques au sein d'une espèce (Boorse, 1977, p. 555). Chaque système vivant d'une classe de référence partage un répertoire de fonctions normales avec les autres systèmes vivants de sa classe. Ce répertoire de fonctions normales est le *design fonctionnel* des individus de cette classe de référence.

On peut illustrer ce qu'est une fonction normale avec le cas de la queue de l'écureuil chez les écureuils mâles de deux ans de la famille des Sciuridés (une famille de mammifères rongeurs). Chez tous les écureuils appartenant à cette classe de référence, en principe, la queue typiquement donne de l'équilibre — et donc les empêche de tomber au sol — quand ils sautent d'arbre en arbre (Kingma, 2010, p. 244). Pour cette raison, la queue des écureuils appartenant à cette classe de référence contribue de manière statistiquement typique à la survie et à la reproduction de ces derniers. C'est

³⁴ Selon Boorse, une fonction normale doit être adaptative dans le moment présent. Cette conception de la fonction est différente de la conception de l'effet sélectionné (que l'on appelle aussi « conception étiologique »). Wakefield (1992, 2006a, 2011), entre autres, dont on présentera la conception du trouble au chapitre 2, adopte cette conception de la fonction. Selon cette dernière, une fonction est un effet d'une partie (d'un mécanisme, d'un organe, etc.) qui a contribué, au cours de l'histoire évolutive, à la survie et à la reproduction des ancêtres de l'organisme présentant cette partie et qui, pour cette raison, a été sélectionné. Il faut noter que la propriété selon laquelle la fonction normale boorséenne doit être adaptative dans le moment présent n'empêche pas que les fonctions normales soient conçues comme étant, dans les faits, issues de l'évolution par la voie de la sélection naturelle. Cela implique toutefois qu'un trait qui aurait contribué au succès reproducteur des ancêtres d'un organisme par le passé et qui n'y contribuerait plus en ce moment ne serait plus considéré, dans le cadre de la TBS, comme ayant une fonction. Je pense que cette implication de la TBS est, en dernière instance, intenable. En effet, comme je le soutiens plus loin dans cette section, suivant Garson (2019a, p. 1148-1149) qui se base lui-même sur les affirmations de Boorse (2002, p. 99), on ne peut pas déterminer ce qui est une contribution causale statistiquement typique à la survie et à la reproduction sans tenir compte de l'histoire évolutive des ancêtres du système vivant ayant cette fonction.

pourquoi on dit que cette fonction de la queue de l'écureuil fait partie du répertoire des fonctions *normales* de ces écureuils.

Il est à noter que ce ne sont pas toutes les contributions causales des parties à la survie et à la reproduction d'un système vivant qui sont des fonctions *normales*. En effet, certaines contributions causales à la survie et à la reproduction ne sont pas typiques (ou habituelles) aux parties des membres d'une classe de référence. Par exemple, le fait que la queue d'un écureuil, appartenant à la classe de référence susmentionnée, puisse se coincer dans une fissure au sol et empêcher cet écureuil de se faire écraser par une auto n'est pas une contribution causale typique de la queue de l'écureuil pour la survie et la reproduction de ce dernier (Boorse, 1977, p. 557). Dans ce cas de figure, la queue de l'écureuil contribue plutôt de manière accidentelle à la survie et à la reproduction de celui-ci. Pour cette raison, cette fonction de la queue de l'écureuil ne fait pas partie du *design fonctionnel* des membres de la classe de référence de ces mammifères rongeurs.

Dans le cadre de la psychiatrie, les fonctions normales dont il est question sont de nature psychologique ou mentale (Boorse, 1976a, p. 63). Boorse soutient que toutes ces fonctions mentales normales sont des fonctions biologiques au même titre que les fonctions sexuelles ou d'excrétion (Boorse, 1976a, p. 64). Pour être qualifiée de mentale, une fonction normale doit présenter nécessairement, selon lui, deux caractéristiques : (1) elles jouent un rôle causal dans l'action et (2) elles contribuent à la production d'actions *typiques* de l'espèce contribuant à la survie et à la reproduction de l'organisme (Boorse, 1976a, p. 63-64, mes *italiques*).

Boorse (1976a, p. 64) mentionne quelques-unes de ces fonctions mentales normales. D'abord, il nomme les capacités perceptuelles, l'intelligence et les capacités mnémoniques. Ces fonctions donnent de l'information sur le monde et guident l'action (Boorse, 1976a, p. 64). Boorse mentionne aussi des fonctions comme l'anxiété et la

douleur (Boorse, 1976a, p. 64). Celles-ci servent, entre autres, à signaler le danger. Enfin, Boorse parle des capacités langagières qui favorisent la coopération entre les individus et contribuent à leur enrichissement cognitif (Boorse, 1976a, p. 64). Selon la TBS, toutes ces fonctions mentales contribuent (et ont contribué dans le passé comme je le soutiens dans les prochains paragraphes) typiquement à la survie et à la reproduction de l'organisme.

Étant donné que la TBS stipule que les fonctions normales des systèmes vivants sont une contribution causale *typique* des parties de ces derniers à leur survie et à leur succès reproducteur [fitness], on peut affirmer que la théorie de l'évolution est une des bases théoriques de la TBS³⁵. On peut soutenir cela parce que, dans le cadre de cette dernière, identifier ce qui est une contribution causale statistiquement typique d'une partie d'un système vivant à sa survie et à sa reproduction nécessite que l'on s'intéresse à l'histoire évolutive des ancêtres de ce système qui possède cette partie. En effet, on ne peut pas établir ou cerner la typicalité d'un trait en se basant seulement sur une occurrence isolée de son activité à un moment donné. Pour accomplir cette tâche, il faut aussi étudier

-

Valles, dans un écrit plus récent, semble encore penser que la théorie de l'évolution est sous-jacente à la TBS. Commentant cette dernière, en effet, il écrit :

L'idée est que nous pouvons utiliser les connaissances de la biochimie, de la pathologie et de la biologie évolutive pour obtenir un moyen objectif de « lire » la nature [...][qui nous permet de déterminer] quels sont les états pathologiques ou maladifs, [et ce] sans l'interférence d'évaluations embrouillées et enracinées/biaisées culturellement. (Valles, 2020, section 2.2, mes italiques)

³⁵ Valles (2012, p. 256-258) aussi pense que la théorie de l'évolution par la voie de la sélection naturelle est sous-jacente à la TBS. En effet, il dit que « [l]a TBS est liée à la médecine évolutionniste, car Boorse définit la "fonction normale d'une partie ou d'un processus" comme "une contribution statistiquement typique de celle-ci à la survie et à la reproduction de l'individu" » (Boorse, 1997, p. 7 dans ; Valles, 2012, p. 256-257).

comment le trait opère depuis très longtemps dans la classe de référence. C'est ce qu'affirme Garson alors qu'il examine la TBS :

[...] lorsque nous considérons ce qui est « statistiquement typique » pour un trait, nous ne pouvons pas tenir compte seulement de ce qui est typique en ce moment. Nous devons examiner le comportement du trait sur une tranche de temps qui inclut le moment présent et remonte loin dans le passé [...] (Garson, 2019a, p. 1148-1149)

Comme le remarque Garson (2019a, p. 1149), Boorse lui-même admet cela. De fait, ce dernier affirme que « [d]e toute évidence, une tranche de l'histoire de l'espèce doit être incluse dans ce qui est *typique* de l'espèce » (Boorse, 2002, p. 99, mes *italiques*). Boorse ajoute à cela que cette tranche de temps dans l'histoire évolutive devrait être plus longue qu'« une vie ou deux » et pourrait en fait inclure « des millénaires » (Boorse, 2002, p. 99). Ainsi, il semble que le concept de fonction normale que Boorse qualifie d'*an*historique présente au fond, comme le mentionne Garson (2019a, p. 1148), une « profondeur historique » [*historical depth*]. Considérant cela, on peut se demander ce qui pousse Boorse à affirmer que sa théorie de la fonction, *quand elle est appliquée aux systèmes vivants*³⁶, est « logiquement indépendante de toute [référence à l']évolution [...] » (Boorse, 2002, p. 75).

À cela, et à l'encontre de ce que dit Boorse (1977, p. 549, 1997, p. 88, 2002, p. 73-75) à propos de la TBS, on peut ajouter que cette dernière et sa conception de la fonction

_

³⁶ Je spécifie cela parce que, lorsqu'elle est utilisée dans d'autres domaines (par exemple, dans ceux des artéfacts ou des systèmes non vivants faits par les humains), la conception de la fonction de Boorse *ne suppose pas* la théorie de l'évolution par la voie de la sélection naturelle ni, en conséquence, que les buts des fonctions sont la survie et la reproduction (Boorse, 1976b, p. 79). Par exemple, si on applique sa conception de la fonction à une chaise, on dira alors que les parties de la chaise contribuent à la fonction utilitaire que les humains lui donnent, à savoir celle d'être un artéfact sur lequel on peut s'assoir (Boorse, 1976b, p. 79-80). Par ailleurs, si on applique sa conception à un missile à tête chercheuse, on dira alors que les parties du missile contribuent au guidage du missile vers sa cible (lequel est un but inhérent à ce système).

normale peuvent être interprétées avec une grille d'analyse adaptationniste (voir Cournoyea, 2013 p. 49 pour un avis semblable; voir aussi Thompson et Upshur, 2017, p. 20-22). L'adaptationnisme est une notion qui peut être comprise de plusieurs manières (Godfrey-Smith, 2001; Lewens, 2009). La façon dont Piggliucci et Kaplan résument les trois interprétations de l'adaptationnisme que formule Godfrey-Smith (2001, p. 336-338) est éclairante :

- (1) *Empirique*: La sélection naturelle est si puissante, et les contraintes sur elle si peu nombreuses, que pratiquement chaque trait sera adapté de manière optimale pour une fonction (ou [sera] un compromis optimal pour un ensemble de fonctions).
- (2) *Explicative*: L'adaptation par la sélection naturelle, aussi omniprésente (ou non) soit-elle, est la réponse aux questions les plus importantes et les plus intéressantes de la biologie évolutive.
- (3) *Méthodologique*: La meilleure approche pour comprendre l'origine historique, la propagation ou le maintien d'un trait est de supposer, d'abord, que ce dernier est une adaptation, puis de travailler à partir de cette hypothèse et aller vers des hypothèses alternatives. (Pigliucci et Kaplan, 2006, p. 121)

On peut penser que la TBS suppose au moins la première et la deuxième de ces trois thèses suivant cet argument :

Prémisse 1: Une fonction normale est une contribution causale typique à la survie et à la reproduction.

Prémisse 2 : Le design fonctionnel d'une classe de référence est le répertoire de fonctions normales que partagent les membres de cette classe de référence.

Prémisse 3 : Le design fonctionnel d'une classe de référence est déterminé par la sélection naturelle. C'est ce que dit Boorse :

[L]e mode de fonctionnement typique de notre espèce, typique par définition, nous a dotés, *par la sélection naturelle*, des capacités adaptées à une manière de vivre dans notre environnement, que nous valorisons. (Boorse, 1977, p. 550, mes *italiques*. Traduction d'Élodie Giroux dans Giroux et Lemoine, 2012, p. 76).

Dans le même ordre d'idées, Boorse affirme aussi ailleurs qu'

en dehors des échelles de temps évolutionnaires, les *designs* biologiques manifestent une grande stabilité qui est vigoureusement maintenue grâce à l'action de la *sélection normalisante* [normalizing selection] (Boorse, 1977, p. 557, 1997, p. 32, mes *italiques*. Traduction d'Élodie Giroux dans; Giroux et Lemoine, 2012, p. 89-90)³⁷

Corolaire 1 : les fonctions normales sont conçues, dans le cadre de la TBS, comme des adaptations retenues par la sélection naturelle.

En effet, si le design fonctionnel d'une classe de référence est déterminé par la sélection naturelle (P3) et si *le design fonctionnel* d'une classe de référence *est le répertoire de fonctions normales* que partagent les membres de cette classe de référence (P2), alors chacune des fonctions normales est retenue par la sélection naturelle (C1).

Conclusion: la TBS suppose au moins l'adaptationnisme empirique et l'adaptationnisme explicatif.

Autrement dit, d'une part, la TBS suppose que la sélection naturelle est le mécanisme *principal* de l'évolution opérant dans le monde du vivant. D'autre part, la TBS *explique*, dans la plupart des cas, la présence des fonctions normales partagées par les membres de la classe de référence (c.-à-d. son design fonctionnel) avec la sélection naturelle.

³⁷ Ce que dit Boorse ici s'harmonise bien avec cette affirmation de Thompson et Upshur qui porte sur le concept de fonction normale de la TBS: «L'utilisation par Boorse de [...][l'évolution] pour opérationnaliser la "fonction normale" semble exiger une conception statique (une tranche temporelle) de l'évolution. La manière dont les choses ont *en ce moment* évolué pour fonctionner [*currently evolved to function*] pour améliorer la survie est [pour Boorse] la norme. » (Thompson et Upshur, 2017, p. 22, mes *italiques*). On peut penser que c'est cette conception statique de l'évolution qui permet d'avancer l'idée (que j'ai remis en question précédemment) selon laquelle les fonctions normales contribuent seulement *au moment présent* à la survie et à la reproduction de l'individu. Il est également plausible de soutenir que c'est dans le cadre de cette conception statique de l'évolution que Boorse conçoit les notions de classe de référence, le design fonctionnel de même que les concepts de santé et de pathologie que l'on verra bientôt.

Je pense que la conception boorséenne de la fonction normale (physique ou mentale) telle qu'elle est définie dans le cadre de cette interprétation adaptationniste de la TBS est en arrière-plan des propos de trois des représentants des RDoC. En effet, ces derniers affirment que

les domaines et les construits [de la matrice des RDoC] sont destinés à refléter les principaux systèmes que le cerveau a développés pour [that the brain has evolved in order for] que les humains se comportent de manière adaptative. (Morris et al., 2016, p. 228)

Certains critiques des RDoC semblent aussi penser que cette conception de la fonction normale (mentale) est opérante dans ce cadre de recherche :

[l]es RDoC conçoivent les troubles mentaux comme des dysfonctionnements des systèmes cérébraux [(listés dans la matrice)] qui ont d'importantes implications adaptatives, tels que les systèmes liés à la réactivité aux récompenses et à la sensibilité aux menaces. (Lilienfeld et Treadway, 2016, p. 445)³⁸

Maintenant que ces concepts de la TBS ont été présentés et mis en lien avec les propos des tenants des RDoC et avec ceux des critiques de ces derniers, on peut définir la *santé* et la *dysfonction* (ou la *pathologie*). Il convient d'abord de définir la santé. Suivant la TBS, on dira qu'un système vivant est en *santé* si toutes ses parties contribuent typiquement (c'est-à-dire de la manière dont elles contribuent habituellement chez les membres de la classe de référence à laquelle il appartient) à sa survie et à sa reproduction. Selon Kingma (2010, p. 245-249) — et Boorse (2014, p. 685) est

-

³⁸ Dans cet article de Lilienfeld et Treadway, les fonctions adaptatives que présentent ces systèmes cérébraux sont conçues, selon mon analyse, comme étant apparues sous l'action de la sélection naturelle au cours de l'histoire évolutive de l'espèce humaine. Cela est encore plus clair à la lecture d'un article d'Harkness, Reynolds et Lilienfeld (2014) auquel ce même Lilienfeld et Treadway (2016, p. 445) renvoient dans leur propre article.

d'accord avec elle sur ce point — il faut ajouter trois précisions à cette définition de la santé :

- (1) Le niveau de ce fonctionnement normal d'une partie n'est pas seulement qualitativement constaté, mais aussi déterminé de manière quantitative (Kingma, 2010, p. 245 et 249). Par exemple, un cœur fonctionne normalement non seulement parce qu'il pompe le sang (c'est ce que Kingma appelle sa fonction normale qualitative), mais aussi (entre autres) parce qu'il le pompe à un certain rythme, à une certaine vitesse et selon une certaine pression artérielle (c'est ce que Kingma appelle sa fonction normale quantitative).
- (2) Une fonction d'une partie ne peut être jugée comme opérant normalement (ou anormalement) que relativement à une situation spécifique ou à un environnement spécifique (Boorse, 1977, p. 561-562; Kingma, 2010, p. 249). Pour illustrer cela, Kingma (2010, p. 248) donne l'exemple des plaquettes sanguines qui sont actives seulement lorsque nos vaisseaux sanguins sont endommagés. Quand nos vaisseaux sanguins ne sont pas endommagés, en effet, les plaquettes sanguines flottent dans le sang et ne font rien (et cela est important parce que si elles étaient actives elles causeraient des caillots de sang pouvant engendrer des thromboses). Kingma (2010, p. 248) donne aussi le cas de la régulation de la glycémie. En effet, les niveaux normaux quantitativement mesurables de production d'insuline par le pancréas, d'absorption du glucose par les cellules et de synthèse du glycogène par le foie varient tous selon des situations spécifiques. Par exemple, une production importante d'insuline est normale au moment où un individu mange un gros gâteau au chocolat, mais elle est anormale s'il est à jeun depuis 12 heures. Enfin, toujours pour illustrer cela, Kingma affirme qu'« une fréquence cardiaque de 60 battements par minute est

normale pendant le sommeil, mais anormale pendant un exercice » (Kingma, 2010, p. 249).

(3) Une partie opère normalement si elle est *disposée* à opérer normalement dans des situations spécifiques potentielles *non présentes* (Boorse, 1977, p. 561-562; Kingma, 2010, p. 246-247 et 249). Par exemple, mon système immunitaire, qui est au repos en ce moment, fonctionne normalement si, dans l'éventualité où mon corps développait une infection, il est *disposé* à se mettre en action. Un autre exemple serait que mon cœur, qui est au repos en ce moment, et qui bat à 55 battements par minute, fonctionne normalement si, dans l'éventualité où j'avais à me mettre à courir, il est *disposé* à battre à un rythme plus rapide (Kingma, 2010, p. 249).

Maintenant que la conception de la santé inhérente à la TBS a été définie, on peut présenter la notion de dysfonction (ou de pathologie) que contient cette dernière. Selon la TBS, on dira qu'une partie d'un système vivant présente une *dysfonction* (ou une *pathologie* ou un *trouble*) quand elle ne fait pas ou a de la difficulté à faire ou ne serait pas disposée à faire ce qu'elle effectue *typiquement* pour contribuer à la survie et à la reproduction individuelle du système vivant dans une situation spécifique. Autrement dit, les dysfonctions à l'origine d'états pathologiques présentent deux caractéristiques : (1) elles nuisent (ou pourraient nuire), dans une situation spécifique, à la capacité de survie et de reproduction du système vivant et (2) leur niveau de fonctionnement, comparativement à celui des fonctions des autres membres de la classe de référence, se situe aux deux ou à l'une des extrémités de la courbe normale suivante ³⁹:

³⁹ Voici une citation de Boorse qui conforte le point (2) : « [...] la caractéristique logique la plus évidente de la normalité médicale est que la plupart des fonctions ont une étendue de valeurs normales. Aucune

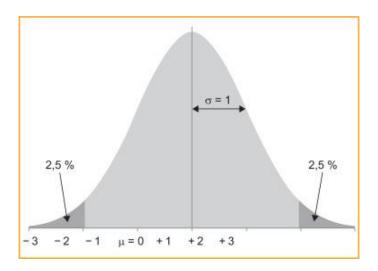


Figure 1.4 Loi normale (Ingrand, 2017)

C'est cette conception de la dysfonction (ou de la pathologie) qui est opérante dans la conception du trouble mental des RDoC. Je pense que l'on peut inférer cela, entre autres, des trois citations suivantes des représentants des RDoC :

- (1) « Les RDoC posent les questions suivantes : quelle est la distribution normale pour un certain trait [adaptatif] ou une certaine caractéristique [...]? » (Cuthbert, 2014b, p. 31).
- (2) «[...] les troubles mentaux sont considérés comme des extrêmes à l'une ou l'autre ou aux deux extrémités des distributions normales [...] » (Cuthbert et Insel, 2010, p. 312)

valeur de fréquence cardiaque, de pression artérielle, de taux sanguin d'azote uréique, de [taux de] sérum glutamo-oxaloacétique transaminase, de force de l'avant-bras, de taille, de QI et ainsi de suite n'est uniquement normale. Il existe plutôt une étendue de variations normales autour d'une moyenne, avec une ou deux extrémités pathologiques. » (Boorse, 2002, p. 101, mes italiques)

(3) « Il est peu probable que l'on se plaigne d'une mémoire exceptionnelle ou d'une vision précise [(Cuthbert et Insel affirment cela alors qu'ils parlent de ce qu'est un trouble mental)]» (Cuthbert et Insel, 2013, p. 5).

La première de ces citations montre que les tenants des RDoC pensent qu'il faut distribuer nos mesures quantitatives du niveau d'activité ou de performance des fonctions normales des individus appartenant à une classe de référence sur une courbe de Gauss. La deuxième citation montre que les tenants des RDoC conçoivent les dysfonctions des parties des systèmes vivants comme des déviances statistiques que l'on peut situer à une ou aux deux extrémités de la courbe normale. La troisième citation montre que les tenants des RDoC n'identifient pas les dysfonctions essentiellement à des déviances statistiques (par exemple, présenter une intelligence hors du commun est statistiquement rare, mais cela n'est pas une pathologie). En effet, suivant la conception de la dysfonction boorséenne qu'ils adoptent, ils les identifient à des déviances statistiques qui nuisent à la survie et à la reproduction individuelle. Dans cette perspective, le diabète de type 1 que présente un adolescent de 15 ans est une pathologie parce que cette condition est statistiquement rare et parce qu'elle nuit à la survie (si celui-ci contrôle mal sa glycémie, par exemple) et, éventuellement, à la reproduction de ce dernier.

Pour compléter ces propos sur la santé et la pathologie (ou la dysfonction), il est important de souligner que, selon la TBS, « la distinction entre normal et pathologique est *naturelle* et objective au sens où elle peut être observée dans la nature et n'est pas le fruit d'une décision » (Giroux, 2012, p. 55). On peut affirmer que les tenants des RDoC adoptent cette thèse que l'on peut qualifier de naturaliste en regard de la santé et des pathologies parce qu'ils pensent que les troubles mentaux sont des *espèces naturelles*. Robert Klee donne une bonne définition générale de ce concept :

[...] une espèce naturelle est une classe d'entités dont chaque membre partage les mêmes propriétés essentielles, et où la formation et l'existence continues de la classe dépendent de processus naturels plutôt que de pratiques ou de conventions sociales humaines. (Klee, 1997, p. 246)

Parce qu'ils croient que les troubles mentaux sont des espèces naturelles, les représentants des RDoC soutiennent que les catégories et, plus précisément, les modèles mécanistes des troubles mentaux (section 1.5.2) que l'on développe doivent saisir ces processus ou mécanismes naturels (qui sont dysfonctionnels au sens de la TBS).

Cette thèse à propos de ce que sont les troubles mentaux et cet objectif épistémique qu'entretiennent les tenants des RDoC sont mentionnés dans ce passage où trois de ces derniers parlent des travaux de Paul Meelh en psychopathologie de la personnalité (ici spécialement ceux portant sur la schizotypie):

L'articulation du concept de taxon de schizotypie [faite] par Meehl esquisse les objectifs des *Research Domain Criteria* (RDoC) du NIMH: « La façon la plus courante d'expliquer le concept de taxon est peut-être de dire qu'un véritable taxon est une espèce naturelle, par opposition à une classe arbitraire. La connotation d'"espèce naturelle" [...] [signifie] [...] [que le taxon de schizotypie] existe en quelque sorte réellement, que les scientifiques humains l'identifient ou non » (Meehl, 1992, p. 122). Des travaux ultérieurs ont apporté la preuve de la validité du taxon de schizotypie en utilisant les outils de l'imagerie cérébrale moderne, de la biologie moléculaire et de la génétique (Faraone *et al.*, 2013; Tsuang *et al.*, 2003). (Morris *et al.*, 2016, p. 227)

Des critiques des RDoC, comme Sisti et ses collègues, soutiennent sensiblement la même chose. De fait, ils disent que « [l]es RDoC supposent que les troubles mentaux sont des espèces naturelles qui peuvent être découvertes et expliquées selon un mélange

axiologiquement neutre de génétique, d'imagerie et de neurosciences » (Sisti *et al.*, 2013, p. 2; voir aussi Zachar, 2014a, p. 145 pour un avis similaire)⁴⁰.

1.5.4 Le principe épistémoquantitatif (et la variante quantitative du réalisme des entités basé sur la robustesse d'Eronen)

Je pense que les défenseurs des RDoC pourraient faire leurs ces considérations épistémologiques que tenait Lord Kelvin au 19^e siècle :

Quand vous pouvez mesurer ce dont vous parlez et l'exprimer en nombre, vous en savez quelque chose; mais quand vous ne pouvez pas le mesurer, quand vous ne pouvez pas l'exprimer en nombre, votre connaissance est pauvre et insatisfaisante : c'est peut-être le début de la connaissance, mais vous avez à peine, dans vos pensées, avancé l'état de la science. (Kelvin cité dans Tal, 2017, note 1)

Les partisans des RDoC pourraient revendiquer les affirmations de ce célèbre physicien anglais parce qu'à mon avis ils idéalisent, comme lui, l'approche quantitative en science. En effet, ils élèvent cette approche en un principe devant guider la recherche en psychiatrie. On pourrait formuler ce dernier de la manière suivante : seulement ce

les jeunes souffrant du trouble déficitaire de l'attention avec hyperactivité (TDAH) présentent un retard de maturation cérébrale [de] trois ans en moyenne dans certaines régions. [...] et que [c]e retard est le plus marqué dans le lobe frontal, [lequel est] important pour la capacité à contrôler la pensée, l'attention et la planification. (NIMH, 2008, p. 14)

_

montré que

⁴⁰ Pour finir, j'aimerais mentionner, dans cette note de bas de page, que si je n'ai pas trouvé de citation montrant *explicitement* que les partisans des RDoC utilisent le concept de *classe de référence* de la TBS, je pense néanmoins que cette notion est tacitement mobilisée quand ils font de la recherche sur les aspects développementaux des troubles mentaux (section 1.3, principe 5). En effet, faire ce type de recherches nécessite que l'on comprenne la manière dont (entre autres) le cerveau se développe chez les humains au cours de leurs différents stades de développement (enfance, préadolescence, adolescence, jeunes adultes, adultes, etc.). Les chercheuses et les chercheurs ont, en effet, besoin d'étudier ces stades développementaux afin d'être en mesure de comparer le fonctionnement et la structure du cerveau des membres qui composent une classe de référence et, par là même, de juger lesquels d'entre eux présentent un développement cérébral et une structure cérébrale typique ou atypique. Par exemple, sur la base de ce qu'ils considèrent comme un développement normal du cerveau, Shaw et ses collègues (2007) ont

qui peut être mesuré quantitativement a une pertinence épistémique (voir Fuchs, 2018, p. 254 pour une idée semblable). Je qualifie ce principe d'« épistémoquantitatif ». Dans ce qui suit, d'abord, j'explique pourquoi je crois que ce principe est sous-jacent aux RDoC. Je présente ensuite la conception du progrès en science psychiatrique que me semble sous-tendre ce dernier. Je montre après que cette conception de l'avancement des connaissances psychiatriques suppose le réalisme de Bunge (2006, p. 29-30) et, plus spécifiquement, ce qui me semble être une version quantitative du réalisme des entités basé sur la robustesse [robustness-based entity realism] (RER) qu'a développé le philosophe des sciences Markus Eronen (2015, 2019a, 2019b)⁴¹. Enfin, j'explique en quoi ce principe épistémoquantitatif et cette variante quantitative du RER influent sur la conception du trouble mental inhérente aux RDoC et, en conséquence, sur la recherche psychiatrique financée par le NIMH.

Un fait qui confirme que le principe épistémoquantitatif est sous-jacent aux RDoC est que les construits de la matrice sont sans exception tous quantitativement mesurables (voir section 1.3, principe 2). Ce fait concorde avec les convictions qu'entretiennent les partisans des RDoC en regard de l'amélioration des connaissances des troubles mentaux qui nécessite, selon eux, de mesurer ces construits. De fait, ils affirment que « [p]our comprendre les troubles mentaux [...] il sera important [...] [d]'élaborer

Dans ce qui suit, puisque ce n'est pas utile à mon propos, je n'en dis pas plus sur le RE.

⁴¹ Le RER formulé par Eronen est une variante du réalisme des entités [*entity realism*] (RE) originellement développé indépendamment par Nancy Cartwright (1983) et Ian Hacking (1983). Le RE qu'ont élaboré cette dernière et ce dernier

^[...] est basée sur les pratiques scientifiques [...] de manipulation expérimentale et d'explication causale qui sont utilisées aujourd'hui. [L'idée] centrale du ER est que, même si nous ne sommes pas justifiés de croire en la vérité des théories scientifiques, nous sommes justifiés de croire que les entités que les scientifiques manipulent causalement sont réelles. (Eronen, 2019b, p. 2342)

des mesures fiables et valides de ces composantes ⁴² fondamentales des troubles mentaux [...]. » (NIMH, 2008, p. 10)

Dans le même ordre d'idées, ils disent que

[...] l'accent mis sur des construits de [fonctions] [...] qui peuvent être mesurés par divers moyens et selon les différentes unités des RDoC [...] vise à jeter les bases d'une approche psychométrique et quantitative rigoureuse de l'évaluation [de la performance de ces fonctions]. (Lee Anna Clark et al., 2017, p. 118, mes italiques)

À la lecture de ces propos, il est plausible de penser que le principe épistémoquantitatif suppose que le progrès en science psychiatrique a lieu par le biais d'un processus de validation, par l'intermédiaire de mesures *quantitatives*, des construits des RDoC (Kozak et Cuthbert, 2016, p. 292). Murphy (2015, p. 67), s'inspirant de Zachar (2012, p. 22), soutient que cette validation par voie de quantification se fait, dans le cadre des RDoC, par une *validation petit-v* et par une *validation grand-*V.

D'abord, il convient d'aborder la validation petit-v. Cette dernière s'opère en produisant, à l'aide de techniques diverses (ex. : imagerie par résonnance magnétique fonctionnelle [IRMf], test psychométrique, etc.), pour chacun des construits des RDoC, des données quantitatives (provenant des différentes unités d'analyses) qui sont mises en relation les unes avec les autres. Autrement dit, on fait de la validation petit-v (1) lorsqu'on mesure des variables observables que l'on pense être des indicateurs du construit des RDoC auquel on s'intéresse et (2) lorsqu'on essaie de voir la manière dont ces variables observables covarient entre elles. Par exemple, on peut penser que des variables observables provenant des unités d'analyse « autoévaluation » (ex. : mesure

⁴² Le terme « composantes » réfère ici aux construits des RDoC.

de la détresse expérientielle en relation avec des objets et des événements aversifs), « comportement » (ex. : mesure des signes d'inconfort lors d'interactions avec des personnes non familières) et « physiologie » (ex. : augmentation du réflexe de sursaut lors du traitement [cognitif] d'indices désagréables ou neutres) covarient entre elles et sont reliées au construit « peur dispositionnelle » [dispositional fear] (Patrick et al., 2013, p. 904).

En principe, le niveau de validité (petit-v) d'un construit — par exemple celui de la peur dispositionnelle — sera (entre autres)⁴³ fonction de l'invariance des covariations existant entre les différentes mesures des variables observables qui pointent vers lui⁴⁴.

Le projet de recherche de Sukhodolsky, vu à la section 1.4.2, me servira pour illustrer plus en détail le concept de validation petit-v et cette idée d'invariance des relations

⁴³ Je mets « entre autres » entre parenthèses parce que la validation d'un construit doit se faire au moins d'une autre manière dont je ne parle pas. En effet, si on valide un construit en montrant qu'il forme un réseau nomologique avec des *variables observables* qui pointent vers lui (Kozak et Cuthbert, 2016, p. 292), il faut aussi valider ce même construit en examinant la manière dont il covarie avec les *autres construits* — reliés eux aussi pour la plupart à des variables observables — de son réseau nomologique (Cronbach et Meehl, 1955, p. 290-291). C'est de cette manière que l'on pourra découvrir graduellement ce que signifie ce construit et le valider.

⁴⁴ À noter, toutefois, que l'obtention de covariations parfaites entre les mesures est un idéal à atteindre. En effet, il semble rare, pour l'instant, qu'on arrive à établir de telles covariations. Par exemple, les mesures d'autoévaluation, de comportements et de physiologie covarient seulement modestement dans le construit de peur (Kozak et Cuthbert, 2016, p. 292). Ce type de situation montre que le travail des chercheuses et des chercheurs n'est pas seulement de prendre des mesures des différentes dimensions du construit et de les faire covarier entre elles, mais aussi d'élaborer ce construit en réévaluant les données qui servent à le valider et en réfléchissant à ce dernier à la lumière de différentes théories pour mieux en comprendre (entre autres) sa dynamique interne. Par exemple, il est pertinent de noter que certaines mesures du construit « peur » sont plus corrélées entre elles si le niveau de peur est élevé (Kozak et Cuthbert, 2016, p. 292). Il faut aussi avoir en tête que le niveau de peur est fonction de la peur perçue et non de la proximité physique de ce qui cause la peur (Kozak et Cuthbert, 2016, p. 292). Enfin, il faut savoir que d'autres facteurs peuvent jouer sur les mesures comme le niveau de fatigue des individus de l'échantillon étudié (Kozak et Cuthbert, 2016, p. 292). C'est donc en réévaluant constamment les données et les théories qui servent à concevoir les différentes facettes d'un construit et la manière dont elles sont reliées entre elles que l'on peut espérer parvenir à développer un construit valide. Et il se peut fort bien que le construit que l'on voulait valider au départ ne soit plus celui qu'on validera à la fin de cette démarche de raffinement et de développement conceptuel.

entre les mesures des variables observables qui pointent vers un construit théorique des RDoC. Comme il a été abordé plus tôt, lors d'une étude portant sur des jeunes âgés de 8 à 16 ans présentant des comportements perturbateurs, les membres de l'équipe de ce chercheur veulent démontrer la validité du construit « non-récompense frustrante ». Pour ce faire, ils prendront des mesures de l'activité cérébrale de ces enfants (à l'aide d'un électroencéphalogramme et de techniques d'imagerie cérébrale) alors qu'ils feront le test cognitif Go/noGo. Suivant la littérature scientifique sur les personnes ayant des problèmes de régulation des émotions, ils s'attendent à ce que leur amygdale soit hyperactive et que leurs régions préfrontales soient hypoactives (Sukhodolsky et al., 2016, p. 39). Si ces enfants réussissent une thérapie cognitivocomportementale (TCC), toujours selon ce que l'on peut lire dans ces écrits, ils prévoient que l'activité de leur amygdale baissera et que l'activité de leurs régions préfrontales augmentera (Sukhodolsky et al., 2016, p. 39). Si on suppose que leurs prédictions sont bonnes, on pourrait alors dire qu'ils auraient produit des données empiriques contribuant à la validation (petit-v) du construit des RDoC « non-récompense frustrante ». On pourrait affirmer cela parce qu'ils auraient montré qu'il existe une relation constante entre deux variables observables qui pointent vers ce construit, à savoir entre (1) des mesures de l'activité des circuits cérébraux responsables de la régulation des émotions et (2) des mesures de comportements agressifs ou normaux. Autrement dit, ils auraient montré que plus ces circuits fonctionnent normalement, moins on compte de comportements agressifs et vice versa⁴⁵.

⁴⁵ À noter que trouver ce genre de relations invariantes ou de généralisations invariantes (Woodward, 2003) permet de faire des prédictions ou des interventions (voir Goyer, 2013, section 3.8, pour une présentation des idées de Woodward 2003 à ce sujet). On pourrait prédire, entre autres, que toutes formes d'interventions — disons pharmacologiques cette fois-ci — visant à rétablir l'activité des circuits neuraux responsables de la régulation des émotions devraient faire baisser la quantité de comportements agressifs chez ces enfants.

Maintenant que la validation petit-v a été vue, on peut présenter la validation grand-V. Celle-ci peut être définie comme une sorte d'idéal régulateur épistémique selon lequel « la compréhension du monde implique de saisir sa structure causale » [...] « en mesurant tout ce qui est accessible » (Murphy, 2015, p. 62, mes italiques). C'est par l'intermédiaire de beaucoup de validations quantitatives petit-v que les partisans des RDoC tentent de saisir cette structure causale. Dans cette perspective, plus on valide un construit de la matrice des RDoC de diverses manières, plus on peut prétendre avoir découvert les circuits neuraux auxquels il correspond (et, éventuellement, les gènes responsables du développement de ces circuits).

En dernière instance, l'objectif des partisans des RDoC est de découvrir, par l'intermédiaire de plusieurs validations petit-v, la structure causale *complète* des différents troubles mentaux qui sont, selon eux, tous des espèces naturelles (section 1.5.3). Cette structure causale des différents troubles mentaux sera saisie par divers modèles mécanistes (section 1.5.2). Chacun de ces modèles expliquera le syndrome (autrement dit, l'ensemble des symptômes) du trouble mental à l'étude, lequel est un *phénomène* produit par un mécanisme dysfonctionnel⁴⁶ (section 1.5.2).

La reformulation faite par Dirner et ses collègues (2011) du modèle cognitif de la dépression majeure de Beck (1967, 1987) en un modèle neurobiocognitif qui montre les structures et les processus cérébraux qui sont sous-jacents aux processus cognitifs (hypothétiquement) impliqués dans l'émergence des symptômes de cette maladie mentale me semble être un bon exemple de ce qui pourrait être l'ébauche d'un tel

⁴⁶ Je crois que cette manière mécaniste de modéliser ou d'expliquer les troubles mentaux est compatible avec ou, à tout le moins, complémentaire à l'approche des réseaux de construits [Construct-Network Approach] proposée par Patrick et ses collaborateurs (2013, p. 905-906). Cette dernière vise ultimement à établir une correspondance entre des mesures de différents aspects cliniques d'une condition psychiatrique quelconque et des mesures de l'activité neurale sous-jacente à cette condition (Patrick et al., 2013, p. 903).

modèle mécaniste saisissant la structure causale (ou l'espèce naturelle) de la dépression majeure⁴⁷ (voir Figure 1.5).

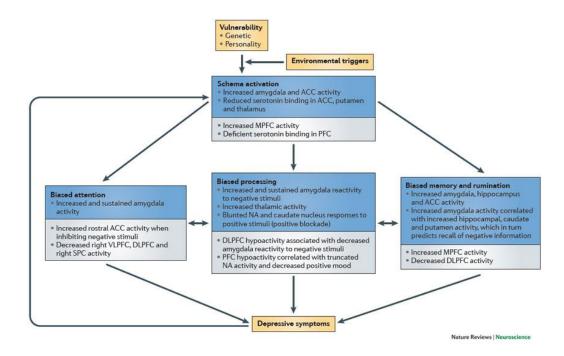


Figure 1.5 Modèle neurobiocognitif de la dépression majeure basé sur le modèle cognitif de Beck (Disner *et al.*, 2011, p. 475)

Je pense que la thèse inhérente à la validation grand-V selon laquelle il existe une structure causale objective du monde, que l'on peut connaître grâce à nos instruments

⁴⁷ Pour mieux comprendre le niveau cognitif de ce modèle neurobiocognitif de la dépression majeure, il faut avoir en tête ce que dit le modèle cognitif de Beck (1967, 1987). Ce modèle suppose que les personnes humaines développent, au cours de leur existence, des schèmes cognitifs, lesquels sont des représentations de stimuli, d'idées ou d'expériences stockées dans leur esprit (Disner *et al.*, 2011, p. 467). Chez les personnes dépressives ou vulnérables à la dépression majeure (pour des raisons génétiques, par exemple), certains de ces schèmes sont défectueux (en raison de traumas durant l'enfance, par exemple). Quand les stimuli environnementaux intérieurs ou extérieurs sont captés par l'individu, ces schèmes cognitifs défectueux sont activés et causent des biais attentionnels, des biais sur le plan du traitement de l'information et des biais sur le plan de l'engrangement et de la récupération des souvenirs. Le modèle de Beck dit que ce sont ces biais causés par les schèmes cognitifs défectueux qui maintiennent et qui causent les symptômes de la dépression majeure.

de mesure quantitative qui permettent plusieurs validations petit-v, est saisie par le système philosophique du réalisme formulé par Bunge (2006, p. 29-30). En effet, si on entretient cette conviction, alors on accepte (1) la thèse réaliste ontologique selon laquelle le monde extérieur existe indépendamment du sujet connaissant et (2) la thèse réaliste épistémique selon laquelle le monde peut être connu⁴⁸.

Bien que le système philosophique du réalisme formulé par Bunge saisisse cette thèse inhérente à la validation grand-V que soutiennent les tenants des RDoC, je crois qu'il est plus précis de dire que ces derniers adoptent une *version quantitative* du RER (Eronen, 2015, 2019a, 2019b). En effet, cette variante du RER, en plus d'être en accord avec les deux thèses réalistes de Bunge, saisit très bien le principe épistémoquantitatif, la conception du progrès scientifique sous-jacente à ce dernier et, ultimement, l'engagement ontologique réaliste inhérent au principe régulateur qu'est la validation grand-V qu'adoptent, à mon avis, dans leur pratique scientifique, les chercheuses et les

⁴⁸ Ces deux thèses du réalisme sont directement tirées des deux premières thèses de la conception du réalisme de Bunge (2006, p. 29-30), laquelle en contient sept (pour mon propos, je n'ai pas besoin de présenter ces sept thèses) :

⁽¹⁾ Réalisme ontologique : le monde extérieur existe indépendamment du sujet connaissant.

⁽²⁾ Réalisme épistémologique :

⁽a) Le monde peut être connu.

⁽b) Toute connaissance des faits est incomplète et faillible, et en grande partie indirecte.

⁽³⁾ Réalisme sémantique :

⁽a) Quelques propositions réfèrent à (ou sont à propos) des faits;

⁽b) quelques-unes de ces propositions (factuelles) sont approximativement vraies;

⁽c) en principe toutes les approximations sont perfectibles.

⁽⁴⁾ Réalisme méthodologique : La meilleure stratégie pour explorer le monde est la méthode scientifique (scientisme).

⁽⁵⁾ Réalisme axiologique : Il existe des valeurs objectives, telles que la santé, la connaissance, la sécurité, la paix, la protection de l'environnement et la justice.

⁽⁶⁾ Le réalisme moral : Il existe (a) des faits moraux, comme des actes généreux et égoïstes; et (b) des principes moraux vrais, tels que « Les droits, pour être justes et respectés, doivent être balancés par des devoirs» et «La solidarité et la démocratie favorisent la coexistence».

⁽⁷⁾ Réalisme pratique : Il existe des paires objectives <moyen-but>, tel que <travail-bien-être>, <connaissance-efficacité> et <participation-démocratie>.

chercheurs œuvrant dans le cadre des RDoC. Je ferai connaître les raisons qui m'amènent à penser cela à la fin de cette section. J'expliquerai aussi en quoi le principe épistémoquantitatif et cette variante quantitative du RER qu'il suppose influent sur la conception du trouble mental inhérente aux RDoC et, conséquemment, sur la recherche psychiatrique. Cela dit, pour le moment, je présente le RER.

Une notion centrale du RER est celle de robustesse (un concept à l'origine développé par Wimsatt, 1981). Il faut donc la présenter avant de formuler ce dernier. Eronen conçoit la robustesse de cette manière :

X est robuste dans une communauté scientifique pertinente à un certain moment pour autant que X est détectable, mesurable, dérivable, productible ou explicable de diverses manières indépendantes [les unes des autres]. (Eronen, 2015, p. 3967)

Il est important de préciser cinq choses à propos de cette définition de la robustesse qui est centrale au RER. La première est que le X en question peut référer à des propriétés, des entités, des phénomènes ou à autre chose dans le monde (Eronen, 2015, p. 3967). En psychopathologie, ce X peut être toutes sortes de choses. Par exemple, la fonction des différents neurotransmetteurs (comme la dopamine, la sérotonine ou l'acide γ-aminobutyrique), les processus mnémoniques et le phénomène de la motivation sont des X possibles. Dans le cadre des RDoC, le X est en dernière instance un trouble mental que les conceptrices et concepteurs des RDoC appréhendent comme des espèces naturelles (section 1.5.3) et que les chercheuses et chercheurs œuvrant en son sein essaient de découvrir.

La deuxième chose à souligner est que la robustesse d'une entité ou d'un phénomène est relative à une certaine communauté scientifique évoluant à une certaine époque où sont disponibles des techniques, des méthodes, des modèles et des théories pour intervenir sur le monde et le comprendre. En effet, les façons de détecter, de mesurer,

de dériver, de produire ou d'expliquer les entités et les phénomènes changent avec le temps. Par exemple, il y a plus de données probantes aujourd'hui sur la dépression majeure qu'il y en avait au 17^e siècle. En effet, aujourd'hui, il y a, entre autres, des technologies qui permettent d'observer ce qui se passe dans le cerveau alors qu'il y a 400 ans ces dernières n'existaient pas. Les chercheuses et chercheurs œuvrant dans le cadre des RDoC sont donc mieux équipés pour découvrir le fonctionnement cérébral que les chercheuses et chercheurs qui vivaient il y a quelques siècles.

La troisième chose à souligner dans cette définition de la robustesse concerne la notion « explicable » qui s'y trouve. Eronen affirme qu'il l'inclut dans sa définition de la robustesse pour tenir compte du « fait qu'une entité ou une propriété [qui apparait] dans des lois scientifiques ou des généralisations explicatives est une preuve (faillible) de sa réalité » (Eronen, 2015, p. 3968). Par exemple, en plus d'être capables de détecter, de mesurer et de produire des potentiels d'action, c'est parce qu'on a plusieurs bons modèles indépendants qui les décrivent et les expliquent qu'on a de bonnes raisons de penser qu'ils sont réels (Eronen, 2019b, p. 2346).

La quatrième chose à souligner est qu'aucune des dimensions susmentionnées (détectable, mesurable, dérivable, etc.) n'est en elle-même nécessaire à l'établissement de la robustesse des propriétés, des entités ou des phénomènes (Eronen, 2015, p. 3968). Par exemple, certains phénomènes qu'on ne peut pas produire, mais qu'on peut par ailleurs détecter, mesurer et dériver sont néanmoins robustes. Eronen donne, entre autres, l'exemple des éruptions solaires pour illustrer cela (Eronen, 2015, p. 3968). Celles-ci sont détectables, mesurables et dérivables de diverses manières indépendantes, mais elles sont (et, probablement, seront à jamais) impossibles à produire en contexte expérimental. Cependant, que l'on soit incapable d'une telle tâche n'empêche pas que les éruptions solaires soient des phénomènes astronomiques très robustes.

La cinquième chose à aborder concerne la notion d'indépendance des stratégies de mesure, laquelle est cruciale pour le concept de robustesse (Eronen, 2019b, p. 2346). Bien que la signification de cette notion soit encore débattue, Eronen semble dire qu'elle réfère minimalement au fait que ces « diverses manières » d'étudier le monde dont il est question dans la définition susmentionnée de la robustesse ont pour fondements différents protocoles de recherche, différents processus physiques ou différentes hypothèses théoriques (Eronen, 2015, p. 3969). Par exemple, la coloration de Golgi et le microscope électronique sont deux stratégies de mesures indépendantes servant à détecter les neurones (Eronen, 2019b, p. 2346). En effet, la première consiste à enduire une portion d'un tissu nerveux avec une substance chimique qui cause le noircissement de quelques neurones que l'on peut alors observer avec un microscope régulier. Quant à l'imagerie des neurones que produit le microscope électronique, comme l'explique Eronen, elle

[...] repose sur l'envoi d'un faisceau d'électrons à travers une mince section de tissu nerveux, ce qui entraîne la diffusion d'une partie des électrons en raison de leur interaction avec le spécimen. Ainsi, [ajoute Eronen,] le faisceau qui traverse l'échantillon contient des informations sur la forme et la structure [de ce dernier] lequel peut ensuite être visualisé sous forme de photographie ou sur un écran d'ordinateur. (Eronen, 2019b, p. 2346)

Maintenant que l'on a expliqué la notion de robustesse, voici le RER : « La robustesse confère une justification pour croire que X est réel et le degré de cette justification correspond au degré de robustesse des données probantes pour X. » (Eronen, 2019b, p. 2345)

Le RER dit donc essentiellement trois choses :

(1) Si X est robuste, alors on a une justification de croire que X est réel.

- (2) Plus X est robuste, plus nous avons de raisons ou sommes *justifiés* de soutenir que X est réel.
- (3) Moins X est robuste, moins nous avons de raisons ou sommes *justifiés* de soutenir que X est réel.

Comme on peut le constater, selon le RER, la notion de robustesse *n'est pas* un critère ontologique⁴⁹. Cela signifie qu'elle ne permet pas de dire qu'une propriété, une entité ou un phénomène existe hors de tout doute. La notion de robustesse est plutôt ici comprise comme un critère *épistémique* qui permet de *justifier* un engagement ontologique. En effet, Eronen soutient qu'elle est « un critère qui nous indique quand nous avons de *bonnes raisons* de considérer quelque chose comme réel » (Eronen, 2015, p. 3966)⁵⁰. Parce qu'il y a des degrés plus ou moins importants de robustesse pour un quelconque X (Eronen, 2015, p. 3966), alors il y a des *raisons* plus ou moins importantes qui justifient d'affirmer que ce X existe. Et sous quelles conditions le degré de robustesse de X augmente-t-il? L'augmentation du degré de la robustesse de X varie en fonction du nombre de manières *indépendantes* que l'on est capable d'utiliser pour détecter, mesurer, dériver, produire ou expliquer X. Plus ces manières indépendantes sont nombreuses, plus le degré de robustesse de X est grand et plus il est justifié de croire en son existence. On peut affirmer cela parce qu'« il est très *improbable* que

⁴⁹ Pour exemplifier à quoi pourrait ressembler un tel critère ontologique, Eronen (2015, p. 3973) donne deux exemples. Il en tire un d'un article de Philipp Pettit : « toutes les choses réelles sont composées de particules physiques fondamentales. » (Pettit, 1993) L'autre critère ontologique qu'il donne en exemple est de Jaegwon Kim (qui lui-même l'emprunte au philosophe britannique Samuel Alexander) (2) « être réel est d'avoir des pouvoirs causaux » [to be real is to have causal powers] (Kim, 1993, p. 202).

⁵⁰ En cela, la conception de la robustesse qu'Eronen(2015, p. 3964-3965) soutient est différente de celle de Wimsatt qui en fait un critère ontologique. Pour ce dernier, en effet, « la robustesse possède le bon type de propriétés en tant que critère pour ce qui est réel » (Wimsatt, 1994, p. 214).

toutes ces voies indépendantes se révèlent erronées, et il est donc très *probable* que [ce X] soit réel » (Eronen, 2019a, p. 515, mes *italiques*).

À l'instar d'Eronen (2019b, p. 2352), je pense que le RER rend compte adéquatement de ce qui se passe dans la pratique scientifique. C'est pourquoi je crois que l'engagement ontologique réaliste qu'adoptent les partisans des RDoC et les chercheuses et chercheurs œuvrant dans son cadre est bien saisi par le RER. En effet, l'idée au cœur du RER selon laquelle c'est par l'augmentation de la robustesse de X que l'on peut justifier l'existence de ce X est très similaire à l'idée selon laquelle c'est par l'intermédiaire de plusieurs validations petit-v des construits saisissant un X que l'on atteint la validation grand-V des construits qui saisissent ce X. Cependant, parce que les tenants des RDoC adoptent le principe épistémoquantitatif mentionné au début de cette section, je crois qu'ils adoptent une version exclusivement quantitative de la robustesse et, par conséquent, du RER. Cela implique qu'ils excluent de leurs outils de recherche les méthodes non quantitatives avec lesquelles ils pourraient détecter ou expliquer un X quelconque⁵¹. En conséquence, dans le cadre des RDoC, pour justifier l'existence d'un quelconque X, ne sont utilisées que des méthodes quantitatives. Il suit de cela que les entités, les phénomènes et les propriétés qui intéressent les tenants des RDoC doivent, en dernière instance, être quantifiables. L'effet direct de l'adoption, par les tenants des RDoC, du principe épistémoquantitatif et de la version quantitative du RER que ce dernier suppose, est que les entités ou phénomènes qui ne sont pas quantifiables ne sont pas saisis par la conception du trouble mental des RDoC. Par exemple, les connaissances tirées d'analyses phénoménologiques ou d'études

_

⁵¹ Il est à noter que les cas d'études scientifiques (en neurosciences de la vision, sur les changements climatiques et sur les cellules amacrines dans la rétine) qu'utilise Eronen (2015, 2019b) pour illustrer le RER mettent tous en scène des méthodes *quantitatives*. Cependant, je crois qu'*en principe* le RER (et son concept de robustesse) saisit aussi les phénomènes étudiés par les méthodes *qualitatives*. C'est pourquoi je pense que cette version quantitative du RER que je mets de l'avant a une pertinence.

qualitatives (comme on le verra aux sections 2.2.1 et 2.2.2) ne sont pas prises en compte par cette conception de la pathologie mentale.

1.5.5 L'hypothèse de l'identité psychoneurale de Bunge

Dans les écrits des représentants des RDoC, on retrouve souvent ce type d'affirmations sur la nature des troubles de santé mentale :

- (1) «[...] [L]es maladies mentales *sont* des troubles d'un organe du corps : le cerveau » (Insel *et al.*, 2015, mes *italiques*).
- (2) « Ce qu'il nous faut [...] pour progresser [...] est de reconceptualiser ces troubles [mentaux] *comme* des troubles du cerveau » (Insel, 2013a, mes *italiques*).
- (3) « Les maladies mentales sont présumées *être* des troubles des circuits cérébraux » (Cuthbert et Morris, 2012, p. 33, mes *italiques*).

Dans ces trois citations, les troubles mentaux sont *identifiés* à des troubles du cerveau. Plus précisément, cela signifie que, dans ces dernières, des mécanismes mentaux dysfonctionnels sont identifiés à des mécanismes neuraux dysfonctionnels. Il semble donc permis de soutenir que l'hypothèse de l'*identité* psychoneurale selon laquelle les mécanismes mentaux *sont* des mécanismes neuraux est sous-jacente à ces propos des défenseurs des RDoC (une formulation plus exacte de cette hypothèse sera donnée bientôt).

Dans la citation suivante, Parnas, un des critiques des RDoC, affirme aussi (mais en d'autres mots) que l'hypothèse de l'identité psychoneurale est sous-jacente à ce cadre de recherche : « Le fondement théorique des RDoC semble être un réductionnisme neuronal "type-type" : les "chunks" (types) spécifiques de la vie mentale [...] *sont*

identiques [...] [à] certains "chunks" (types) spécifiques d'activités neuronales [...]. » (Parnas, 2014, p. 46, mes italiques)

Andersen, qui est comme Parnas une critique des RDoC, soutient une thèse similaire à ce dernier : «[...] il semble clair qu'il est implicite, dans le cadre des RDoC, qu'on suppose une réduction ontologique. Elle prend la forme d'une identité type-type, où les troubles mentaux sont considérés comme n'étant rien d'autre que des troubles cérébraux. » (Andersen, 2017, p. 115)

Il convient de dire trois choses à propos de l'hypothèse de l'identité psychoneurale. La première chose que l'on peut affirmer est qu'elle s'inscrit dans une conception métaphysique matérialiste du monde. Selon cette dernière, (1) toutes les entités qui existent dans l'univers — incluant les mécanismes mentaux — sont matérielles et (2) une entité est matérielle (et, donc, fait partie de l'univers) si (a) elle est mutable et si (b) elle contient de l'énergie (Bunge, 2006, p. 10,12 et 29). Jodoin (2010, p. 446), suivant Bunge (2006, p. 10, 12 et 29), exprime de manière plus concise le contenu de cette conception métaphysique dans l'énoncé suivant :

 $\forall x \ (x \subset M = x \text{ est matériel} = x \text{ est mutable} = x \text{ a une énergie}), où M \text{ est le monde (ou l'univers)}^{52}.$

La deuxième chose que l'on peut dire à propos de l'hypothèse de l'identité psychoneurale est qu'elle est une hypothèse ontologique. En effet, elle porte sur la nature de l'esprit et sur le type de relation que celui-ci entretient avec le cerveau. Selon

_

⁵² Ce postulat métaphysique est utile parce qu'il permet de déterminer si une entité existe ou non. Par exemple, le chiffre « 2 » ne contient pas d'énergie. En conséquence, il ne change pas. Donc, il n'existe pas. Suivant cela, à moins d'être une partisane ou un partisan du platonisme, il convient de dire que le chiffre 2 est une fiction mathématique. Le cerveau, au contraire, existe. Il est une entité dynamique matérielle dans le monde. De fait, il change, mute et contient de l'énergie.

elle, et conformément à la conception métaphysique matérialiste décrite précédemment dans laquelle elle s'inscrit, un mécanisme mental *est* un mécanisme neural. Vu cela, comme le dit Bunge, ces deux mécanismes entretiennent une relation d'*identité* factuelle l'un avec l'autre (Bunge, 2010, p. 176). En d'autres termes, ils sont un seul et même mécanisme de la même façon que la chaleur et le mouvement atomique ou moléculaire aléatoire sont un seul et même mécanisme (Bunge, 2010, p. 176).

La formulation de l'hypothèse de l'identité psychoneurale qui convient le mieux pour les RDoC est à mon avis celle de Bunge⁵³ (2010, p. 161) : « Pour chaque processus mental M, il existe un processus N dans un système cérébral, tel que M=N. Pareillement : pour chaque fonction mentale F, il existe un système cérébral B qui effectue F. » (Bunge, 2010, p. 161)⁵⁴

Au moins trois corollaires découlent de cette hypothèse ontologique. Le premier est que la conscience phénoménale et les différents mécanismes cognitifs sont essentiellement constitués par des mécanismes neuraux. On peut prendre l'exemple de la mémoire de travail pour illustrer cela. Celle-ci est «[...] l'ensemble des processus qui représentent temporairement et conservent dans la conscience des informations vécues ou récupérées très récemment lorsque ces informations ne sont plus présentes dans l'environnement » (Welshon, 2011, p. 155). Conformément à l'hypothèse de

⁵³ D'autres personnes ont défendu l'hypothèse de l'identité psychoneurale (Smart, 2017). Cependant, étant donné que la formulation de Bunge (et son application à la recherche en neurosciences et en psychopathologie) convient au cadre des RDoC, je n'ai pas cru bon de les présenter. Ainsi, dans ma thèse, lorsque je parle de cette hypothèse, c'est celle que Bunge a formulée que j'ai en tête.

⁵⁴ Pour être plus cohérent avec Bechtel (2008) qui utilise l'expression «mécanisme mental» pour parler des processus mentaux (section 1.5.2), peut-être faudrait-il formuler cette hypothèse de la manière suivante : « Pour chaque mécanisme mental M, il existe un mécanisme N dans un système cérébral, tel que M = N. Pareillement : pour chaque fonction mentale F, il existe un système cérébral B qui effectue F.» J'ai néanmoins choisi de laisser la formulation de Bunge tel quel parce que je ne veux pas trahir sa pensée et parce que, de toute façon, dans mon travail, je ne fais pas de différence entre les notions de mécanisme et de processus.

l'identité psychoneurale, seulement des mécanismes neuraux constituent cette fonction mnémonique, spécialement ceux se trouvant dans le cortex préfrontal (Welshon, 2011, p. 158-159).

Le deuxième corollaire s'applique spécialement aux neurosciences, à la neuropsychologie, à la neurologie et à la psychiatrie. Il dit que « si B est lésé ou absent, [alors] F est perturbé ou ne se produit pas » (Bunge, 2010, p. 161). Par exemple, si dans le système cérébral responsable des fonctions langagières l'aire de Wernicke est lésée, alors la compréhension du langage (oral et écrit) est perturbée.

Le troisième corollaire s'applique spécialement à la psychiatrie. Il dit qu'un trouble mental est toujours réalisé exclusivement par un cerveau. Ce troisième corollaire s'harmonise parfaitement avec la conception du trouble mental des RDoC selon laquelle les troubles mentaux sont des troubles du cerveau. Comme l'explique Fuchs, cette façon d'appréhender les troubles mentaux tend à (a) isoler la personne malade des interactions qu'elle entretient avec l'environnement et à (b) considérer son trouble comme ne relevant que des systèmes à l'intérieur d'elle. Et ce, malgré le fait que l'hypothèse de l'identité psychoneurale (à l'instar des RDoC, voir section 1.3, principe 5) ne soit pas incompatible avec l'idée selon laquelle le cerveau est influencé, par l'intermédiaire des processus épigénétiques, par certains facteurs environnementaux comme des traumas à l'enfance et des relations d'attachement perturbées (Fuchs, 2018, p. 253). Je pense que ce sont les conséquences (a) et (b) de ce troisième corollaire qui expliquent au moins en partie pourquoi, au sein des RDoC, on ne fait pas de recherches

sur les mécanismes environnementaux (physiques, sociaux et culturels) favorisant le développement des troubles mentaux⁵⁵ (voir aussi section 2.1.1 à ce sujet).

La troisième chose que l'on peut dire à propos de l'hypothèse de l'identité psychoneurale est qu'elle est au fondement d'une très grande partie et même peut-être de toute la recherche en neurosciences cognitives (Bechtel et McCauley, 1999; Bunge, 2010, p. 161 et 172; voir aussi Mandik, 2014, chapitre 6, qui semble endosser cette idée; McCauley et Bechtel, 2001). Deux raisons me font croire cela. La première est qu'elle oriente de manière générale les études faites sur l'esprit. En effet, elle « motive une recherche neurocognitive sérieuse» (Bunge, 2010, p. 162) qui vise à faire « correspondre [les activités de l']esprit à celles du cerveau » (Bunge, 2010, p. 161) ou, autrement dit, à « localiser les fonctions mentales dans le cerveau » (Bunge, 2003b, p. 172). Dans le même ordre d'idées, Bechtel et McCauley (1999, p. 71) soutiennent que les énoncés contenant « des relations d'identité entre des processus psychologiques et des mécanismes neuronaux sont [...] [utilisés] comme des heuristiques qui servent à orienter les recherches [...] ». Je soutiens que les construits des RDoC jouent ce rôle d'outils heuristiques. En effet, ces derniers sont conçus comme des entités conceptuelles qui supposent une correspondance entre un mécanisme mental et un mécanisme biologique (voir section 1.3, principes 2 et 3), laquelle doit être validée à l'aide de méthodes quantitatives en faisant de la recherche dans le cadre des RDoC (voir section 1.5.4).

⁵⁵ On pourrait contre-argumenter que l'hypothèse de l'identité psychoneurale n'empêche pas, *en principe*, de faire de la recherche sur ces mécanismes environnementaux. J'en conviens. Mais je persiste à penser que cette hypothèse, à tout le moins, ne *commande* pas de recherches sérieuses sur les propriétés intrinsèques des mécanismes environnementaux favorisant le développement des troubles mentaux (voir à la fin de la section 2.1.1 pour un exemple de ce type de recherches où l'on s'intéresse de près aux mécanismes environnementaux). Que l'hypothèse de l'identité psychoneurale de Bunge n'oblige pas ce type de recherches est peut-être dû au fait que, bien qu'elle soit compatible avec l'idée selon laquelle l'esprit est incrusté dans un environnement, *elle ne contient pas cette idée*.

La deuxième raison qui me porte à croire que l'hypothèse de l'identité psychoneurale est à la base des recherches en neurosciences est qu'elle saisit le fait incontestable que le cerveau est une entité *matérielle* qui (a) se développe, mute ou change dans le temps et qui (b) contient de l'énergie. En effet, toutes sortes de mécanismes (fonctionnels et potentiellement dysfonctionnels) sont opérants dans le cerveau. Ces mécanismes sont étudiés lors des recherches menées dans le cadre des RDoC (voir section 1.4). On compte parmi eux les mécanismes mentaux (cognitifs et comportementaux)/cérébraux (les circuits) et les mécanismes moléculaires et cellulaires qui leur sont sous-jacents. On compte également les mécanismes physiologiques (comme la sécrétion de cortisol). Enfin, il y a les mécanismes que l'on pourrait appeler «neurocognitivo développementaux » (voir section 1.3, principe 5). Parmi eux, on retrouve les mécanismes normaux du vieillissement. Par exemple, le niveau de performance des facultés cognitives chez un humain change avec le temps. Notamment, les capacités mnémoniques diminuent au fur et à mesure qu'il avance en âge. Les mécanismes neurocognitivo développementaux anormaux, quant à eux, sont opérants dans plusieurs troubles mentaux. Par exemple, on observe une diminution progressive de la quantité de matière grise chez les enfants qui développent à l'adolescence une schizophrénie, et ce, bien avant que n'apparaissent chez eux les symptômes de ce trouble (NIMH, 2008, p. 16).

1.6 Conclusion du premier chapitre

Dans ce premier chapitre, j'ai montré d'où proviennent les RDoC (section 1.1) et présenté les raisons principales qui ont motivé le NIMH à les développer (section 1.2). J'ai aussi expliqué en détail ce que sont les RDoC (section 1.3) et le type de recherches que l'on mène en leur sein (section 1.4). Enfin, à la section 1.5, j'ai présenté ce qu'à mon avis sont la conception du trouble mental des RDoC (section 1.5.1) et les bases philosophiques sur lesquelles elle repose (sections 1.5.2-1.5.5). Ces dernières sont une

version à la fois pluraliste et réductionniste du néomécanisme de Bechtel qui partage un air de famille avec le réductionnisme fragmentaire de Schaffner (section 1.5.2), une interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Boorse (section 1.5.3), le principe épistémoquantitatif (qui suppose ce que je pense être une variante quantitative du réalisme des entités basé sur la robustesse d'Eronen) (section 1.5.4) et l'hypothèse de l'identité psychoneurale formulée par Bunge (section 1.5.5). À la lumière de ces bases philosophiques, on peut dire au moins deux choses. La première est qu'au sein des RDoC les troubles mentaux sont identiques (au sens de l'hypothèse de l'identité psychoneurale de Bunge) à des mécanismes neuraux (au sens du néomécanisme de Bechtel) dysfonctionnels (au sens de l'interprétation adaptationniste de la TBS). La deuxième est que, pour expliquer ces troubles mentaux, conformément au principe épistémoquantitatif, on mesure quantitativement l'activité de ces mécanismes à l'aide des méthodes et technologies utilisées en neurosciences moléculaires, cognitives et comportementales.

CHAPITRE II

CRITIQUE DE LA CONCEPTION DU TROUBLE MENTAL DES RDOC

Dans ce second chapitre, je critique l'arrière-plan philosophique (vu à la section 1.5) sur lequel repose la conception du trouble mental des RDoC que j'ai formulée et explicitée au chapitre précédent. D'abord, je critique l'hypothèse de l'identité psychoneurale sous-jacente à la conception du trouble mental des RDoC en montrant qu'elle véhicule une conception de l'esprit inadéquate (sections 2.1 et 2.1.1-2.1.4). Je soutiens également que cette dernière, parce qu'elle est problématique, nuit à la recherche psychiatrique (sections 2.1 et 2.1.1-2.1.4). Puis, j'explique en quoi ce que j'ai nommé le principe épistémoquantitatif sous-jacent à la conception du trouble mental des RDoC empêche de saisir certaines facettes des troubles mentaux et, conséquemment, nuit, à l'instar de l'hypothèse de l'identité psychoneurale, à la recherche en santé mentale (sections 2.2, 2.2.1 et 2.2.2). Après, je soutiens que les tenants des RDoC ne devraient pas adopter la TBS (section 2.3). En effet, elle est incapable de saisir ce qu'est la normativité biologique et, en conséquence, ce qu'est une dysfonction. Cela pose un problème puisqu'un concept de trouble mental doit être capable de saisir les dysfonctions chez les individus. Enfin, pour ajouter un complément à cette critique de l'arrière-plan philosophique des RDoC, j'examine les conceptions du trouble mental contemporaines de Wakefield (2010; 1992, 2006a; 2019) (section 2.4.1) et de Powell et Scarffe (2019b, 2019a) (section 2.4.2) et montre qu'elles ne sont pas des alternatives viables à la TBS (section 2.4.3). En effet, ces dernières

échouent elles aussi à appréhender la normativité biologique. Ainsi, elles ne peuvent pas remédier aux insuffisances de la TBS (et des RDoC).

2.1 L'hypothèse de l'identité psychoneurale de Bunge est problématique et nuit à la recherche psychiatrique

Dans cette section, je montre que (1) l'hypothèse de l'identité psychoneurale de Bunge⁵⁶ que suppose la conception du trouble mental des RDoC est problématique et que, pour cette raison, (2) elle nuit à la recherche en psychiatrie. Pour démontrer (1) et (2), je fais deux choses (sections 2.1.1 à 2.1.4). Premièrement, je présente des résultats d'études empiriques (sections 2.1.1-2.1.3) et/ou des notions théoriques⁵⁷ qui vont à l'encontre de l'hypothèse de l'identité psychoneurale. Deuxièmement, tirant des leçons pour la psychiatrie de ces résultats empiriques et/ou de ces considérations théoriques, je présente des types de recherches *non* neurocentrées qui devraient, à mon avis, être menées en santé mentale, mais qui ne le sont pas dans le cadre des RDoC.

Plus précisément, à la section 2.1.1, suivant plusieurs résultats d'études empiriques que je tire en grande partie de Jasanoff (2018), je montre que les mécanismes mentaux sont fortement *incrustés* dans l'environnement (ou, en d'autres termes, en relation étroite avec ce dernier). Après, je présente une étude de Joshi et son équipe (2017) sur l'impact de la pauvreté de quartier sur le développement de la dépression majeure (DM) chez les personnes âgées.

À la section 2.1.2, je présente des résultats d'études empiriques qui indiquent que les mécanismes mentaux sont *incarnés* (c'est-à-dire *constitués* de mécanismes corporels).

⁵⁷ Celle de couplage et celle de système dynamique à la section 2.1.3 et celle d'institution cognitive à la section 2.1.4.

-

⁵⁶ L'hypothèse de l'identité psychoneurale de Bunge a été vue à la section 1.5.5.

Après, j'expose la conception incarnée du trouble déficitaire de l'attention avec ou sans hyperactivité (TDAH) de Maiese (2012). Cette conception corporelle du TDAH me sert à montrer qu'il pourrait (et devrait) exister des conceptions *non* neurocentrées des différents profils cognitifs existant chez les humains pour guider la recherche psychiatrique.

À la section 2.1.3, suivant Varga (2018) et Harris et ses collègues (2014), je présente des études empiriques qui montrent qu'il existe des mécanismes mentaux qui émergent et sont *constitués* de mécanismes mentaux ayant lieu chez deux (ou éventuellement plus de deux) personnes *couplées* qui forment l'une et l'autre un *système dynamique*. Ensuite, m'inspirant de Bray (2008), j'illustre ce type de cas avec un exemple où une dysfonction sous-jacente à une dépression majeure, que présente une personne, ne se trouve pas dans cette personne, mais dans une autre personne. Je pense que ce cas de figure montre qu'il serait pertinent de mener des recherches sur les mécanismes multipersonnels à l'origine des troubles mentaux.

Enfin, à la section 2.1.4, suivant Gallagher qui pense que l'esprit peut être socialement étendu, je soutiens que les mécanismes mentaux des individus sont *constitués* d'institutions cognitives avec lesquelles ces mêmes individus sont couplés et forment un système dynamique (Gallagher, 2011, 2013, 2017, 2018; Gallagher *et al.*, 2019; Gallagher et Crisafi, 2009; Petracca et Gallagher, 2020; Slaby et Gallagher, 2015). Après, je présente la thèse de Merrit (2013) selon laquelle la dysfonction sexuelle féminine est constituée principalement par l'institution du genre (féminin et masculin). Ce cas montre que les dysfonctions à l'origine des troubles mentaux ne sont pas toujours nécessairement dans les personnes malades, mais peuvent également se trouver à l'extérieur d'elles, dans ce cas dans des institutions cognitives, et qu'il serait pertinent de s'intéresser à ces dernières afin de mieux comprendre l'étiologie de divers troubles mentaux.

Avant de poursuivre, j'aimerais noter que la thèse selon laquelle l'hypothèse de l'identité psychoneurale empêche que soient menées (a) des études sérieuses sur les mécanismes environnementaux causant les troubles mentaux et (b) des études sur les mécanismes corporels, multipersonnels et cognitivo-institutionnels qui sont sous-jacents à ces troubles n'est pas seulement *théorique*. En effet, cette thèse semble aussi confirmée *empiriquement*. Deux raisons me portent à soutenir cela.

Premièrement, selon mon examen, parmi les 48 études représentatives des recherches menées dans le cadre des RDoC que recensent Carcone et Ruocco (2017)⁵⁸, aucune d'entre elles ne porte sur la constitution corporelle et/ou extrapersonnelle et/ou cognitivo-institutionnelle des troubles mentaux.

Deuxièmement, parmi ces mêmes 48 études, seulement celle de Newman et ses collègues (2016) et celle de Sharp et ses collègues (2014) font référence à l'environnement, mais rien dans cette considération de l'environnement ne réfère à l'expérience ⁵⁹ qu'en ont les gens ou à la dynamique intrinsèque des mécanismes environnementaux qui favorisent le développement de troubles mentaux (pour un avis similaire, voir Paris et Kirmayer, 2016, p. 27). En effet, dans l'étude de Newman et ses collègues (2016), si on apprend à sa lecture qu'il existe une corrélation entre la consommation régulière de cannabis (un facteur environnemental) et le fait d'avoir un gyrus frontal inférieur plus mince que la normale (lequel est une structure impliquée dans la réponse d'inhibition), ce qui intéresse avant tout les chercheuses et chercheurs est de vérifier si on peut associer le fait de mal performer à la tâche Go/noGo (un test

⁵⁸ Je parle plus en détail de cette étude dans la note 15 en bas de page à la section 1.4.

⁵⁹ De telles recherches pourraient, par exemple, porter sur la manière dont les gens *perçoivent* différentes situations stressantes comme la pauvreté, une situation familiale compliquée ou un événement violent dont ils ont été témoins ou victimes. Ce type d'études, comme on peut le constater, ne porte pas sur la phénoménologie du trouble mental en soi, mais sur la manière dont les personnes présentant un trouble mental perçoivent la situation environnementale dans laquelle elles se trouvent.

psychologique qui examine les capacités d'inhibition) et le fait de présenter cette même structure cérébrale présentant une minceur plus importante que la moyenne. Bref, ce n'est qu'en passant qu'on mentionne cette information concernant l'impact de la consommation du cannabis sur le cerveau. Ainsi, on ne s'intéresse pas aux *raisons* qui motivent ces gens à consommer du cannabis. On n'examine pas non plus l'environnement dans lequel ceux-ci évoluent (normes sociales, disponibilité des produits, modes de consommation, etc.) et la manière dont ils expérimentent ce dernier.

Quant à l'étude de Sharp et ses collègues (2014), ce qu'on y examine prioritairement ne concerne pas l'environnement. En effet, on cherche essentiellement à valider ou infirmer l'hypothèse selon laquelle une faible activité du striatum ventral est un endophénotype⁶⁰ prédicteur de la dépression majeure (DM). Cette étude comporte trois échantillons : (1) un groupe d'adolescentes ne souffrant pas de DM ayant une mère présentant une histoire de DM (c'est le groupe à risque), (2) un groupe d'adolescentes actuellement déprimées ayant elles aussi une mère présentant une histoire de DM et (3) un groupe d'adolescentes qui ne sont pas déprimées et dont les mères n'ont pas d'histoire de DM (c'est le groupe contrôle). Pour vérifier leur hypothèse de départ, l'équipe de recherche tente de voir si le striatum ventral des adolescentes du groupe 1 fonctionne comme celui de celles du groupe 2, et ce, en comparant l'activité du striatum ventral des adolescentes des groupes 1 et 2 au fonctionnement (supposé normal) du striatum ventral des adolescentes du groupe 3. Si leur conjecture tient la route, alors l'activité du striatum ventral des adolescentes des groupes 1 et 2 devrait être semblable et moins importante que celle de celui des adolescentes du groupe 3. Comme on peut le constater, l'utilisation d'un facteur environnemental (l'histoire dépressive de la mère) sert ici essentiellement l'objectif central de l'étude qui est de vérifier si une activité

⁶⁰ Voir la note 27 en bas de page, à la section 1.5.2, pour une définition du terme « endophénotype ».

faible du striatum ventral est un endophénotype de la DM. Dans cette optique, la manière dont les adolescentes déprimées ont *vécu ou expérimenté* cette relation avec une mère présentant une histoire de DM n'intéresse pas les chercheuses et chercheurs.

Maintenant que j'ai précisé que l'hypothèse de l'identité psychoneurale n'empêche pas seulement en théorie, mais aussi dans les faits, certains types d'études en santé mentale, je vais commencer la critique de cette dernière et présenter des types de recherches psychiatriques qui, à mon avis, ne devraient pas être négligés par les tenants des RDoC.

2.1.1 Incrustation environnementale des mécanismes mentaux

Bien qu'elle ne le nie pas en principe, l'hypothèse de l'identité psychoneurale de Bunge semble faire oublier aux chercheuses et chercheurs travaillant dans le cadre des RDoC que nos mécanismes mentaux sont fortement incrustés dans un contexte physique et social⁶¹. En fait, ces derniers sont en relation étroite avec ces contextes. C'est une erreur de négliger cela parce que nos systèmes sensoriels principaux — la vue, l'ouïe, le toucher, le goût et l'odorat — captent continuellement des stimuli de l'environnement. La quantité de ces derniers que traitent les diverses structures et aires cérébrales est, en effet, énorme. De fait, « [...] le volume d'informations sensorielles convergeant vers le cerveau est de l'ordre de dizaines de millions de potentiels d'action chaque seconde [...] » (Jasanoff, 2018, p. 123). Comme le montre Jasanoff à l'aide de plusieurs études, ces intrants sensoriels provenant de l'environnement physique et social affectent de

-

⁶¹ Je suis parfaitement conscient que l'hypothèse de l'identité psychoneurale n'empêche pas la recherche sur les mécanismes environnementaux favorisant le développement des troubles mentaux. C'est pourquoi j'écris « en principe ». Mais comme je l'explique à la section 1.5.5, je pense que cette hypothèse, à tout le moins, n'*oblige pas* à ce que soient menées des recherches de ce type. Cela est peut-être dû au fait que l'hypothèse de l'identité psychoneurale, bien qu'elle soit compatible avec l'idée selon laquelle l'esprit est incrusté dans un environnement, *ne contient pas cette idée*.

manière fondamentale (et insoupçonnée) le fonctionnement de notre système nerveux et, conséquemment, nos comportements, nos mécanismes cognitifs et nos états émotifs.

Deux des études que rapporte Jasanoff portent sur l'effet de la température ambiante sur les émotions et les comportements des gens (c'est donc de l'influence de l'environnement physique sur le cerveau dans un corps dont il est question ici). La première étude que Jasanoff (2018, p. 127) résume est celle qu'un groupe de psychologues a menée sur des policiers hollandais qui se pratiquaient dans une salle de tir (Vrij et al., 1994). L'expérience consistait à faire varier la température de ce lieu d'entraînement. Les chercheuses et chercheurs ont observé que plus il faisait chaud, plus les policiers étaient hostiles et belligérants. Cette augmentation de leur agressivité se traduisait par une augmentation des tirs sur les faux contrevenants. Par exemple, les policiers tiraient 50 % de fois plus à 81 °F qu'à 70 °F. La deuxième étude (Hsiang et al., 2013) que Jasanoff (2018, p. 127) rapporte va dans le même sens que les résultats de l'étude menée sur les policiers. Elle montre qu'il existe un lien fort entre le climat et les divers types de conflits (ex. : taux de criminalité, violence conjugale, etc.), et ce, peu importe l'époque et l'emplacement géographique. Ces faits empiriques portent Jasanoff (2018, p. 127) à penser qu'un mécanisme physiologique quelconque (encore inconnu) serait impliqué dans ce phénomène où une température élevée cause une augmentation de l'agressivité chez l'humain.

Jasanoff présente également des études qui portent sur l'impact de l'environnement sonore sur nos mécanismes cognitifs. Il rapporte que des psychologues-chercheuses et psychologues-chercheurs (comme Salamé et Baddeley, 1982) ont montré que les bruits ambiants humains ou non humains avaient un effet néfaste sur la mémoire visuelle à court terme. Pour illustrer l'impact des bruits humains, Jasanoff rapporte les résultats d'une étude faite par la psychologue Elliott (2002). Celle-ci a fait passer des tests de mémoire visuelle à quatre cohortes composées de personnes d'âges différents (8,3 ans,

9,6 ans, 11,8 ans et 19 ans), et ce, tout en les obligeant à écouter des mots non pertinents en boucle. En présence de ces sons, et relativement aux résultats qu'ils ont obtenus dans le silence, les adultes (19 ans) performent 10 % moins bien et les élèves de deuxième (8,3 ans) performent 40 % moins bien. Cela montre que « les enfants sont plus affectés par les sons distrayants que les adultes » (Elliott, 2002, p. 483), et ce, probablement parce que les facultés attentionnelles sont moins bien développées chez les premiers que chez les derniers (Elliott, 2002, p. 486).

Enfin, Jasanoff rapporte des études sur l'impact de l'environnement social sur nos mécanismes cognitifs. Il présente, entre autres, l'expérience classique d'Asch (1951). Celle-ci met scène un groupe de personnes dont une seule est le sujet de l'expérience (mais celle-ci ne sait pas que toutes les autres sont des complices de l'expériencateur et qu'elle est la seule dans cette situation). L'expérience se déroule en trois étapes que l'on répète un certain nombre de fois. D'abord, on montre à ces personnes une feuille sur laquelle se trouve dessinée une ligne. Ensuite, on leur montre une feuille sur laquelle se trouvent trois lignes. Finalement, on leur demande de dire à voix haute tour à tour laquelle des trois lignes sur cette deuxième feuille ressemble le plus à celle qui se trouve sur la première feuille. Comme la différence entre les lignes est visuellement évidente, il est facile de répondre. Cependant, à certains moments pendant l'expérience, tous les complices ont pour consigne de donner la même mauvaise réponse. Dans ce cas de figure, la personne qui est le sujet de l'expérience, au moins à quelques reprises, adopte la position des autres membres du groupe, laquelle va pourtant totalement à l'encontre de ce que le sujet perçoit comme la bonne réponse.

On peut inférer trois leçons des expériences (et de bien d'autres) que je viens de décrire. La première est que les mécanismes mentaux sont incrustés dans les mécanismes environnementaux (autrement dit, ils sont en relation étroite avec eux) et fortement influencés par ceux-ci (on peut penser à l'effet de la température sur l'humeur). La

deuxième leçon concerne la psychiatrie et est une conséquence de la première leçon. Elle dit que dans certains types d'environnements, que l'on pourrait qualifier de malsains, il y a de très bonnes chances que le fonctionnement de l'esprit soit gravement perturbé (les mécanismes environnementaux ont, de fait, un effet important sur le développement de ce dernier). On peut observer cela quand un individu vit dans la pauvreté, subit de la violence ou évolue dans un contexte de guerre. La troisième leçon qui découle des deux premières est qu'il devrait y avoir des recherches psychiatriques sérieuses portant sur les effets de l'environnement sur la santé mentale. En effet, parce que les mécanismes mentaux (et les mécanismes biologiques qui leur sont sous-jacents) sont fortement incrustés dans les mécanismes environnementaux physiques, sociaux et culturels, ces investigations doivent avoir pour objectif de comprendre *en détail* ces derniers.

Pour illustrer ces trois leçons, on peut prendre l'impact de la pauvreté de quartier (un facteur environnemental) sur le développement de la dépression majeure (DM) chez les personnes âgées. Plusieurs mécanismes environnementaux sont opérants dans ce type de quartier et pourraient favoriser le développement de la DM chez ces dernières. Joshi et ses collègues (2017, p. 138-139) mentionnent quelques mécanismes candidats potentiels :

- (1) Les gens vivant dans les quartiers pauvres sont exposés à plus d'événements stressants de la vie (ex. : une perte d'emploi) et le stress cause la DM.
- (2) Dans les quartiers pauvres, il y a généralement un manque de commodités (comme des organismes communautaires ou des églises) qui favorisent l'organisation sociale et l'engagement social. Ce manque contribue à l'isolement des gens et cet isolement contribue au développement de la DM.

- (3) Les quartiers où le niveau de pauvreté est élevé présentent un niveau faible de cohésion sociale et de contrôle social. Ces quartiers présentent également un niveau élevé de criminalité et de défauts physiques (ex. : déchets abondants, bâtiments abandonnés, etc.). D'une part, un niveau élevé de criminalité et de défauts physiques peut générer de la peur chez les gens. D'autre part, un niveau bas de la cohésion sociale et un niveau bas du contrôle social peuvent réduire la confiance que les gens ressentent les uns envers les autres et, pour cette raison, peuvent nuire aux efforts des gens du quartier qui visent à faire réduire la criminalité et la quantité de défauts physiques qu'on y trouve. Joshi et son équipe notent, en outre, que des niveaux faibles de cohésion et de contrôle sociaux peuvent faire diminuer le niveau du support que les gens du quartier s'accordent mutuellement. Cela est regrettable parce qu'un bon niveau de ce support est nécessaire pour faire baisser, d'une part, la peur et le stress qu'engendre le haut niveau de criminalité et de défauts physiques dans le quartier et, d'autre part, la prévalence de la DM (en effet, la peur et le stress causent ce trouble mental).
- (4) Les quartiers pauvres contiennent moins d'espaces verts et d'endroits pour marcher et on a remarqué que de telles caractéristiques étaient associées à des prévalences plus élevées de DM.

Parmi ces mécanismes, lequel semble favoriser le plus le développement de la dépression majeure (DM) chez les personnes âgées? À la suite d'une étude longitudinale, Joshi et son équipe ont trouvé que « la violence de quartier est un mécanisme clé par lequel la pauvreté de quartier détermine le risque de dépression chez les personnes âgées » (Joshi *et al.*, 2017, p. 141). En effet, « [1]e taux d'homicide représente 30,0 % de l'association de la pauvreté du quartier avec la dépression » (Joshi *et al.*, 2017, p. 141).

La corrélation établie par Joshi et ses collègues a de la valeur parce qu'elle permet de faire des interventions bénéfiques qui ciblent un facteur de risque important (la violence) dans certains quartiers pauvres. En effet,

[l]'investissement dans la prévention de la violence dans les quartiers très pauvres permettra non seulement de réduire la violence, mais pourrait également avoir des avantages secondaires en améliorant la santé mentale dans les quartiers les plus vulnérables. (Joshi *et al.*, 2017, p. 142)

Ce genre d'intervention est encore plus pertinent considérant le fait que les personnes âgées sont vulnérables, entre autres, parce qu'elles sont moins mobiles (Joshi *et al.*, 2017, p. 138). En effet, les personnes faisant partie de cette tranche d'âge n'ont généralement accès qu'aux ressources offrant un support qui sont situées près de chez elles. Les personnes âgées pauvres ont également moins la capacité de déménager que les personnes moins âgées.

Ce cas montre que, lorsqu'on fait de la recherche sur les troubles mentaux, il ne suffit pas de s'intéresser seulement à ce qui se passe dans le cerveau des gens. En effet, il faut aussi s'intéresser à l'environnement où évoluent ces derniers. Et cette prise en compte de l'environnement ne doit pas être superficielle. Par exemple, si l'on reprend l'étude présentée précédemment, il n'aurait pas suffi, comme on aurait pu s'en contenter si l'on avait fait une recherche sur la DM dans le cadre des RDoC, de seulement mentionner au passage le fait avéré que la pauvreté a des effets néfastes sur le cerveau des gens et sur leurs fonctions cognitives (Lende, 2012) et le fait établi que les gens pauvres sont plus susceptibles de développer une dépression majeure que les gens ne vivant pas une situation de pauvreté (Gouvernement du Québec, 2019). En effet, on doit aussi faire des études qui portent sur les propriétés intrinsèques de l'environnement (ex. : les défauts physiques dans le quartier, la diminution de la cohésion sociale) où évoluent les gens développant un tel problème de santé mentale. Il faut aussi mettre en lumière les mécanismes environnementaux auxquels les

personnes pauvres sont confrontées et qui favorisent le développement de divers troubles mentaux chez elles (ex. : quartier pauvre + taux élevé d'homicide → baisse de la cohésion sociale et du contrôle social → baisse du niveau de support que les gens du quartier s'accordent mutuellement → augmentation de la probabilité chez les gens de vivre du stress → DM). C'est ce qu'ont fait Joshi et ses collègues (2017).

2.1.2 Composantes corporelles des mécanismes mentaux

On peut opposer à l'hypothèse de l'identité psychoneurale la conjecture selon laquelle les mécanismes mentaux sont constitués en partie de mécanismes corporels. Plusieurs études, en effet, montrent que le cerveau et le corps entretiennent des liens intimes lors de la génération de la vie mentale.

On peut prendre le cas des émotions pour illustrer cela. En effet, de nombreuses études indiquent que, chez les individus, chacune des émotions est en relation avec certaines mesures de phénomènes corporels, comme l'activité musculaire, le rythme du cœur et de la respiration de même qu'avec la conductance cutanée (Jasanoff, 2018, p. 101). Jasanoff rapporte que Kreibig (2010) a fait une analyse de centaines d'études où sont rapportées ce type de corrélations. Cette dernière a découvert que l'on pouvait mettre en relation des activités précises du corps pour chaque type d'émotion. Par exemple, elle observe que le sentiment de bonheur est caractérisé, chez les gens, «[p]ar une augmentation de l'activité cardiaque due à une réduction de l'activité du nerf vague, à une vasodilatation, à une augmentation de la conductance cutanée et à une augmentation de l'activité respiratoire » (Kreibig, 2010, p. 406).

Kreibig note également que bien que le sentiment du bonheur partage avec diverses émotions négatives une augmentation de l'activation cardiaque (en raison de la baisse de l'activité du nerf vague), il se distingue de ces dernières par la présence d'une vasodilatation périphérique (Kreibig, 2010, p. 407). L'analyse de Kreibig révèle, en

outre, que l'on peut distinguer des émotions similaires grâce aux activités corporelles différentes associées à ces dernières. Par exemple, Jasanoff rapporte que

[b]ien que l'anxiété et la tristesse (sans pleurs) sont toutes deux associées à une augmentation de la fréquence de la respiration, ces deux émotions tendent à être associées à des changements opposés sur le plan du rythme cardiaque, de la conductance cutanée et du volume respiratoire. (Jasanoff, 2018, p. 101)

Jasanoff (2018, p. 101-102) présente également les résultats d'une étude de Nummenmaa et ses collègues (2014) dans laquelle sont rapportés les propos de 701 personnes qui témoignent de la manière dont les différentes émotions s'inscrivent en elles et sont ressenties dans leurs corps. Ces témoignages ont été recueillis alors que les participants prenaient connaissance de mots émotifs, d'histoires, de films et d'expressions faciales et identifiaient en même temps, à l'aide de l'outil informatique d'auto-évaluation topographique *emBODY*, sur une silhouette où, dans leur corps, ils ressentaient une augmentation de l'activité somatique et où, sur une autre silhouette, ils ressentaient une baisse d'activité somatique (voir Figure 2.1).

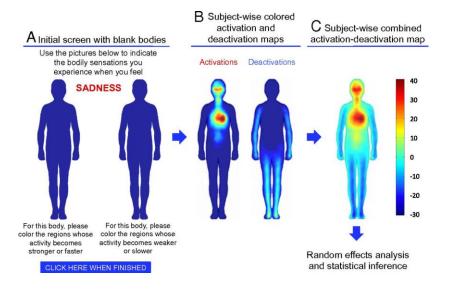


Figure 2.1 Outil informatique d'auto-évaluation topographique emBODY (Nummenmaa *et al.*, 2014, p. 647)

À la suite d'une analyse statistique de toutes ces données fournies par les participantes et participants, les auteures et auteurs de l'étude ont tiré des cartes de l'inscription corporelle de quatorze émotions (colère, peur, dégoût, bonheur, tristesse, surprise, neutralité émotionnelle, anxiété, amour, dépression, contemplation, fierté, honte et envie) dont on peut prendre connaissance dans la figure 2.2 suivante :

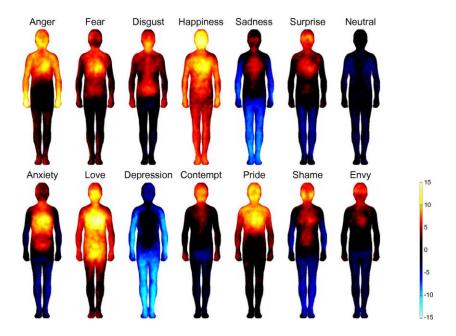


Figure 2.2 Topographie corporelle des émotions basiques (en haut) et non basiques (en bas) (Nummenmaa *et al.*, 2014, p. 647)

Dans ce schéma, pour chacune des émotions, une augmentation de l'activité somatique est représentée par une couleur chaude et une diminution de l'activité somatique est représentée par une couleur froide. Les auteures et auteurs de cette recherche pensent que ces schémas sont universels parce que des études menées sur des personnes provenant de pays différents (Finlandais, Suédois et Taiwanais) donnent des résultats similaires (Nummenmaa *et al.*, 2014, p. 646; cette thèse est aussi défendue dans Volynets *et al.*, 2019).

Jasanoff note que « les réponses émotionnelles à l'échelle du corps comme celles que Nummenmaa a cartographiées peuvent jouer un rôle intégral dans la cognition » (2018, p. 102). Pour défendre cette idée, Jasanoff utilise ce que dit l'hypothèse des marqueurs somatiques de Damasio (1994). Selon cette dernière, les décisions que l'on prend sont influencées sans qu'on en ait conscience par des sensations émotionnelles (que Damasio nomme des « marqueurs somatiques »). Cette hypothèse stipule que les

décisions sont influencées par les expériences somatiques qui ont suivi les décisions similaires qu'on a prises dans le passé. On évite donc certaines décisions qui ont provoqué des expériences désagréables et on est renforcé à prendre certaines décisions qui ont été suivies par des expériences positives.

On vient de voir que les *états* du corps sont impliqués dans nos mécanismes cognitifs. Mais la contribution de ce dernier à notre cognition ne s'arrête pas là. En effet, les *caractéristiques* de notre corps influencent également notre activité cognitive. Pour illustrer cela chez l'humain, on peut prendre le cas du grand violoniste, guitariste et compositeur Niccolò Paganini que présente Jasanoff (2018, p. 104) et celui du très célèbre et important physicien anglais Stephen Hawking qui a vécu, à partir de la vingtaine, avec la sclérose latérale amyotrophique (SLA).

Paganini avait, pense-t-on, une dysfonction des tissus conjonctifs qui dotait ses mains d'une très grande flexibilité. Cette propriété influençait sa manière de composer de la musique. En effet, certaines des pièces pour violon qu'il a composées présentent des difficultés techniques extrêmes et lui sont probablement venues à l'esprit parce qu'il possédait cette immense flexibilité des mains. Une personne ne possédant pas de telles caractéristiques n'aurait pas pu, ou du moins pas aussi facilement, concevoir de telles partitions musicales⁶².

Quant à Hawking, Mialet (2014, p. 93-94) suggère que sa manière de penser était influencée par ses limitations physiques. En effet, comme il ne pouvait pas effectuer

62 On pourrait contre-argumenter qu'une personne possédant d'immenses capacités créatrices musicales ou n'étent per violoniste pourrait à l'instant de Paganini, composer des pièces de musique pour violon

ou n'étant pas violoniste pourrait, à l'instar de Paganini, composer des pièces de musique pour violon présentant des difficultés techniques extrêmes. Le *Concerto pour violon en ré majeur, op. 77* de Johannes Brahms pourrait illustrer ce cas de figure. Néanmoins, je persiste à croire que les caractéristiques physiques d'une personne peuvent avoir une influence sur sa cognition.

des calculs mathématiques avec ses mains (en raison de sa maladie qui l'empêchait d'écrire, de marcher, de bouger, etc.), Hawking a développé une manière visuelle et géométrique de réfléchir. Il utilisait notamment les diagrammes de Penrose dans lesquels ils se projetaient mentalement afin d'intégrer une réalité virtuelle où il tentait alors de résoudre les problèmes de cosmologie qui l'intéressaient. Cette manière de réfléchir, on peut le penser, a eu un impact sur le type de travaux qu'Hawking a fait en physique théorique.

Ces exemples suggèrent que les mécanismes cognitifs peuvent être façonnés et constitués (au moins en partie) par les états et les caractéristiques du corps. Pour cette raison, il est raisonnable de penser que, si le corps est altéré, alors l'esprit changera. C'est justement ce que montre Jasanoff dans son livre à l'aide de plusieurs études. Deux d'entre elles portent sur les impacts des transplantations fécales sur l'esprit (Jasanoff, 2018, p. 112-113). La première porte sur des humains. Chez ces derniers, non seulement la transplantation fécale modifie le microbiote du receveur et affecte leurs fonctions digestives, mais elle cause aussi des changements de leur niveau d'anxiété, de stress et de dépression (Dinan et al., 2015; Smith, 2015). La deuxième étude porte sur des souris de deux souches différentes. Les membres d'une équipe de recherche ont fait une double transplantation fécale chez ces souris (Bercik et al., 2011). Les membres de cette équipe ont pris des matières fécales des souris BALB/c (qui sont plutôt timides, évitent la lumière et explorent peu leur environnement) et les ont transplantées dans les souris NIH Swiss (plus exploratrices et frondeuses que les souris BALB/c). Les résultats de cette étude sont spectaculaires. En effet, après que les souris NIH Swiss aient reçu les matières fécales des BALB/c, elles sont devenues plus anxieuses et timides. Les membres de l'équipe de Collins ont aussi transplanté des matières fécales des souris NIH Swiss dans les souris BALB/c. Le résultat a été que ces dernières sont devenues plus exploratrices et moins anxieuses. Bref, ces deux recherches montrent clairement que, chez les humains comme chez les souris, la composition du microbiome a un impact direct sur la vie psychique. On pense que c'est par l'intermédiaire de l'axe microbiote-intestin-cerveau que les microorganismes dans l'estomac influencent les mécanismes cérébraux (Jasanoff, 2018, p. 113).

On peut tirer encore une fois trois leçons des travaux que j'ai présentés. La première est que certaines « des fonctions mentales que nous attribuons habituellement au cerveau sont en fait des fonctions du corps tout entier » (Jasanoff, 2018, p. 97). Ce constat va manifestement à l'encontre de l'hypothèse de l'identité psychoneurale au cœur des RDoC qui dit que les mécanismes mentaux sont *essentiellement* constitués par des mécanismes neuraux (section 1.5.5). La deuxième leçon concerne la psychiatrie. Elle dit que « quand les choses tournent mal pour le corps, l'esprit peut aussi souffrir [...] » (Jasanoff, 2018, p. 97). Cela semble compatible avec le fait que plusieurs troubles de santé mentale sont comorbides avec des troubles physiques. La troisième leçon découle de la première et de la deuxième et dit qu'en présence d'un trouble de santé mentale, on devrait (ou pourrait) intervenir sur le corps et pas seulement sur le cerveau.

Pour illustrer ces trois leçons, il convient de présenter le travail qu'a fait Maiese (2012) sur le TDAH. Celle-ci tient compte de l'importance du corps lorsqu'elle développe sa conception (spéculative)⁶³ de ce profil mental. Sa conception de ce profil cognitif s'articule principalement autour de l'idée selon laquelle tous les systèmes vivants possèdent des *capacités affectives d'encadrement [affective framing capacities*]. Celles-ci sont « [des] processus neurobiologiques *pleinement incarnés* qui impliquent

⁶³ Je spécifie qu'il s'agit d'une conception spéculative parce que Maiese dit que sa « caractérisation des causes sous-jacentes du TDAH ainsi que les implications correspondantes pour le traitement [(je présente ces dernières à la fin de cette section)] sont tout à fait spéculatives et nécessitent une étude empirique plus poussée » (Maiese, 2012, p. 914-15).

non seulement le cerveau, mais aussi *les autres systèmes vitaux de notre corps vivant* » (Maiese, 2012, p. 904, mes *italiques*)⁶⁴.

Les capacités affectives d'encadrement sont innées et opèrent sans qu'on en ait conscience. Elles sont guidées par les désirs et besoins d'un individu. Elles fonctionnent généralement au cours des expériences subjectives qu'un individu a du monde et de lui-même. Elles ont pour rôle de porter son attention sur ce qui est important pour lui dans son environnement (Maiese, 2012, p. 901). Elles sont également à l'origine de la production d'une interprétation non conceptuelle et préréflexive du monde (Maiese, 2012, p. 895). Elles opèrent donc en arrière-plan de l'expérience phénoménale générale et influent, en conséquence, sur la manière dont on interprète les personnes, les objets, les faits, les états de choses et les situations (Maiese, 2012, p. 901).

On peut également dire avec assurance que la manière dont fonctionnent ces capacités affectives d'encadrement joue un rôle dans la façon dont opèrent les *fonctions* cognitives de haut niveau. En effet,

[l]e cadrage affectif joue un rôle important dans la détermination des informations qui seront conservées dans la *mémoire de travail*, de celles qui disparaîtront et de celles qui feront, si [cela est] nécessaire, l'objet d'une attention consciente. [Le cadrage affectif] sous-tend également la capacité de stimuler, de soutenir et de renouveler l'énergie et la *motivation* nécessaires à la réalisation des plans et à l'atteinte des objectifs que l'on s'est fixés. En effet, le cadrage affectif attribue des poids et des priorités

-

⁶⁴ Peut-être est-ce parce qu'il s'agit d'une conception spéculative qui n'est pas confirmée empiriquement qu'il est difficile de comprendre à quoi réfèrent exactement, dans le corps, pour Maiese, ces processus neurobiologiques pleinement incarnés responsables de ce qu'elle appelle les *capacités affectives d'encadrement*?

afin que nous puissions *filtrer l'information* et *séquencer* les activités en fonction de leur importance [...]. (Maiese, 2012, p. 904, mes *italiques*)

Comme les capacités affectives d'encadrement influent sur la manière dont opèrent les fonctions exécutives, Maiese fait l'hypothèse que le dysfonctionnement des premières entraîne une perturbation des secondes (attention, planification, etc.). C'est ce dysfonctionnement de ces capacités et de ces fonctions qui serait, selon elle, à l'origine des symptômes cognitifs et comportementaux du TDAH, à savoir les déficits d'attention, de la concentration, de la mémoire de travail et le manque d'inhibition comportementale (Maiese, 2012, p. 895 et 903). Elle soutient donc que si un individu présente des capacités affectives d'encadrement dysfonctionnelles, alors il lui sera difficile « de garder à l'esprit les événements et les informations saillantes et [que] [s]es processus de rétrospection, de prévoyance et d'anticipation seront compromis » (Maiese, 2012, p. 907). C'est ce qui explique, selon elle, par exemple, que « les personnes atteintes du TDAH oublient souvent ce qu'elles étaient sur le point de dire ou pourquoi elles sont entrées dans une pièce » (Maiese, 2012, p. 907). C'est aussi ce qui pourrait expliquer qu'elles « [...] peuvent avoir de la difficulté à garder le sens d'un texte dans leur esprit lorsqu'elles lisent les parties suivantes ou qu'elles répondent à des questions après la lecture » (Maiese, 2012, p. 907).

En somme, comme on peut le constater, selon cette conception qui accorde de l'importance au rôle du corps dans le TDAH que formule Maiese, les systèmes neurobiologiques responsables des fonctions exécutives sont impliqués dans l'émergence des symptômes de ce profil mental, mais ils ne sont pas les premiers responsables. Ce sont les dysfonctionnements des capacités affectives d'encadrement qui le sont.

Une des conséquences thérapeutiques de cette conception incarnée du TDAH de Maiese est qu'intervenir intensément sur le plan comportemental « pourrait être le

traitement le plus efficace et durable du TDAH » (Maiese, 2012, p. 911). Concernant les enfants, dans les écoles, qui présentent ce profil cognitif, Maiese donne les conseils thérapeutiques suivants :

- (1) Dans le contexte où l'enseignante pose une question au groupe, (a) apprendre aux élèves à dire leurs réponses tranquillement à eux-mêmes plutôt que de les verbaliser à voix haute sans lever la main (et ainsi interrompre la classe) ou (b) apprendre aux élèves à verbaliser tranquillement à voix basse des expressions qui les aide à ce retenir telle que « Je peux attendre mon tour » (Maiese, 2012, p. 913).
- (2) Faire des routines d'entraînement (par exemple, faire des répétitions de mouvements corporels intentionnels ou faire des jeux sur la façon de rester assis sans bouger pendant les cours ou pendant les études) (Maiese, 2012, p. 913).
- (3) Utiliser un horaire ou un plan quotidien pour aider les élèves à développer leurs compétences organisationnelles et leur capacité à fixer un objectif et à avoir la motivation de l'atteindre, d'une part, et, d'autre part, récompenser les élèves qui respectent ce plan (Maiese, 2012, p. 913).

On peut penser que ce type comportemental d'interventions thérapeutiques est d'autant plus pertinent et souhaitable que les médicaments pour traiter le TDAH ont probablement, à long terme, des effets délétères sur le cerveau (Higgins, 2009).

Bien qu'elle soit spéculative et qu'elle n'ait pas été confirmée empiriquement (Maiese, 2012, p. 914-915), cette conception du TDAH de Maiese est néanmoins présentée ici pour donner une idée de la manière dont une prise en compte de la corporéité peut permettre de modifier la manière habituellement neurocentrée de concevoir de nos

jours les conditions psychiatriques. De manière générale, on peut penser que cette grille d'analyse pourrait éventuellement favoriser la mise en place de projets de recherche qui permettraient de conceptualiser à nouveau plusieurs troubles mentaux et les manières de les traiter. Malheureusement, en raison de l'hypothèse de l'identité psychoneurale qui est sous-jacente aux RDoC, ces derniers ne semblent pas être un cadre propice à ce type de recherches.

2.1.3 Composantes multipersonnelles de certains mécanismes mentaux

En plus d'opposer à l'hypothèse de l'identité psychoneurale le fait qu'il existe des mécanismes mentaux fortement *incrustés* dans l'environnement (section 2.1.1) et/ou *constitués* corporellement (et pas seulement cérébralement) (section 2.1.2), on peut faire valoir qu'il existe aussi des mécanismes mentaux qui sont *constitués* par l'activité cognitive de *deux* (ou de *plusieurs*) *personnes*. On observe ce type de constitution quand un mécanisme mental émerge⁶⁵ de l'activité cognitive ayant lieu chez deux personnes ou plus qui sont *couplées* l'une à l'autre. Des personnes sont couplées lorsque l'état et l'activité de chacune d'elles sont fonction de l'état et de l'activité de l'autre ou des autres (Thompson, 2007, p. 45). Ces personnes couplées sont chacune conçues comme des systèmes dynamiques qui sont définis comme des systèmes dont les propriétés changent au fil du temps selon une loi ou des lois⁶⁶ (Thompson, 2007, p. 38-39). Quand deux personnes sont couplées, alors elles forment ensemble un autre système dynamique physiquement plus gros que chacune d'elles prise individuellement. C'est de ce plus grand système dynamique qu'émergent les mécanismes mentaux constitués multipersonnellement.

65 Je présente, à l'annexe C, la conception de l'émergence que j'adopte dans cette thèse.

_

⁶⁶ Je présente plus en détail, à l'annexe B, ce qu'est un système dynamique.

Dans ce qui suit, je présente deux exemples de mécanismes mentaux émergeant chacun d'un système dynamique. Ces deux exemples invalident l'hypothèse de l'identité psychoneurale et on peut tirer d'eux une leçon pour la recherche en psychiatrie que j'illustre à la fin de cette section.

Le premier exemple est la *régulation des émotions chez les nourrissons*. Ce mécanisme mental a habituellement pour but de diminuer la durée et l'intensité des émotions négatives et de maintenir la présence des émotions positives chez l'agent cognitif (Varga, 2018, p. 442). Ce mécanisme mental émerge de certaines activités cognitives du nourrisson *et* de certaines de celles de ses parents. Il en est ainsi, entre autres, parce que le nourrisson et ses parents sont *couplés* ensemble. Le mécanisme mental de la régulation des émotions chez le nourrisson émerge donc du grand système dynamique que forment ce dernier et ses parents. Dans ce cas, on semble en droit de dire que le mécanisme mental de la régulation des émotions, chez le nourrisson, est *constitué* par l'activité cognitive du nourrisson et par celle de ses parents.

Il est bénéfique pour le nourrisson de faire partie de ce grand système dynamique. En effet, le nourrisson a besoin d'être partie prenante de ce dernier pour réguler ses émotions parce que les mécanismes mentaux dont il dispose pour gérer ces dernières, contrairement à ceux de ses parents, ne sont pas développés. Par exemple, un adulte qui vit une émotion négative trop intense et désagréable en raison d'une situation stressante est généralement capable de réévaluer ou de mettre à distance cette dernière alors qu'un bébé ne le peut pas.

La relation de couplage entre le nourrisson et ses parents et, en conséquence, l'existence du système dynamique « nourrisson-parents » est possible parce que le nourrisson et ses parents mettent en œuvre des habiletés interactionnelles (Varga, 2018, p. 442). On remarque, en effet, que le nourrisson est capable de répondre avec les

mouvements de son corps aux microcomportements de ses parents tels que leur « niveau d'éveil, la direction [de leurs] regards, [leurs] expressions faciales et corporelles et le ton de [leurs] voix » (Varga, 2018, p. 442).

Cette interaction entre le nourrisson et les parents est caractérisée, selon Varga, par une certaine synchronie, c'est-à-dire par une « configuration en séquences rythmiques de microcomportements socialement pertinents, dans lesquelles les comportements [du nourrisson et du parent] sont appariés et non simplement copiés » (Varga, 2018, p. 442). Plus précisément, cette configuration synchrone suppose

[...] l'achèvement des vocalisations et des gestes de chacun, [achèvement qui] implique la mise en correspondance non seulement des comportements et des états émotionnels [du nourrisson avec ceux des parents], mais aussi de [quelques-uns de leurs] rythmes biologiques. (Varga, 2018, p. 442)

Il est à noter deux choses à propos de cette synchronie. La première est qu'elle est un phénomène qui conforte la validité de l'hypothèse selon laquelle il existe un couplage entre le nourrisson et ses parents parce qu'elle implique que l'état du premier est fonction (au moins en partie) de l'état des seconds et vice versa (quelques-uns de leurs rythmes biologiques covarient, par exemple).

La deuxième est que cette synchronie est nécessaire au bon fonctionnement du système dynamique que forment au fil du temps le nourrisson et les parents. Il importe que ce système fonctionne adéquatement (ou, plus précisément, selon le cadre énactif autopoïétique, qu'il reste viable, comme on le verra à la section 3.1) parce que c'est en participant à ce dernier que le nourrisson apprendra (inconsciemment) et réussira à réguler correctement ses émotions. Cet apprentissage est nécessaire parce que les émotions jouent un rôle très important dans nos existences. En effet, si le nourrisson n'apprend pas à bien gérer ses émotions, il développera des stratégies de régulation dysfonctionnelles, lesquelles pourraient diminuer sa qualité de vie. Par exemple, des

émotions trop intenses, qui durent trop longtemps et qui sont inappropriées pour certains contextes pourraient, quand il sera plus vieux, nuire à ses capacités de prises de décision ou à la qualité de ses relations sociales.

Le second exemple qui invalide l'hypothèse de l'identité psychoneurale est celui du *mécanisme mnémonique de récupération de souvenirs partagés par les personnes d'un couple de longue date*. Harris et ses collègues (2014) l'ont étudié. Ce mécanisme émerge de l'activité cognitive des *deux* personnes âgées formant ce couple. L'existence de ce mécanisme dépend donc du fait que deux personnes sont couplées l'une à l'autre dans la vie de tous les jours et, plus particulièrement, lors de cette tâche cognitive mnémonique de récupération de souvenirs. On peut donc affirmer que ce mécanisme mental émerge du grand système dynamique que forment ces dernières. En raison du couplage que ces personnes âgées entretiennent l'une avec l'autre et parce qu'elles sont des composantes de ce système dynamique, on peut dire de ce mécanisme mental mnémonique qu'il est *constitué* par les mécanismes cognitifs de ces dernières⁶⁷.

On peut faire l'hypothèse que ce sont les comportements de communication de chacune des personnes du couple qui rendent possible le couplage qu'elles entretiennent dans la vie quotidienne, couplage qui rend lui-même possible l'existence du système dynamique qu'elles forment ensemble. On peut définir ces comportements de communication comme des « comportements verbaux et non verbaux se produisant

⁶⁷ Dans la vie quotidienne, ce système dynamique est composé de deux personnes *et* d'éléments dans l'environnement (par exemple, un agenda) qui sont *couplés* (au sens donné précédemment). Ainsi, la composition du système dynamique sur lequel émerge la fonction de récupération des souvenirs du couple (et plus généralement ses fonctions mnémoniques) n'est pas sociale (les deux membres du couple) *ou* matérielle (un cahier de notes), mais sociale *et* matérielle (Harris *et al.*, 2014, p. 293-294). Il suit de cela que, dans la vie de tous les jours, ce mécanisme mental de récupération est *constitué* par l'activité cognitive des *deux* personnes âgées et rendu possible par certaines entités de leur environnement (par exemple, un cahier de notes).

[lorsque les personnes d'un couple sont] en interaction » (Kelly *et al.*, 2003, p. 724). Parmi les comportements non verbaux, on retrouve le contact visuel, le ton de la voix, l'expression faciale et la posture du corps qui peuvent tous être utilisés dans le cadre d'une communication (Kelly *et al.*, 2003, p. 726). Quant aux comportements verbaux, ils renvoient aux contenus des actions verbales des personnes. L'expression d'une approbation, l'expression d'un désaccord, l'expression d'un reproche ou l'expression de félicitations sont tous des exemples de ce type de comportements.

Dans le contexte expérimental où est étudié le mécanisme mnémonique de récupération de souvenirs partagés par les personnes d'un couple de longue date, les comportements de communication qui rendent possible le couplage des membres du couple prennent, selon mon interprétation des études présentées par Harris et ses collègues, la forme de stratégies communicationnelles de rappel. Il semble que, selon celles qui sont utilisées, le niveau de performance du mécanisme mnémonique du couple augmente ou diminue. De manière générale, on peut dire que, si les membres du couple se corrigent mutuellement ou perçoivent chez eux une inégalité d'expertise relativement à la tâche à accomplir, alors cela fait baisser le niveau de performance de ce mécanisme mental (Harris et al., 2014, p. 289). À l'inverse, si les personnes du couple se partagent des indices, élaborent ensemble de nouvelles informations et répètent ce que l'autre dit, alors cela augmente le niveau de performance de ce mécanisme mental, et ce, plus précisément sur le plan de la mémoire épisodique (Harris et al., 2014, p. 290-291). En effet, en contexte expérimental, quand les personnes d'un couple collaborent pour se souvenir d'endroits dans le monde où elles ont voyagé et, plus généralement, d'éléments autobiographiques qu'elles partagent, on observe trois propriétés émergentes relativement à leur mémoire épisodique (Harris et al., 2014, p. 291-292). Premièrement, il y a émergence de nouveaux détails qu'aucune des deux personnes à elle seule n'aurait pu se rappeler. Deuxièmement, il y a émergence d'une qualité supérieure des souvenirs au sens où ces derniers sont plus riches émotionnellement et plus vifs. Troisièmement, il y a émergence d'un surplus de compréhension (au sens où la compréhension d'un même événement qu'entretient chacun des membres du couple est différente après que ces derniers se le soient remémoré ensemble).

Il est pertinent de noter que l'amélioration de la mémoire épisodique n'est pas évaluée ici en fonction de la *quantité* de réminiscences différentes qui refont surface dans l'esprit des personnes âgées formant le couple. En effet, Harris et ses collègues soutiennent plutôt qu'il y a augmentation de la performance de la mémoire parce que les souvenirs partagés présentent ces *trois nouvelles propriétés*. Comme quoi « [...] la fonction de la réminiscence conjointe n'est pas toujours de se souvenir le plus possible ou le plus précisément possible » (Harris *et al.*, 2014, p. 292) et, qu'en conséquence, il faut développer des méthodes qui mesurent autre chose que la quantité de souvenirs (Harris *et al.*, 2014, p. 292).

Pour clore la présentation de ces résultats empiriques, j'aimerais mentionner qu'il est surprenant de découvrir que la performance de la mémoire épisodique est augmentée lorsqu'elle émerge d'une interaction entre les membres d'un couple de longue date parce que les personnes âgées prises individuellement présentent généralement des déficits sur le plan de leur mémoire épisodique (alors que leur mémoire sémantique — celle qui sert à récupérer des mots, des idées et des concepts — reste généralement intacte) (Harris *et al.*, 2014, p. 291).

Comme on peut le constater, ces considérations sur la constitution multipersonnelle de mécanismes mentaux chez le nourrisson et chez les personnes âgées formant un couple de longue date vont à l'encontre de l'hypothèse de l'identité psychoneurale parce qu'elles montrent qu'au moins certains mécanismes mentaux ne sont pas seulement constitués par les mécanismes neuraux d'une seule personne.

On peut tirer au moins une leçon pour la psychiatrie de ces considérations sur la cognition. La leçon est la suivante : si les mécanismes mentaux d'une personne particulière sont parfois constitués de mécanismes mentaux d'autres personnes, alors pour expliquer une dysfonction mentale chez une personne particulière, on doit aussi parfois s'intéresser aux mécanismes mentaux des autres personnes avec lesquels sont couplés les mécanismes mentaux de cette personne. En d'autres termes, quand on veut expliquer un trouble mental, l'*unité d'analyse* ne devrait pas être uniquement la personne qui présente le trouble mental (ou seulement son esprit/cerveau, comme c'est le cas dans les RDoC), mais plutôt le système dynamique dont cette dernière fait partie.

Pour illustrer cette leçon, on peut prendre le cas fictif d'une personne présentant un trouble de la personnalité limite (TPL) qui est depuis peu en couple avec une autre personne qui ne présente pas de condition mentale particulière. Les personnes présentant un TPL ont les traits suivants :

- Elles sont impulsives;
- Elles ont des relations interpersonnelles instables et intenses;
- Elles ont une immense peur de l'abandon;
- Elles ont une identité perturbée;
- Elles ont un sentiment de vide chronique;
- Elles ont de la difficulté à réguler leurs émotions. (Voir American Psychiatric Association, 2013, p. 663-666 pour plus de détails)

Étant donné ces traits, il est plausible de penser, suivant Bray (2008), que cette personne présentant un TPL pourrait éventuellement utiliser le cerveau de la personne avec laquelle elle est nouvellement en couple pour pallier les dysfonctions de son esprit. Par exemple, elle pourrait utiliser, selon Bray (2008), les mécanismes mentaux de cette autre personne pour réguler ses émotions. Il est plausible de penser que cette personne

présentant un TPL puisse utiliser les mécanismes mentaux de l'autre personne parce qu'elle entretient avec cette dernière, étant en couple avec elle, une relation de proximité. Considérant cela, il est plausible de penser qu'avec le temps, la personne dont les mécanismes mentaux sont (trop) utilisés pour réguler les émotions de la personne présentant un TPL se fatigue, s'aliène peu à peu et développe elle-même une détresse psychologique. Par exemple, elle pourrait développer une humeur dépressive.

Ce cas de figure et les considérations sur les composantes multipersonnelles de la cognition que l'on a vues précédemment montrent qu'il est plausible de penser que la dysfonction responsable d'un trouble mental chez une personne peut se trouver *dans une autre personne*. De la recherche psychiatrique devrait donc examiner ce type de phénomène. Mais la conception du trouble mental des RDoC ne permet pas de faire ce genre d'investigations parce qu'elle suppose que les dysfonctions à l'origine d'un trouble mental se trouvent *toujours* dans la personne qui présente le trouble mental.

À la section 3.2.1, j'expose la conception énactive et objective de la santé et du trouble (CÉOST). À cette occasion, je montre, entre autres, que la CÉOST contient cette thèse selon laquelle la dysfonction de la personne malade peut se trouver dans une autre personne, c'est-à-dire dans le système dynamique sociale dont elle fait partie et qui la traverse. En conséquence, à l'inverse de la conception du trouble mental des RDoC, la CÉOST permet de faire de la recherche sur ce type de systèmes.

2.1.4 Composantes cognitivo-institutionnelles des mécanismes mentaux

On a vu que plusieurs études empiriques montrent que les mécanismes mentaux sont fortement incrustés dans des mécanismes environnementaux (section 2.1.1). On a aussi défendu l'idée selon laquelle les mécanismes mentaux sont *constitués* de mécanismes cérébraux *et* corporels (section 2.1.2). On a également soutenu que les mécanismes mentaux d'une personne sont *constitués* par des mécanismes mentaux n'ayant pas lieu

seulement en cette dernière, mais aussi dans une autre personne avec laquelle elle est couplée et forme un système dynamique (section 2.1.3). Les RDoC, comme on l'a vu, ne permettent malheureusement pas d'étudier ces mécanismes que l'on pourrait généralement qualifier d'extraneuraux. Dans cette section, je fais valoir l'idée selon laquelle les mécanismes mentaux peuvent être *constitués* par ce que Gallagher et ses collègues appellent des « institutions cognitives »⁶⁸ (Gallagher, 2013; Gallagher *et al.*, 2019; Gallagher et Crisafi, 2009; Petracca et Gallagher, 2020; Slaby et Gallagher, 2015).

Gallagher conçoit la notion d'institution cognitive dans le cadre de la conception énactive autopoïétique de la cognition (CÉAC) (Petracca et Gallagher, 2020, p. 8, note 7)⁶⁹. Dans ce cadre, la cognition est conçue en termes d'activités de l'agent (lesquelles activités impliquent toutes sortes de processus à l'intérieur et à l'extérieur de cet agent). Plus précisément, la cognition y est appréhendée comme étant constituée et émergeant d'interactions qu'entretient continuellement un agent incarné avec son environnement (Petracca et Gallagher, 2020, p. 19). Cet environnement peut être physique, biologique, social et culturel. Les institutions cognitives font partie de l'environnement culturel de l'individu. Sans elles, les agents cognitifs ne pourraient pas opérer certains processus cognitifs. Plus précisément, elles permettent à ces agents de s'engager dans des activités cognitives qui ne peuvent pas être produites par une seule tête et même par plusieurs têtes (Gallagher, 2013, p. 7).

_

⁶⁸ Ici, j'expose essentiellement ce qui est nécessaire pour comprendre la notion d'institution cognitive. À l'annexe A, laquelle est au fond un développement de la section 2.1.4, j'expose plus en détail l'origine de cette notion et le paysage conceptuel dans lequel elle s'inscrit.

⁶⁹ La CÉAC sera vue en détail à la section 3.1. Dans ce qui suit, je ne fais que présenter les éléments de la CÉAC qui sont nécessaires pour la compréhension de ce qu'est la notion d'institution cognitive.

De manière générale, on peut concevoir une institution cognitive comme une affordance sociale (Gallagher, 2013, p. 4) qui offre des possibilités d'activités cognitives. Toujours de manière générale, parce qu'elles changent avec le temps, on peut aussi les concevoir comme des systèmes dynamiques ⁷⁰. De manière plus particulière, on peut définir l'institution cognitive comme un ensemble de pratiques cognitives produites à des moments et des lieux spécifiques (Gallagher, 2013, p. 6). Cet ensemble de pratiques, selon Gallagher et Petracca, « [...] fonctionne comme une entité de résolution de problèmes lorsqu'il est correctement *couplé* avec des agents cognitifs » (Petracca et Gallagher, 2020, p. 19, mes *italiques*).

À quelles conditions peut-on dire qu'un agent est *couplé* à une institution de ce type? Selon Gallagher, on peut affirmer qu'il y a un couplage à deux conditions (à mon avis, complémentaires). Premièrement, on peut dire qu'il existe un couplage s'il y a modification neurale dans le cerveau du sujet [(c.-à-d., l'agent cognitif)] qui s'engage avec les institutions [cognitives] (Gallagher, 2013, p. 10). Deuxièmement, on peut dire que « [d]eux ou plusieurs systèmes dynamiques sont couplés lorsque la conduite de chacun est fonction de la conduite de l'autre » (Thompson, 2007, p. 45) (cette conception du couplage a été introduite à la section 2.1.3).

C'est lorsqu'un tel couplage existe entre l'agent cognitif et l'institution cognitive (ou, autrement dit, c'est lorsque celui-ci et celle-ci forment un système dynamique) que l'institution cognitive peut être conçue comme une des *composantes* des processus cognitifs de l'agent. C'est aussi à ce moment qu'il y a extension de certains processus cognitifs de l'agent cognitif dans (ou vers) les institutions cognitives. L'extension de la cognition, dans cette perspective, se fait de *l'intérieur* de l'individu vers l'extérieur de celui-ci (dans l'institution cognitive). Mais on peut aussi concevoir l'extension de

⁷⁰ Je présente plus en détail à l'annexe B ce qu'est un système dynamique.

la cognition selon un mouvement inverse. En effet, on peut affirmer que, depuis *l'extérieur*, les processus cognitifs et les pratiques cognitives inhérentes aux institutions cognitives entrent à *l'intérieur* du cerveau des agents cognitifs. C'est ce que soutient Gallagher:

[L]a cognition ne s'étend pas simplement du cerveau vers l'extérieur pour incorporer des outils, des technologies et des institutions, [...] elle fonctionne [aussi] parfois de l'extérieur vers l'intérieur; les outils, les technologies et les institutions façonnent souvent nos processus cognitifs, font fonctionner notre cerveau de certaines manières et peuvent même provoquer des changements plastiques dans la structure neuronale. (Gallagher, 2013, p. 7)

Les institutions cognitives sont très nombreuses dans ce que l'on pourrait appeler le monde des significations humaines. Parmi elles, on retrouve, entre autres, l'institution de la science (Slaby et Gallagher, 2015), l'institution des genres féminin et masculin (Merritt, 2013), l'institution économique (Petracca et Gallagher, 2020) et, à mon avis, l'institution psychiatrique⁷¹.

Pour illustrer concrètement ce qu'est une institution cognitive, Gallagher a souvent utilisé le cas du système légal (Gallagher, 2011, 2013; Gallagher et Crisafi, 2009; Slaby et Gallagher, 2015, p. 36-37). Comme toutes les institutions cognitives, ce dernier ne doit pas être conçu comme une entité abstraite, mais plutôt comme un ensemble de *pratiques* cognitives qui le constitue (Slaby et Gallagher, 2015, p. 36). Ces pratiques cognitives sont à l'origine de ce que l'on pourrait appeler des *produits cognitifs* parmi lesquels on retrouve, par exemple, le contrat (mais aussi le jugement de culpabilité, la

_

⁷¹ Ici, l'expression « institution (cognitive) psychiatrique » réfère aux normes, règles, concepts, pratiques institutionnelles et savoirs que l'on trouve en psychiatrie. Elle ne réfère pas spécifiquement à l'institutionnalisation en psychiatrie, c'est-à-dire au fait de placer des gens sans leur consentement dans des instituts psychiatriques.

définition des statuts des parties, la révision des lois, etc.). Plus précisément, un contrat (comme un bail) « est l'expression (dans ce cas un accord juridique) de plusieurs esprits [qui établissent], dans une mémoire extérieure, une décision convenue [...] » (Slaby et Gallagher, 2015, p. 36). Quand on fait un contrat, celui-ci s'ajoute au système légal « qui transcende les particularités de l'esprit de tout individu » (Gallagher, 2013, p. 6). Les contrats, comme tous les produits cognitifs, « contribuent à [nos processus cognitifs] et façonnent nos processus cognitifs à l'occasion d'une réflexion plus poussée ou lors de la résolution d'un problème » (Slaby et Gallagher, 2015, p. 36). Il en est de même pour les jugements rendus au tribunal. Ceux-ci, en effet, doivent respecter un « ensemble de règles qui sont établies par le système [légal] » (Slaby et Gallagher, 2015, p. 36). Aussi, Gallagher ajoute à ce sujet que

[l]e processus par l'intermédiaire duquel les jugements sont rendus dépendra de l'engagement cognitif d'un certain nombre de personnes à l'égard d'un ensemble de lois, dont les composantes pertinentes sont mises en évidence en fonction des particularités précises de l'affaire, au fur et à mesure que la procédure se déroule. (Gallagher, 2013, p. 6)

Bref, les juges qui formulent des jugements juridiques ne formulent pas ces derniers seulement à l'aide de leur esprit. En effet, elles sont capables de formuler ces jugements parce qu'elles sont activement couplées (au sens énactif mentionné précédemment) avec le système légal sans lequel, je le répète, elles ne pourraient pas formuler de tels jugements. C'est ce que Gallagher dit (en citant Clark et Chalmers) : « [r]etirer la partie externe de ce processus cognitif — enlever l'institution légale — et "la compétence comportementale du système va s'effondrer, tout comme cela adviendrait si on lui enlevait une partie de son cerveau". » (Clark et Chalmers, 1998, p. 9, cité dans Gallagher, 2013, p. 7)

Il est important d'ajouter que le système légal se constitue lui-même grâce à l'activité cognitive des agents qui sont couplés avec lui et en fonction des problèmes qu'il aide

à résoudre. En effet, comme toutes les institutions cognitives, le système légal n'est pas un système statique, mais un système dynamique. De fait, il change avec le temps. Par exemple, la structure légale sous-jacente au système juridique reposant essentiellement sur le serment et l'ordalie opérante jusqu'au milieu du Moyen-Âge est bien différente de la structure légale sous-jacente au système accusatoire anglais que l'on voit apparaître après et où y figure un jury (dont la composition est toutefois bien différente de celle d'aujourd'hui) (Poirier *et al.*, 2019, section 3.1). De manière plus générale, on pourrait dire que le couplage entre l'agent cognitif et le système légal (ou toute autre institution cognitive) crée de la signification (Slaby et Gallagher, 2015, p. 37) et, en cela, contribue à la construction du monde de sens des êtres humains (c'est ce que dit d'ailleurs la CÉAC [section 3.1] que suppose la notion d'institution cognitive).

De ces considérations, on peut tirer au moins une leçon en psychiatrie que je formulerais ainsi : si l'esprit des personnes est constitué d'institutions cognitives, alors il se peut que certaines dysfonctions sous-jacentes à certains troubles mentaux ne se trouvent pas dans la personne, mais dans ces institutions. C'est ce qu'affirme notamment Merritt (2013) qui pense que la composante dysfonctionnelle principale du trouble « dysfonction sexuelle féminine » est l'institution du genre (féminin et masculin).

L'argument de Merritt repose sur l'idée selon laquelle l'institution du genre influe de façon générale sur la manière dont les femmes se conçoivent, pensent, socialisent et agissent dans le monde, mais aussi de manière plus particulière sur la façon dont ces dernières conçoivent et expérimentent leur vie sexuelle. Selon Merritt (2013, p. 52), la façon dont elles expérimentent leur sexualité est influencée par des représentations culturelles comme les publicités que les compagnies pharmaceutiques véhiculent dans les médias. Des représentations du même type se retrouvent aussi dans les émissions de télévision et les articles de magazines féminins. Ces représentations sont ainsi des

composantes des mécanismes cognitifs « qui soutiennent les expériences sexuelles incarnées des femmes – y compris l'image corporelle, le plaisir et l'orgasme » (Merritt, 2013, p. 51). Par exemple, une des idées que véhiculent ces représentations est que les femmes doivent s'attendre à atteindre l'orgasme à chaque épisode des rapports sexuels (Merritt, 2013, p. 52).

Merritt pense que cette institution du genre et les représentations qu'elle véhicule sur la sexualité féminine sont dysfonctionnelles. Ces représentations seraient dysfonctionnelles et responsables des symptômes du trouble « dysfonction sexuelle féminine » parce que, sans cette institution du genre, les femmes n'auraient pas en tête ces façons de se représenter ce que devrait être leur sexualité et, en conséquence, il est plausible de faire l'hypothèse que le trouble « dysfonction sexuelle féminine » n'existerait pas.

À la section 3.2.1, je montre que, contrairement à la conception du trouble mental des RDoC, la CÉOST est en mesure de saisir le fait que l'institution du genre puisse constituer ce trouble chez la femme. La CÉOST permet donc de concevoir qu'une dysfonction puisse se trouver physiquement à *l'extérieur* de l'enveloppe corporelle de la personne qui présente un trouble. En conséquence, à l'inverse de la conception du trouble mental des RDoC, la CÉOST permet de faire de la recherche sur les institutions de ce type.

2.2 Le principe épistémoquantitatif sous-jacent aux RDoC nuit à la recherche psychiatrique

Dans cette section, je défends l'idée selon laquelle les connaissances qui sont issues des études phénoménologiques et qualitatives ont une pertinence épistémique et pratique. Malheureusement, les RDoC ne saisissent pas ce type de connaissance (pour un avis semblable, voir Berenbaum, 2013, p. 897, cité dans Lilienfeld et Treadway,

2016, p. 449). Les RDoC sont aveugles à ce type de connaissances parce que le principe épistémoquantitatif qui leur est sous-jacent suppose que les phénomènes non quantifiables (et que l'on ne peut pas précisément faire correspondre à des mécanismes neuraux)⁷² n'ont pas d'intérêt épistémique (voir la section 1.5.4 pour avoir plus de détails sur le principe épistémoquantitatif).

Cette limitation de la recherche que ce principe épistémoquantitatif prévoit est empiriquement confirmée. En effet, dans le travail de Carcone et Ruocco (2017) portant sur les types de recherches qui ont été menées dans le cadre des RDoC depuis le lancement de ces derniers 73, aucune des études qui y sont citées n'est de nature qualitative ou phénoménologique. En fait, le seul type d'études empiriques qui portent sur l'expérience subjective des patientes et patients et qui semblent pouvoir être menées dans le cadre des RDoC utilise des outils psychométriques avec lesquels on peut produire des données quantitatives qui doivent être mises en lien avec des mesures de fonctionnement de mécanismes (parmi lesquelles, conformément au principe 3 des RDoC, vu à la section 1.3, doivent nécessairement figurer des mesures de mécanismes biologiques). Les propos de deux représentants des RDoC vont dans le sens de cette conjecture. De fait, ils disent que «la recherche qui repose exclusivement sur les données [issues] de rapports d'auto-évaluation ne relève pas de l'approche des RDoC » (Cuthbert et Kozak, 2013, p. 933). Un bon exemple pour illustrer ce desideratum des RDoC est donné par Carcone et Ruocco (2017) qui rapportent que Patrick et ses collègues (2013) ont montré que les résultats à l'Externalizing Spectrum Inventory (ESI) (un outil psychométrique qui permet d'évaluer quantitativement les aspects de la désinhibition chez une personne) « prédisent une variété de problèmes externalisants et

⁷² En effet, c'est ce qu'exige en principe l'hypothèse de l'identité psychoneurale de Bunge (voir section 1.5.5).

⁷³ Je parle aussi de ce travail de Carcone et Ruocco à la section 1.4 dans la note 15 en bas de page et à la section 2.1.

corrèlent avec des indices neuraux liés à une tendance à la désinhibition » (mes italiques) 74 .

Je procéderai en deux temps pour mettre en lumière ces carences épistémiques des RDoC. Premièrement, à la section 2.2.1, j'explique ce qu'est la recherche phénoménologique. Puis, je donne des exemples de ce type d'études appliquées aux troubles mentaux (Doerr-Zegers, 2018; Ratcliffe, 2015) et j'en fais connaître la pertinence épistémique et pratique.

Deuxièmement, à la section 2.2.2, j'expose ce qu'est la recherche qualitative en faisant connaître ses approches théoriques et ses méthodes de collectes de données. Ensuite, je présente les résultats d'une étude qualitative de Keyes, Noltes et Williams (2018) sur le trouble obsessionnel compulsif (TOC) et je m'en sers pour faire valoir l'intérêt de ce type d'études en psychiatrie.

2.2.1 Ce principe épistémoquantitatif empêche de faire des études phénoménologiques

La phénoménologie est « une approche philosophique qui se concentre sur les phénomènes (ce que l'on perçoit et expérimente) plutôt que sur la réalité des choses (ce qu'il y a)» (Carel, 2016, p. 19-20). Plus précisément, « la phénoménologie est l'étude des structures de la conscience telles qu'expérimentées à la première personne » (Woodruff Smith, 2018). Cette étude de la manière dont on expérimente les structures de la conscience est menée la plupart du temps individuellement. Aussi, quand on fait de la phénoménologie, « on n'essaie pas de trouver une explication causale de ces

⁷⁴ Cette recherche de Patrick et ses collègues (2013) figure dans l'annexe de l'article de Carcone et Ruocco (2017). Cette annexe contient un résumé des résultats de chacune des études que ces chercheurs ont analysées. Cette annexe n'est pas paginée. C'est pourquoi le passage que je cite n'est pas associé à une page.

expériences » (Carel, 2016, p. 21). On ne tente pas nécessairement, non plus, comme dans les RDoC, de mesurer quantitativement ces expériences. On ne fait que décrire *qualitativement* ces expériences.

Il est possible d'expérimenter plusieurs choses. Par exemple, on expérimente quelque chose lorsqu'on existe en tant qu'être humain, lorsqu'on perçoit une peinture, lorsqu'on imagine une histoire, lorsqu'on rêve, lorsqu'on pense à un projet ou lorsqu'on ressent une émotion. Chaque fois qu'on expérimente quelque chose, notre conscience *a une visée*, c'est-à-dire qu'elle est *à propos* de quelque chose. Dans le cadre de la phénoménologie, cette idée que la conscience a toujours une visée est appelée « intentionnalité ». L'intentionnalité est une structure centrale de la conscience. Outre celle-ci, les structures de la conscience qui intéressent la ou le phénoménologue diffèrent en fonction de l'expérience qui est vécue. Par exemple, lors de l'écoute d'une mélodie, ce sont les structures temporelles de rétention et de protention qui attireront son attention. Suivant Husserl (1990 [1928]), Carel définit ces structures qui sont mobilisées lors de l'écoute d'une pièce de musique de la manière suivante :

La rétention nous permet de conserver en mémoire les notes qui viennent d'être jouées [...] [et] contre lesquelles on entend le ton présent. La protention [, pour sa part,] se prolonge dans le futur et nous fait [...][entrevoir][, par l'intermédiaire] d'un sentiment d'anticipation, ce que les notes à venir pourraient être. (Carel, 2016, p. 22)

Selon le cadre phénoménologique, le fait que l'on prévoit ou que l'on ne prévoit pas ces notes « détermine, en partie, la mélodie que nous entendons » (Carel, 2016, p. 22).

Quand on écoute une mélodie, on peut aussi ressentir des émotions comme de la joie ou de la tristesse. Celles-ci font également partie de notre expérience musicale. En effet, les choses « ne nous apparaissent pas seulement comme simplement $l\hat{a}$, mais d'une manière particulière, par exemple, comme attrayantes, repoussantes, déplacées, utiles,

et ainsi de suite » (Carel, 2016, p. 20). Autrement dit, les expériences qu'on fait sont signifiantes, en ce sens qu'elles veulent dire quelque chose de personnel pour celui ou celle qui les expérimente. Bref, la phénoménologie s'intéresse aussi à ce qui est significatif pour le sujet.

Ce qui a été dit jusqu'ici montre que la recherche en phénoménologie a au moins deux objectifs. Le premier est de rendre compte de façon *individuelle*, *non causale*, *non quantifiée* et très détaillée des aspects subjectifs vécus innombrables des expériences cognitives, perceptuelles, émotionnelles et somatiques (entre autres) qui échappent la plupart du temps aux humains (parce qu'ils ne sont pas attentifs à ces aspects). Le deuxième objectif est de mettre en lumière la signification que les humains donnent (consciemment ou inconsciemment) à ces expériences subjectives.

Dans le cadre de la phénoménologie, Havi Carel (2013, 2016) a développé un modèle pour rendre compte de la *maladie*, laquelle est définie comme la pathologie *vécue* par un sujet. Le modèle phénoménologique de Carel suppose que le trouble mental⁷⁵ n'est pas seulement « un dysfonctionnement biologique à être corrigé par des experts médicaux », mais aussi « un mode de vie, une manière d'expérimenter le monde et d'interagir avec d'autres personnes » (Carel, 2013, p. 10, traduction libre). Ainsi, conformément au modèle phénoménologique, « pour comprendre totalement la maladie, il faut aussi l'étudier comme une expérience vécue » (Carel, 2016, p. 1).

Dans le cadre de son modèle phénoménologique, Carel, dans son livre *Illness* (2013), fait une analyse de la pathologie pulmonaire très grave et invalidante qu'elle s'est fait diagnostiquer en 2006. Trois choses sont, entre autres, abordées dans son ouvrage. La

⁷⁵ À noter que le modèle phénoménologique s'applique aussi aux troubles physiques. En fait, c'est pour les troubles physiques qu'il a d'abord été développé par Carel.

première concerne l'altération, chez elle, de la perception de l'espace et du temps occasionnée par sa condition. Par exemple, concernant l'espace, elle dit que « les distances augmentent, les collines deviennent des montagnes et les escaliers deviennent des obstacles plutôt que des passages » (Carel, 2013, p. 16). La deuxième chose que Carel aborde est la transformation du monde social. Elle souligne, entre autres choses, qu'une condition pulmonaire comme la sienne joue à deux niveaux sur les relations sociales : « elle réduit radicalement notre capacité à contrôler ce que les autres pensent de nous et elle place l'interaction, du moins au début, dans l'ombre de la maladie. » (Carel, 2013, p.70) La troisième chose que soulève Carel (2013, chapitre 4) est la question de la peur de la mort. En effet, quand elle est très sérieuse et éventuellement létale comme la sienne, la maladie rend cette question particulièrement concrète.

À l'instar des troubles physiques, les troubles mentaux ont eux aussi une phénoménologie particulière. On peut illustrer cela à l'aide de deux exemples.

Le premier provient d'un article d'Otto Doerr-Zegers (2018). Celui-ci s'est intéressé à la phénoménologie de l'espace et du temps chez les personnes présentant un trouble obsessionnel compulsif (TOC). Il soutient que, chez ces dernières, l'espace est étroit, oppressif et invasif (Doerr-Zegers, 2018, p. 34). En effet, cet espace est rempli par des objets pour lesquels la personne obsessive donne une signification menaçante qui ne concorde pas avec la signification que ces objets devraient avoir normalement. C'est pourquoi la personne présentant un TOC, pour se protéger d'un malheur ou de quelque chose de dangereux, n'a d'attention que pour ces objets. Les croyances obsessives qu'entretient la personne la paralysent et la restreignent dans ses actions. Ces croyances font aussi que la personne présentant un TOC habite un monde, non seulement dont l'espace se rétrécit, mais qui, aussi, ne correspond pas au monde réel. Par exemple, une personne qui a peur, de manière obsessionnelle, de contracter une pathologie se transmettant par le sang développera un ensemble de croyances irrationnelles en regard

de certains objets susceptibles d'être tachés de sang ou en regard de lieux où pourraient se trouver de tels objets. Elle tentera d'éviter ces objets ou développera des comportements compulsifs à leur égard. Par exemple, elle pourrait être très vigilante si elle marche dans un endroit où pourraient se trouver, selon elle, des seringues souillées de sang contaminé par le virus de l'immunodéficience humaine (VIH-1) ou par l'hépatite C. Ces croyances qu'entretient la personne à propos de la dangerosité de ces objets pourraient même l'amener à éviter ces endroits (les parcs, les trottoirs dans certains endroits de la ville, etc.). Bref, chez les personnes présentant un TOC, l'espace se rétrécit et est menaçant.

Par ailleurs, chez la personne présentant un TOC, le temps, comme le dit Doerr-Zegers (2018, p. 35-36), se déploie de manière circulaire plutôt que de manière linéaire. Autrement dit, le temps tourne en rond et revient sur lui-même. Aussi, ce temps n'est pas fécond, mais stérile parce qu'il se répète et crée du même ou, autrement dit, du semblable. Pour illustrer cela, on peut prendre le cas du monologue intérieur d'une personne qui veut contrôler obsessionnellement ses pensées, qu'elle juge mauvaises ou impures, en se lavant selon une méthode stricte :

Enfin, c'est la douche. Donc, le gant de toilette, le savon. C'est parti. Je me lave une première fois entièrement. Puis une deuxième, et ce, jusqu'à plus de 16 fois d'affilée. Pour que je m'arrête, il faut toujours que je n'aie eu aucune pensée obsédante pendant mon dernier savonnage et rinçage. Entre chaque, j'ai du mal à me lever pour souffler : j'ai la plupart du temps le dos plié vers le robinet. Il faut que je tienne! C'est la fin de la journée. C'est la dernière grosse étape... Ah non! Il me reste encore le lavage des cheveux... Et encore une pensée angoissante... Si seulement je pouvais les anéantir... Il faut que je recommence le lavage... Concentration : il faut que je me lave plus vite que la dernière fois. Vite, vite, le rinçage! Le séchage... Ouf, je peux m'habiller, du moins en bas, car il faut encore que je me lave les cheveux. Et c'est reparti comme pour la douche : plié en deux au-dessus de la baignoire, je me rince les cheveux, passe le shampoing, me frotte et me rince. Ceci entre 8 et 16 fois avec toujours cette même condition pour m'arrêter : pas de pensée « interdite » pendant le dernier lavage jusqu'au

séchage sinon il faut recommencer. Je compte : j'en suis à la 19^e fois. Mes yeux me piquent, car je rince mal pour gagner du temps sur les pensées.

Finalement, à la 20^e fois, dos courbé, j'attrape une serviette, la 5^e depuis le début de mon rituel. Je me sèche les cheveux rapidement, j'essaye de me redresser, la douleur est assez vive, j'arrive avec grande difficulté à me remettre droit. J'ai les yeux qui me piquent et je me sens très fatigué. Je regarde ma montre : ça fait 2 heures que j'y suis. (Neveu, 2005, p. 22-23)

Comme on peut le constater à la lumière de ce cas, « chez tous les patients obsessifs, nous trouvons ce retardement du temps de la vie, cette incapacité à avancer [...][et] ce roulement vain autour d'objets concrets et immanents » (Doerr-Zegers, 2018, p. 37) qui sont tous causés par des pensées obsessives et irrationnelles.

Le deuxième exemple qui montre que les troubles mentaux ont une phénoménologie particulière provient des travaux de Matthew Ratcliffe (2015) qui s'intéresse à la dépression majeure (DM)⁷⁶. Ce dernier explique qu'une personne présentant une DM expérimente une perturbation du sentiment d'être confortablement immergé dans le monde (2015, p. 15). Il dit aussi qu'en raison de cette perturbation cette personne habite un monde phénoménologiquement différent de celui qu'elle habite en temps normal (2015, p. 15). Ainsi, l'expérience du monde de la personne déprimée est altérée. Par exemple, un sentiment de désespoir teinte cette expérience existentielle qu'est la DM (2015, p. 11). Selon Ratcliffe, deux types de désespoir peuvent être distingués chez la

⁷⁶ J'aimerais mentionner que cette analyse phénoménologique de la dépression majeure que fait Ratcliffe est au fond une étude empirique qualitative où est adoptée une *approche* phénoménologique (voir section 2.2.2). En effet, pour faire son analyse phénoménologique, Ratcliffe utilise, entre autres, les réponses qu'ont fournies plus de 145 personnes à un questionnaire visant à sonder l'expérience à la première personne de la dépression majeure (Ratcliffe, 2015, p. 27). J'aurais donc pu utiliser ce travail de Ratcliffe sur la DM à la section 2.2.2 où je présente ce que sont les études empiriques qualitatives. Néanmoins, je pense que l'analyse de la DM que fait Ratcliffe a sa place dans la section 2.2.1 parce qu'elle est bel et bien une analyse phénoménologique. Je pense que le fait que les résultats de l'étude de Ratcliffe conviennent autant à la section 2.2.1 qu'à la section 2.2.2 montre qu'il arrive quelquefois que se chevauchent le contenu des études qualitatives et celui des analyses phénoménologiques.

personne déprimée. Le premier type de désespoir peut être exprimé par un énoncé qui exprime une attitude propositionnelle. L'objet du désespoir est alors quelque chose de précis comme dans « J'ai perdu l'espoir d'exercer un jour un métier que j'aime ». Le deuxième type de désespoir ne peut pas être exprimé par un énoncé qui exprime une attitude propositionnelle. Ce type de désespoir est plus fondamental. Ratcliffe le qualifie d'« existentiel ». Selon lui, c'est un désespoir existentiel qu'expérimentent la plupart du temps les personnes présentant une dépression majeure (Ratcliffe, 2015, p. 99). Ce type de désespoir peut être généralement compris comme une perte de la phénoménologie qu'accompagne ce que Ratcliffe appelle l'espoir existentiel (Ratcliffe, 2015, p. 103-110). L'espoir existentiel est

[...] une toile de fond phénoménologique sur laquelle des états du type « j'espère que p » sont possibles. Elle [la toile de fond] peut — en principe — survivre à la perte de tous les espoirs intentionnels : on peut perdre « tous les espoirs » sans perdre « tout espoir ». (Ratcliffe, 2015, p. 103)

Cette perte ou même cette absence totale d'espoir existentiel qu'expérimentent les personnes déprimées présente une phénoménologie particulière qui a, entre autres, trois caractéristiques. La première d'entre elles est que le monde perçu par ces dernières est « dépourvu de potentiel pour certains types de changements significatifs, et donc [...] semble en quelque sorte différent, diminué » (Ratcliffe, 2015, p. 111). Pour illustrer cela, Ratcliffe cite une patiente ou un patient parmi celles et ceux qu'il a questionnés :

#23. Le monde a l'air très différent lorsque je suis déprimé.e, car tout semble sombre/noir et morne. Pour moi, c'est comme si la couleur et la joie avaient été aspirées du monde et que le monde était complètement terne.(Ratcliffe, 2015, p. 111)

Vivre dans un monde où le champ des possibles est substantiellement diminué est une deuxième caractéristique de cette phénoménologie du désespoir. En effet, comme l'explique Ratcliffe, chez les personnes déprimées, « [c]ertains types de possibilités ne

peuvent plus être imaginés [...] et sont inaccessibles à la pensée plus généralement, disparus du monde dans lequel [ils] résident [...] » (Ratcliffe, 2015, p. 111). Pour illustrer cela, Ratcliffe cite une autre patiente ou un autre patient parmi celles et ceux qu'il a questionnés :

#22. La vie semble complètement inutile quand on est déprimé. La dépression est le pire sentiment au monde et lorsque vous êtes absorbé dans ses profondeurs, vous ne voulez même pas être là, tout pour arrêter l'engourdissement et la douleur. Vous ne pouvez pas voir loin dans l'avenir, donc vous ne pouvez pas voir vos aspirations ou vos rêves. Tout ce que j'avais toujours voulu faire dans ma vie avant me semblait impossible à ce moment. Je pensais aussi que je ne pourrais jamais m'en sortir, que je serais déprimée pour toujours. Cela entraîne des pensées tout à fait irrationnelles, car ce n'est pas une maladie rationnelle. Elle vous fait penser toutes sortes de choses sur la vie et sur vous-même qui ne sont pas vraies. Je pensais que je ne sortirais jamais des profondeurs de la dépression et que je n'accomplirais jamais rien de ma vie. (Ratcliffe, 2015, p. 111)

Une troisième caractéristique de cette phénoménologie du désespoir existentiel est que le monde apparait comme dangereux (Ratcliffe, 2015, p. 111). Voici deux témoignages tirés de l'étude de Ratcliffe qui illustrent bien cela :

#14. Le monde semble être rempli de plus de haine, de mal et de peur. Quand je suis déprimé, le monde est un endroit vraiment horrible.

«#34. Le problème de la dépression, c'est qu'on perd l'espoir et qu'on devient très autodestructeur. Je trouve aussi que le monde devient un endroit sombre et dangereux et je suis incapable d'y trouver de la joie ou du bonheur. (Ratcliffe, 2015, p. 112-113)

Cette phénoménologie du désespoir, qu'expérimentent les gens déprimés décrite par Ratcliffe, peut éventuellement s'accompagner de ce que Benson, Gibson et Brand (2013) appellent le « sentiment d'être suicidaire ». Suivant ces derniers, Ratcliffe explique que ce sentiment ne peut pas être conçu comme un espoir ou un désir de mort

chez un individu. Autrement dit, ce sentiment ne peut pas être traduit en une attitude propositionnelle. C'est ce qu'affirme Ratcliffe :

Le sentiment suicidaire [...] ne consiste pas en un système de croyances, d'émotions ou d'intentions, du moins pas dans un sens familier. La forme même de la motivation a été altérée. Parfois, la personne déprimée ne se perçoit plus ou ne se pense plus en tant qu'agent d'une manière que la plupart d'entre nous tiennent pour acquise. Il se produit un processus par lequel son monde est dépourvu des possibilités que les expériences d'agentivité et de choix présupposent. Elle [la personne] en est déjà tellement dépourvue qu'elle a l'impression de ne plus être pleinement vivante. Le suicide se présente comme l'étape finale d'un processus qui est déjà sur le point de s'achever [...] (Ratcliffe, 2015, p. 114)

Le témoignage suivant illustre très bien cette idée que le sentiment d'être suicidaire (et, on peut le penser, le suicide lui-même) n'est pas un désir ni un choix, mais un événement qui s'inscrit dans un processus naturel :

Vous êtes tenté de mettre fin à la souffrance — mais ce n'est pas une décision calculée (de mettre fin à la souffrance); c'est le sentiment que c'est la prochaine étape naturelle à franchir (tout comme les animaux cherchent la solitude pour mourir) (#117). (Ratcliffe, 2015, p. 114)

Ces études phénoménologiques ont un intérêt pratique. Par exemple, elles peuvent aider à mieux comprendre et à mieux détecter « les expériences de dépression associées à des risques de suicide particulièrement élevés » (Ratcliffe, 2015, p. 115). Par conséquent, elles peuvent aider à empêcher qu'un suicide se produise.

Ces études phénoménologiques ont aussi un intérêt épistémique. Par exemple, elles pourraient aider à raffiner les catégories diagnostiques des différents troubles mentaux que l'on retrouve dans les classifications psychiatriques du moment et, par le fait même, elles pourraient influencer la pratique clinique. Par exemple, Ratcliffe affirme que « [1]es pratiques diagnostiques [relatives à la dépression majeure] à l'heure actuelle ne

sont pas guidées par une distinction explicite entre la perte d'espoir existentiel et la perte d'espoir intentionnel » (Ratcliffe, 2015, p. 115). Je crois qu'il serait pertinent de faire cette distinction puisque les gens dépressifs expérimentent plutôt, si l'on suit Ratcliffe, un désespoir existentiel.

En vertu du principe épistémoquantitatif sur lequel reposent les RDoC, ces derniers ne permettent pas le type d'analyse des troubles mentaux et physiques que font Carel (2016, 2013), Doerr-Zegers (2018) et Ratcliffe (2015). En conséquence, les RDoC ne saisissent pas les propriétés phénoménologiques fines et très variées des troubles mentaux. Cela est malheureux parce que plusieurs chercheuses et chercheurs supposent qu'en les mettant en lumière et en les analysant, on peut collecter des données qui ont une valeur épistémique et pratique (comme on l'a vu précédemment).

2.2.2 Ce principe épistémoquantitatif empêche de faire des études empiriques qualitatives

La recherche qualitative poursuit grosso modo deux objectifs. Le premier objectif est d'expliquer «[...] certains (aspects de) phénomènes sociaux [...]» (Kohn et Christiaens, 2014, p. 68). Le deuxième objectif est de comprendre les diverses expériences personnelles des individus (Kohn et Christiaens, 2014, p. 68). Ce second objectif montre que la recherche qualitative présente une ressemblance avec la phénoménologie décrite précédemment qui tient pour fondamental les expériences que vivent les humains et le sens que donnent ces derniers à ces dernières (voir section 2.2.1). Cependant, contrairement à la recherche phénoménologique qui est, en principe, une entreprise individuelle où l'on ne cherche pas à comprendre les mécanismes causaux sous-jacents aux phénomènes, la recherche qualitative est une enquête empirique souvent collective où sont, d'une part, utilisées diverses *approches* (que l'on pourrait définir comme des postures théoriques) pour appréhender ces expériences individuelles et ces phénomènes sociaux et où sont, d'autre part, mises à

profit des *méthodes* pour recueillir des données principalement (mais pas exclusivement) qualitatives (Saldana, 2011, p. 3) avec lesquelles sont éventuellement construites des connaissances et des concepts (que l'on peut utiliser, si on le souhaite, pour formuler des explications causales)⁷⁷.

Les *approches* utilisées en recherches qualitatives sont nombreuses. Selon Saldana (2011, p. 4-20), les plus utilisées sont l'ethnographie, la théorie ancrée, la phénoménologie, l'étude de cas, l'analyse de contenu, les méthodes mixtes (qualitatives *et* quantitatives) de recherche, l'enquête narrative, l'enquête poétique, la recherche basée sur les arts, l'autoethnographie, la recherche évaluative, la rechercheaction, le journalisme d'enquête et l'enquête critique. On peut dire au moins deux choses à propos de ces approches.

La première chose est que quelques-unes d'entre elles (mais pas toutes) peuvent être combinées. L'ethnographie et l'étude de cas, par exemple, peuvent l'être (Saldana, 2011, p. 4). En effet, analyser en profondeur une unité quelconque comme une organisation (autrement dit, faire une étude de cas) et observer et prendre des notes sur la vie sociale ayant lieu dans cette unité afin de rendre compte de sa culture (autrement dit, faire de l'ethnographie) sont des démarches compatibles. Par exemple, Barrett, dans une étude qu'il a menée au cours des années 1980, dans un hôpital psychiatrique d'Adélaïde, a combiné ces deux approches. Plus précisément, il s'est intéressé à l'institution psychiatrique, aux idées et pratiques de ses membres, à la manière dont ces derniers définissent ce qu'est une personne, à la façon dont ces derniers définissent un

⁷⁷ La pertinence de l'élaboration d'explications causales en recherche qualitative est défendue, par exemple, par Maxwell (2004, 2012). Cependant, je n'aborderai pas ce sujet dans ma thèse.

cas psychiatrique et, plus précisément, à la manière dont ces derniers conçoivent ce que c'est que d'avoir une schizophrénie (Barrett, 1996, p. 272).

La deuxième chose à noter est que le choix de l'approche dépendra de la question de recherche et des buts des chercheuses et chercheurs. Par exemple, il convient d'adopter une approche phénoménologique «[...] lorsque le *but* est de parvenir à une conscience intime et à une compréhension profonde de la manière dont les humains vivent quelque chose » (Saldana, 2011, p. 8, mes *italiques*).

C'est ce qu'ont notamment fait Stanghellini et ses collaborateurs pour saisir les caractéristiques distinctives de l'expérience anormale du temps dans le trouble de la dépression majeure (2017). Ceux-ci ont effectué une étude qualitative rétrospective d'un corpus de 550 dossiers cliniques de patients rencontrés entre 1979 et 1993. Pendant cette période, le contenu de ces dossiers a été construit à partir d'une entrevue semi-dirigée d'environ 90 minutes menée par l'un des chercheurs de l'étude (qui est un psychiatre expérimenté). Les questions posées étaient, entre autres, les suivantes :

« Parlez-moi s'il vous plaît de votre expérience du temps »; « Pensez à une expérience, une période de votre vie où vous avez été particulièrement conscient du temps et parlez en moi »; « Avez-vous fait l'expérience d'une certaine étrangeté dans l'écoulement du temps, par exemple, dans la durée du temps? »; « Êtes-vous plus concentré sur le présent, le passé ou le futur? Ressentez-vous la vitesse du temps comme accélérée ou décélérée? Ressentez-vous que votre expérience temporelle diverge d'une manière ou d'une autre de l'expérience temporelle du sens commun? (Stanghellini *et al.*, 2017, p. 6)

Ce sont donc 550 dossiers cliniques qui ont été développés à partir de ces questions. Suivant certains critères (présence de trauma à la tête, histoire de dépendance à une substance, etc.), l'équipe de Stanghellini a écarté 149 dossiers de l'étude (2017, p. 4). Ensuite, elle a opéré, selon les critères du DSM-5, une réévaluation des 401 dossiers qui restaient. Il a été conclu que 100 patients présentaient un diagnostic de dépression

majeure et que 301 avaient un diagnostic de schizophrénie (2017, p. 4). De ces 100 patients dépressifs, l'équipe de Stanghellini a jugé que 96 d'entre eux présentaient une expérience anormale du temps comportant trois caractéristiques : (1) les patients dépressifs ont la sensation que leurs fonctions vitales sont ralenties (2017, p. 8-9), (2) ils ont de la difficulté à donner du sens à leur vie et à se projeter dans l'avenir, ce qui engendre chez eux un « profond et horrifiant sentiment d'incapacité » (2017, p. 10) et (3) ils sont « préoccupés par les événements passés et moins concentrés sur les événements futurs ou présents » (2017, p. 11). Plus précisément, pour ces patients,

[l]'histoire est vécue dans son irrévocabilité absolue, le passé comme une culpabilité impardonnable, le futur comme une catastrophe inévitable et le présent comme une ruine irréparable. (Stanghellini *et al.*, 2017, p. 12)

Pour montrer que le choix de l'approche dépend de la question de recherche et des buts des chercheurs, on peut aussi prendre l'exemple de l'utilisation des approches mixtes. En effet, il est pertinent d'adopter ces dernières lorsqu'un des objectifs est de corroborer les données quantitatives avec une approche qualitative. Pour illustrer cela, voici un exemple fictif d'un sondage sur la perception qu'ont les patients de la qualité des soins qu'ils reçoivent de la part du personnel soignant travaillant dans une unité psychiatrique⁷⁸ (voir Tableau 2.1).

⁷⁸ Je crée cet exemple fictif à partir d'un exemple, lui aussi fictif, créé par Saldana (2011, p. 11) et qui porte, quant à lui, sur la perception qu'ont les gens de l'efficacité de leur gouvernement.

Tableau 2.1 Sondage fictif sur la perception qu'ont les patients de la qualité des soins qu'ils reçoivent de la part du personnel soignant travaillant dans une unité psychiatrique (inspiré de Saldana, 2011, p. 11)

Énoncés				
J'ai le sentiment que le personnel soignant de l'unité de soin psychiatrique comprend ma détresse mentale.	1	2	3	4
Les médicaments que je prends chaque jour allègent ma détresse mentale.	1	2	3	4
La/le psychiatre qui me suit comprend ma détresse mentale.	1	2	3	4
Légende : 1 = Tout à fait d'accord 3 = Pas d'accord		2 = D'accord 4 = Pas du tout d'accord		

Lors de cette étude fictive, dans un premier temps, on demande aux participants d'encercler le chiffre correspondant à leur évaluation quantitative du contenu de chacun des énoncés (Saldana, 2011, p. 11). Après, pour chacun de ces derniers, on calcule une moyenne. Si des données démographiques (genre, ethnicité, âge, etc.) ont été recueillies pour chaque participant, on pourrait les mettre en lien avec les moyennes obtenues (Saldana, 2011, p. 11).

Dans un deuxième temps — et c'est à ce moment qu'intervient l'approche qualitative —, on sélectionne un échantillon dans le groupe des participants au sondage et on fait passer à ses membres une entrevue individuelle *afin de valider les évaluations quantitatives* qu'ils ont fournies. Par exemple, l'évaluation qu'ils ont donnée pour l'énoncé « La/le psychiatre qui me suit comprend ma détresse mentale » pourrait être corroborée ou invalidée par la réponse qu'ils donneront pour la question suivante : « Quelles perceptions avez-vous de la/du psychiatre qui vous suit? » (Saldana, 2011, p.

11). Bref, l'analyse des réponses données à cette question ouverte pourrait aider la chercheuse ou le chercheur à s'assurer de la justesse des résultats quantitatifs de l'étude.

Maintenant que les approches ont été présentées, on peut aborder les *méthodes* qualitatives de cueillette de données. On retrouve, parmi elles, les entrevues individuelles de recherche (souvent semi-dirigées), le groupe de discussion, l'observation (directe, participante, etc.), les sondages, l'histoire de vie et l'analyse de documents, d'objets culturels et de matériels médiatiques (comme un site Internet). Comme on peut le constater, certaines méthodes supposent une interaction directe avec des personnes (comme l'entrevue individuelle) et d'autres ne supposent pas de telles interactions (comme l'analyse d'objets culturels).

Le choix et le nombre de méthodes utilisées lors d'une enquête qualitative dépendront, entre autres choses, du but de la recherche et de la question que se posent les chercheuses et les chercheurs (Saldana, 2011, p. 31-32).

Faute d'espace, on ne peut pas ici expliquer ce que sont toutes ces méthodes de cueillette de données. Mais on peut mettre à profit un exemple concret pour en définir quelques-unes et pour faire connaître les raisons qui expliquent leur utilisation.

Il convient d'utiliser à cet effet la recherche empirique qualitative de Marlène Pominville-Racette (2018) qui est d'autant plus intéressante qu'elle présente, comme on le verra un peu plus loin, un intérêt certain pour le champ de la psychiatrie pris au sens large.

Pominville-Racette a fait une étude de cas portant sur l'impact de la démarche d'empowerment de l'Épicerie solidaire de Rosemont (ÉSR) sur le niveau de pauvreté de ses membres et de ses bénévoles. Cette chercheuse définit la pauvreté comme un

processus de disqualification sociale (Pominville-Racette, 2018, p. 36). Elle suppose également que la pauvreté a des aspects psychologique, social, économique et politique.

Au terme de son étude de cas, Pominville-Racette a mis en lumière cinq découvertes relatives à l'impact de la démarche d'empowerment de l'ÉSR sur le niveau de pauvreté de ses membres et de ses bénévoles. Premièrement, elle a trouvé que l'ÉSR fait baisser le niveau de pauvreté *psychologique* de ces derniers. En effet, l'ÉSR a un impact positif sur l'estime de soi de ses membres et spécialement sur les bénévoles qui y travaillent. Comme elle le remarque, l'ÉSR « valorise [ces derniers], leur permet de se sentir utiles et leur offre l'opportunité de redéfinir positivement leur statut social » (Pominville-Racette, 2018, p. 167). Deuxièmement, cette chercheuse a découvert que l'ÉSR fait diminuer le niveau de pauvreté *sociale* de ses membres et bénévoles. L'ÉSR permet, en effet, à ses membres et aux bénévoles qui y travaillent de créer du lien social et de briser l'isolement. Cette socialisation des membres et des bénévoles est favorisée parce qu'à l'ÉSR, ceux-ci se sentent acceptés « tels qu'ils le sont, avec leurs limites » (Pominville-Racette, 2018, p. 168). La troisième chose que Pominville-Racette a découverte est que l'ÉSR fait très peu diminuer le niveau de pauvreté *économique* de ses membres et bénévoles. De fait, cette dernière dit que

[b]ien que l'ÉSR améliore la situation économique de certains bénévoles, elle ne leur permet pas en revanche de sortir de la pauvreté. Notamment, celle-ci n'a pas eu d'effet sur l'intégration professionnelle pour les personnes ayant participé à notre étude. (Pominville-Racette, 2018, p. 166)

Quatrièmement, Pominville-Racette a découvert que l'ÉSR fait baisser le niveau de pauvreté *politique* de ses membres et bénévoles. Cette affirmation est confortée par le fait que les bénévoles participent de manière importante aux décisions concernant l'ÉSR (relativement au choix des produits et des activités qui y sont menées, par exemple) (Pominville-Racette, 2018, p. 117). Cette participation donne donc aux bénévoles un certain pouvoir décisionnel au sein de l'ÉSR. Cependant, et c'est la

cinquième découverte de Pominville-Racette, dans le quartier et la société en général, l'impact qu'a l'ÉSR sur le niveau de pauvreté politique de ses membres et de ses bénévoles est très faible. Comme le dit Pominville-Racette, « [t]out ce que l'on peut dire est que ce processus d'empowerment [que promeut l'ÉSR] est reconnu et soutenu par le milieu où il se trouve » (Pominville-Racette, 2018, p. 169) et probablement pas par la société en général.

Afin de mieux les connaître, il convient de s'intéresser aux trois méthodes de cueillette de données utilisées par Pominville-Racette (2018) lors de sa recherche, à savoir l'observation participante, l'entrevue de recherche et le groupe de discussion.

L'observation participante ⁷⁹ est une méthode qui a été à l'origine utilisée lors de recherches ethnographiques et anthropologiques menées dans des pays étrangers. Elle servait à recueillir des données qualitatives détaillées sur la culture et le mode de vie des gens (Saldana, 2011, p. 46). On utilise aujourd'hui l'observation participante pour étudier d'autres endroits comme une salle de classe d'une école ou une organisation dans une société (Saldana, 2011, p. 46). Dans ces lieux, lorsqu'elle observe et prend des notes, la chercheuse essaie de « saisir les actions, les réactions et les interactions

On ne trouve pas spontanément de tels groupes dans la société, mais on s'approche de cette forme et de ce fonctionnement dans les groupes expérimentaux où l'observateur [(ou l'observatrice)] a un rôle formel comme dans les situations créées dans les laboratoires de dynamique de groupe. (Junker, 1960, cité dans Peretz, 2004, p. 49).

⁷⁹ Le terme « participante », dans l'expression « observation participante », réfère au « mode de présence du chercheur au sein du milieu observé » (Peretz, 2004, p. 49). Le type d'observation participante variera en fonction du niveau de présence de l'observatrice ou de l'observateur. Junker (1960) que cite Peretz (2004, p. 50) a formulé quatre types d'observation participante. À une extrémité du continuum où se situent ces modes de présence, l'observatrice (ou l'observateur) n'est pas connue des individus qui évoluent dans le milieu qu'elle observe (cela lui permet, entre autres, d'avoir accès à des informations qui ne lui auraient pas été accessibles si elle avait été connue de ces individus). À l'autre extrémité du continuum, non seulement l'observatrice (ou l'observateur) est connue des individus du milieu qu'elle étudie, mais les rôles de chacun (y compris celui de l'observatrice) sont clairement définis :

des gens dans leur milieu naturel et d'en déduire leurs modes de pensée et leurs sentiments » (Saldana, 2011, p. 46). Par exemple, l'observation participante a permis à Pominville-Racette (2018) de comprendre la dynamique des acteurs évoluant autour et dans l'ÉSR. Mais ce n'est pas tout. En effet, en plus de permettre la cueillette de données, l'observation participante a d'autres fonctions. L'une d'entre elles est de permettre la création d'un lien de confiance avec les personnes qui participent à l'étude (Pominville-Racette, 2018, p. 165). Une autre fonction est d'être utile « pour préciser, exemplifier et valider les informations obtenues lors des entrevues [individuelles semistructurées] » (Pominville-Racette, 2018, p. 165).

L'entrevue de recherche peut être définie simplement comme « une interaction verbale, une conversation entre un interviewer [...][qui est la chercheuse ou le chercheur] et un interviewé » (Savoie-Zajc, 2009, p. 338). On peut adopter une interprétation béhavioriste ou constructiviste de cette définition (Savoie-Zajc, 2009, p. 338-339). Selon la première, « le chercheur est vu comme celui qui émet un stimulus, en général sous la forme d'une question, à l'interviewé qui y réagit en retour par la production d'un nouveau stimulus » (Savoie-Zajc, 2009, p. 338-339). L'interprétation béhavioriste suppose également que la connaissance existe en soi et que le chercheur doit la découvrir lors de l'entrevue (Savoie-Zajc, 2009, p. 339). À l'inverse, selon l'interprétation constructiviste, la connaissance est plutôt construite par le chercheur *et* l'interviewé en ce sens que le

chercheur et [le] participant construisent un « texte », une « narration » constituée d'histoires, empreintes des connotations personnelles, interpersonnelles, sociales et culturelles des individus en présence. (Savoie-Zajc, 2009, p. 339)

Autrement dit, selon cette perspective, la connaissance est conçue comme « une construction interpersonnelle, un produit de la rencontre des personnes engagées dans la relation » (Savoie-Zajc, 2009, p. 339)⁸⁰.

Il existe plusieurs types d'entrevues de recherche. Parmi elles, on retrouve l'entrevue structurée⁸¹, l'entrevue non structurée⁸² et, à mi-chemin entre ces dernières, l'entrevue semi-structurée. Lors de sa recherche, Pominville-Racette a utilisé l'entrevue semi-structurée. Cette dernière présente une structure que l'on pourrait qualifier de malléable. Savoie-Zajc qui défend une conception constructionniste de ce type d'entretien adopte la définition suivante :

L'entrevue semi-dirigée consiste en une interaction verbale animée de façon souple par le chercheur. Celui-ci se laissera guider par le rythme et le contenu unique de l'échange dans le but d'aborder, sur un mode qui ressemble à celui de la conversation, les thèmes généraux qu'il souhaite explorer avec le participant à la recherche. Grâce à cette interaction, une compréhension riche du phénomène à l'étude sera construite conjointement avec l'interviewé. (Savoie-Zajc, 2009, p. 340)

⁸⁰ À noter qu'une telle élaboration intersubjective de la connaissance n'a pas seulement lieu entre une chercheuse ou un chercheur et la personne qu'elle ou il interviewe. En fait, elle peut potentiellement avoir lieu chaque fois que deux personnes (ou plus) s'engagent dans une relation de communication. Par exemple, Potter, Gupta et Goyer (2019) soutiennent qu'une telle co-construction de connaissances s'opère entre une psychiatre (ou un psychiatre) et sa patiente (ou son patient) lorsque la première raisonne afin de poser un diagnostic sur la seconde.

⁸¹ Ce type d'entrevue contient des questions préparées en fonction de la question de recherche que la chercheuse ou le chercheur doit absolument poser, et ce, idéalement, dans l'ordre où elles apparaissent sur son questionnaire. Bref, dans cette perspective, « l'interaction verbale est produite dans les limites d'un questionnaire administré oralement » (Savoie-Zajc, 2009, p. 340).

⁸² L'entrevue non dirigée est, dans sa forme, symétriquement différente de l'entrevue dirigée. En effet, un questionnaire pour une entrevue non dirigée contient tout au plus des thématiques qui peuvent potentiellement être abordées et explorées lors de l'entretien. L'interviewé a, dans cette situation, carte blanche sur la manière dont il veut traiter de la thématique à l'étude. On utilise souvent ce type d'entretien pour la constitution de récits de vie (Savoie-Zajc, 2009, p. 340).

Par exemple, mener les entretiens semi-dirigés a permis à Pominville-Racette de recueillir des informations qui lui ont permis de comprendre plus en profondeur l'impact qu'avait l'ÉSR sur les aspects psychologique, social, économique et politique de la pauvreté (Pominville-Racette, 2018, p. 165). Faire ces entretiens a été aussi l'occasion pour elle de « saisir le sens que les acteurs accordent à la démarche de l'ÉSR et [celui] que les bénévoles donnent à leur engagement dans cet organisme » (Pominville-Racette, 2018, p. 165).

Enfin, le groupe de discussion est conçu comme « une forme d'entrevue de groupe qui mise sur la communication entre les participants de la recherche pour générer des données » (Kitzinger, 1995, p. 299). Un tel groupe « réunit de six à douze participants et un animateur » (Geoffrion, 2009, p. 391). La forme que prend le groupe de discussion sera modifiée en fonction des multiples objectifs de la chercheuse (Morgan, 2008, p. 352). En ce qui la concerne, Pominville-Racette (2018, p. 165) a utilisé le groupe de discussion pour valider et bonifier les données qu'elle avait récoltées lors des entretiens individuels. Comme elle avait un objectif précis, on peut penser que Pominville-Racette a influé activement sur le cours de la discussion. Pour cette raison, on peut qualifier de structuré⁸³ le groupe de discussion qu'elle a mené (Morgan, 2008, p. 354).

Parce qu'une étude de cas comme celle de Pominville-Racette (2018) porte sur *un* seul cas et que ces résultats ne sont pas nécessairement généralisables, on peut penser qu'elle sera dévalorisée par les chercheuses et chercheurs — comme celles et ceux défendant les RDoC — qui valorisent l'utilisation de *grands* échantillons (qui contiennent plusieurs cas particuliers, à savoir des personnes) et de méthodes

⁸³ Je spécifie cela parce que le groupe de discussion peut aussi être peu structuré. Quand la discussion de groupe prend cette forme, les participantes et les participants sont alors relativement libres de parler comme elles et ils le souhaitent à propos du sujet à l'étude (Morgan, 2008, p. 352). Cette forme de groupe de discussion est utilisée lors de recherches exploratoires (Morgan, 2008, p. 352 et 354).

statistiques pour produire des connaissances. Mais dénigrer l'intérêt épistémique de cette approche est, selon moi, une erreur parce qu'elle est très utile quand elle est appliquée aux phénomènes qui lui conviennent, comme l'efficacité d'une démarche d'empowerment dans une organisation quelconque ou les degrés d'adhésion à une mesure dans un groupe social particulier.

En outre, une étude qualitative comme celle de Pominville-Racette a l'intérêt épistémique de faire comprendre le phénomène de la pauvreté sous un angle nouveau. En effet, à la lecture de l'étude de cette dernière, on comprend que la pauvreté n'est pas seulement définie en termes de revenu disponible. Quand l'on comprend cela, d'ailleurs, on comprend mieux pourquoi la prévalence des troubles mentaux est plus élevée chez les personnes pauvres (Lechaume et Savard, 2019). En effet, la baisse de l'estime de soi, l'isolement, le manque de moyens pour agir et le sentiment d'impuissance face à sa situation qui caractérisent le fait d'être pauvre sont tous des facteurs qui ont un impact négatif sur la santé mentale.

En plus d'informer de manière nouvelle sur les facteurs environnementaux (comme la pauvreté) favorisant une augmentation de la prévalence des troubles mentaux, les études qualitatives peuvent aussi fournir des connaissances très utiles sur les troubles mentaux eux-mêmes. Le travail que Keyes, Noltes et Williams (2018) ont fait sur l'*expérience* du trouble obsessionnel compulsif (TOC) chez des jeunes de 13 à 18 ans montre bien cela. Ces chercheurs ont trouvé, en faisant des entrevues qualitatives semi-structurées auprès de jeunes de 14 à 17 ans présentant un TOC, plusieurs éléments de connaissance intéressants sur ce trouble et sur la manière d'intervenir sur ce dernier :

(1) Ils ont observé que le développement de comportements obsessifs et compulsifs chez les jeunes était très souvent précédé par l'expérience d'événements stressants (hostilité dans la famille, maladie et mortalité) (Keyes *et al.*, 2018, p.

- 180). D'un point de vue clinique et thérapeutique, cela signifie que l'on ne doit pas seulement intervenir sur les comportements obsessionnels et compulsifs. En effet, il faut aussi s'intéresser aux traumas vécus (Keyes *et al.*, 2018, p. 177). Plus précisément, les cliniciennes et cliniciens doivent porter leur attention sur la manière dont les jeunes donnent un sens à ces événements. Cela est d'autant plus important que ces événements sont reliés à une augmentation du sentiment de responsabilité chez ces jeunes (Keyes *et al.*, 2018, p. 182).
- (2) Ils ont trouvé que la recherche d'aide relativement aux comportements obsessifs et compulsifs des jeunes était lente à se mettre en marche. Elle varie, en effet, de trois mois à huit ans (Keyes et al., 2018, p. 180). Deux raisons expliqueraient principalement cette lenteur. La première est que le TOC et les symptômes qui le caractérisent ne sont généralement pas bien compris par plusieurs personnes. En effet, les jeunes, pendant un certain temps, pensent que leurs comportements et pensées obsessives sont normaux. Les parents de ces enfants, quant à eux, croient que les comportements obsessifs et compulsifs de ces derniers sont bizarres, mais qu'ils finiront par passer. De plus, certains jeunes interrogés affirment que leurs enseignants et même leur médecin généraliste ne comprenaient pas ce qu'était le TOC. Enfin, les gens de la société se font une idée caricaturale de ce qu'est le TOC (en effet, les gens présentant un TOC ne sont pas nécessairement des gens qui aiment que les choses soient en ordre et bien rangées). La seconde raison qui explique que la recherche d'aide psychiatrique est lente à démarrer est que les jeunes sont gênés de parler de leurs obsessions et compulsions. On peut tirer de cette étude trois motifs qui expliquent cela. D'abord, leurs obsessions et compulsions sont intimes et étranges. Ensuite, elles leur font perdre le contrôle d'eux-mêmes et de leur vie. Enfin, en raison d'elles, ils ont peur de passer pour des fous et se sentent incompris par les autres autour d'eux. Ils préfèrent donc ne pas parler de leur

problème et tardent à aller chercher de l'aide. Cela est regrettable parce que, lorsqu'une professionnelle ou un professionnel diagnostique le TOC chez les jeunes, ceux-ci disent se sentir soulagés (Keyes *et al.*, 2018, p. 181). Par-dessus tout, ces jeunes disent être heureux de pouvoir discuter, dans un cadre thérapeutique, avec d'autres personnes présentant un problème similaire au leur. Ils se sentent alors compris et vont chercher des conseils qui pourraient les aider à aller mieux. Comme le disent les auteurs, une des leçons cliniques que l'on peut tirer de leur étude est que

[l]es interventions qui aident les jeunes ayant un diagnostic de TOC à entrer en contact avec d'autres jeunes ayant des difficultés similaires devraient être considérées comme prioritaires. Il peut s'agir de groupes de soutien, mais aussi de dépliants, de brochures et de vidéos élaborés par des experts ou en collaboration avec des experts par expérience. (Keyes *et al.*, 2018, p. 177)

(3) Ces chercheurs ont trouvé que les jeunes percevaient leur vie avec le TOC comme une bataille à faire (Keyes *et al.*, 2018, p. 181). Deux raisons me semblent montrer cela. La première est que les jeunes présentant un TOC sont en constante recherche émotionnelle pour se sentir bien. Les auteurs de l'étude font l'hypothèse que cette quête omniprésente de la sensation de bien-être pourrait être reliée aux événements stressants vécus par le jeune présentant le TOC. Malheureusement, comme il est difficile pour les jeunes présentant un TOC d'atteindre cet état de bien-être, ceux-ci ont tendance à s'isoler pour faire baisser leur anxiété. La deuxième raison est que le TOC a un impact négatif sur toutes les activités des jeunes, spécialement à l'école. En effet, en raison des idéations et des comportements qu'il implique, le TOC « vole » du temps aux jeunes et génère beaucoup d'anxiété chez eux. Bref, on peut penser que les jeunes présentant un TOC perdent cette bataille.

(4) Les chercheurs ont trouvé que les jeunes présentant un TOC ont une relation ambivalente relativement à l'aide thérapeutique dont ils peuvent bénéficier (Keyes et al., 2018, p. 182). D'abord, lorsqu'ils font une demande d'aide, parce que les listes d'attente sont très longues, cela prend du temps avant qu'on leur réponde. Cela génère de la frustration chez eux et les décourage dans leur démarche. Ensuite, les jeunes jugent que les thérapies d'exposition sont problématiques. En effet, ce n'est pas suffisant, selon eux, d'être exposé aux stimuli problématiques en présence de la psychologue lors de la séance clinique. À leur avis, il faudrait trouver un moyen d'être accompagné par la psychologue ou d'autres personnes faisant partie de leur entourage à l'extérieur de la clinique (Keyes et al., 2018, p. 182). Enfin, les jeunes pensent que, à court terme, si le fait que les membres de leur famille s'adaptent à leurs obsessions les soulage, ce n'est pas une bonne chose à long terme. Les jeunes sont ainsi face à un problème : si les parents s'adaptent à leur TOC, leur anxiété baisse sur le moment; cependant, sur le long terme, cela n'est pas bon parce qu'ils ne se défont pas de leurs obsessions.

Ces connaissances qualitatives sur le TOC que fournit cette étude de Keyes, Noltes et Williams (2018) ont un intérêt épistémique, diagnostic et thérapeutique. Par exemple, la connaissance selon laquelle l'émergence de comportements obsessifs et compulsifs chez les jeunes est très souvent précédée par l'expérience d'événements stressants a un intérêt *épistémique*. C'est le cas parce qu'elle fait appréhender le TOC autrement que comme uniquement « une hyperactivité de certaines parties du cortex frontal et des ganglions de la base » (Insel, 2010, p. 47) accompagnée de pensées intrusives (comme une obsession pour les germes) et d'une envie irrépressible d'agir pour maîtriser ces dernières (comme la compulsion de se laver les mains). Ainsi, comme pour la notion de pauvreté, vue précédemment et dont il est question dans l'étude de Pominville-

Racette (2018), l'approche qualitative de Keyes, Noltes et Williams (2018) fait connaitre de nouvelles dimensions du TOC.

Ensuite, cette connaissance qualitative a un intérêt pour poser un *diagnostic* chez les jeunes présentant un TOC. En effet, si à l'origine de cette condition invalidante il y a une expérience traumatisante (Keyes *et al.*, 2018, p. 177), il faut que la clinicienne ou le clinicien s'y intéresse quand il ou elle évalue sa patiente ou son patient. De fait, l'équipe de Keyes soutient « [qu'au] cours de l'évaluation, il convient d'examiner systématiquement les traumatismes et les événements stressants de la vie, ainsi que le sens que les jeunes leur donnent. » (Keyes *et al.*, 2018, p. 182)

Enfin, cette connaissance qualitative a une valeur sur le plan *thérapeutique*. En effet, considérant qu'il y a souvent un traumatisme à l'origine du TOC chez les jeunes, afin d'intervenir efficacement, il est important que la clinicienne ou le clinicien aide la patiente ou le patient qui souffre de cette condition psychiatrique à identifier les événements déclencheurs traumatisants de l'apparition de son TOC et le récit qu'elle ou qu'il s'en fait (Keyes *et al.*, 2018, p. 182). C'est du moins ce que montre, comme l'indiquent Keyes et ses collègues, une étude récente sur les obsessions de peur de contaminations (Coughtrey *et al.*, 2013).

Je prends le temps de mentionner que ce traumatisme doit être pris en considération lors du traitement du TOC parce que ce dernier a été réduit par Insel (et par d'autres) à un trouble psychobiologique⁸⁴ qui, en conséquence, doit être prioritairement traité pharmacologiquement (par exemple, avec certains antidépresseurs) conjointement à

_

⁸⁴ Il est à noter à ce sujet que l'*International OCD Foundation*, à l'instar d'Insel, véhicule une conception strictement neurobiologique du TOC (International OCD Foundation, 2021).

une thérapie cognitive comportementale (TCC) visant la diminution de la présence, chez l'individu, des obsessions et des compulsions.

En vertu du principe épistémoquantitatif sur lequel reposent les RDoC et qui stipule que les données strictement qualitatives n'ont pas de valeur épistémique, les données que je viens de présenter sont sans intérêt et les études qui les ont générées ne sont pas désirables. Cela est malheureux parce que plusieurs chercheuses et chercheurs mènent en psychiatrie ce type d'enquêtes qui sont, comme on vient de le constater pour le TOC, d'une part, intrinsèquement pertinentes d'un point de vue épistémique et, d'autre part, utile d'un point de vue clinique et thérapeutique.

2.3 La TBS a de la difficulté à saisir la normativité dans le monde du vivant

Comme on l'a vu à la section 1.5.3, la TBS est une des bases philosophiques des RDoC. Il convient de se rappeler rapidement ce qu'elle est. Essentiellement, cette théorie suppose que les systèmes vivants sont constitués de parties (par exemple, de circuits neuraux, d'organes, de cellules, d'un cœur, etc.). Elle pose aussi que chacune de ces parties a une (ou des) fonction normale. Dans un système vivant, suivant la situation spécifique où ce dernier se trouve, une telle fonction est définie comme une contribution causale typique à deux buts, soit à ceux de la survie et de la reproduction individuelle du système vivant. Selon la TBS, un système vivant est en santé quand, selon diverses situations spécifiques, toutes les fonctions de ses parties contribuent typiquement (c'est-à-dire de la manière dont habituellement elles contribuent chez les membres de la *classe de référence* à laquelle il appartient) à sa survie et à sa reproduction. Par ailleurs, un système vivant présente une pathologie quand au moins une de ses parties présente une dysfonction. Il y a présence d'une dysfonction quand une partie ne fait pas, a de la difficulté à faire ou n'est pas disposée à faire ce qu'elle doit effectuer pour contribuer typiquement, dans une situation spécifique, à la survie et

à la reproduction du système vivant. En d'autres termes, et plus succinctement, comme l'affirme Garson, « si un trait ne peut pas faire ce que les autres occurrences de ce trait font typiquement et qui aide au succès reproducteur [fitness], alors celui-ci est dysfonctionnel » (Garson, 2019b, p. 134).

Suivant cela, on peut affirmer qu'au sein de la TBS un trouble mental est défini essentiellement comme un mécanisme biologique dysfonctionnel opérant dans un système ayant une vie mentale et appartenant au monde du vivant⁸⁵. On est donc ici en présence d'une conception, en principe, strictement naturaliste du trouble mental. Celle-ci contient, en effet, seulement une composante factuelle — la dysfonction —, laquelle réfère, selon Boorse, à une entité dans le monde ou, en termes plus techniques, à une espèce naturelle.

Je pense que la TBS ne peut pas être à la base des RDoC. En effet, elle échoue à saisir la normativité inhérente aux traits ou mécanismes dans le monde du vivant et, conséquemment, elle a de la difficulté à saisir les traits ou les mécanismes dysfonctionnels chez les humains (chose qu'un concept de trouble mental devrait être capable de faire). Quatre raisons me portent à penser cela⁸⁶.

Premièrement, la TBS est incapable de saisir la normativité de manière *objective*. Autrement dit, elle n'arrive pas à saisir, chez les systèmes vivants, suivant des critères strictement empiriques, les mécanismes dont le fonctionnement serait normal ou anormal. Comme l'a bien vu Kingma (2007), il en est ainsi parce que la classe de référence à la lumière de laquelle on détermine si le fonctionnement d'un mécanisme

⁸⁶ À noter que cette critique s'applique à la TBS qu'elle soit interprétée de manière adaptationniste ou non. Pour mon interprétation adaptationniste de la TBS, voir la section 1.5.3.

⁸⁵ La TBS s'applique aussi aux plantes (voir Boorse, 1977, p. 565).

d'un membre de cette classe est normal ou anormal est, en dernière instance, choisie suivant des considérations non empiriques. En effet, pour distinguer un état de santé d'un état pathologique, qu'est-ce qui justifie Boorse de choisir la classe de référence des systèmes vivants présentant le même sexe, le même âge et appartenant à la même espèce ? Pourquoi ne pas choisir la classe de référence des systèmes vivants présentant la même classe sociale, le même revenu et la même couleur d'yeux? Il semble que Boorse ne puisse pas justifier l'adoption de ces caractéristiques de manière strictement empirique. En effet, pour motiver l'adoption de cette classe de référence dont les membres présentent ces caractéristiques, il doit ultimement faire appel à des considérations normatives utilisées pour évaluer des connaissances théoriques d'arrière-plan (Stegenga, 2018a, p. 11). Parmi ces savoirs, il y a l'énoncé théorique selon lequel les « fonctions physiologiques diffèrent selon le sexe et l'âge » (Stegenga, 2018a, p. 10). Il y a aussi celui selon lequel les « fonctions physiologiques diffèrent en fonction de toutes sortes de caractéristiques chez différents types de personnes» (Stegenga, 2018a, p. 10). Par exemple, « [n]ous savons que les personnes atteintes de diabète de type 1 sont incapables de produire de l'insuline [...]» (Stegenga, 2018a, p. 10). Cependant, en raison des conséquences funestes qui s'en suivraient, il serait absurde et immoral — ce qui est une considération normative — d'ajouter la caractéristique « présenter un diabète de type 1 » à la classe de référence à la lumière de laquelle on évalue si un trait d'une personne fonctionne bien ou non. En effet, si on faisait cela, alors

[...] les personnes atteintes de diabète de type 1 seraient automatiquement jugées en santé parce que l'efficacité de leurs pancréas à produire de l'insuline serait comparée à celle des autres personnes atteintes du diabète de type 1 [...]. (Stegenga, 2018a, p. 10-11)

Conséquemment, elles mourraient puisque, ne les jugeant pas malades, on ne les traiterait pas. Bref, suivant ces considérations sur le concept de classe de référence, on peut conclure que la TBS est chargée de valeurs (Kingma, 2017, p. 53) et, en

conséquence, qu'elle n'est pas une théorie strictement naturaliste de la santé et de la maladie. Donc, elle ne saisit pas la normativité de manière objective dans le monde du vivant. Idéalement, il serait préférable que le concept de trouble contienne un principe qui permet d'appréhender objectivement les dysfonctions du monde du vivant⁸⁷.

Deuxièmement, Garson (2019b, p. 134) soutient que la TBS ne saisit pas les fonctions atypiques des traits ou des mécanismes et, conséquemment, qu'elle ne peut pas évaluer quand un trait ou un mécanisme présentant une fonction de ce genre opère normalement ou anormalement. Une fonction atypique est une fonction qui est rarement active chez un trait (ou un mécanisme) qui a plus d'une fonction⁸⁸.

Cet argument de Garson – à l'origine formulé par Neander (Boorse, 2002, p. 93, note de bas de page numéro 34) – suppose une définition boorséenne de la typicalité (Boorse, 2002, p. 92-93) qu'il formule de la manière suivante :

que « certaines fonctions sont rarement performées ».

⁸⁷ En effet, je pense qu'il existe, dans le monde du vivant, indépendamment de nos intérêts et de nos valeurs, des fonctions et des dysfonctions et, en conséquence, je crois que le concept de trouble doit les saisir. J'ai la conviction qu'il existe des fonctions et dysfonctions, dans le monde du vivant, parce que, conformément au cadre de l'énactivisme autopoïétique que j'adopte et présente à la section 3.1, je pense qu'une normativité objective et naturelle est opérante dans tous les systèmes vivants (et même dans les systèmes dynamiques faits de systèmes vivants comme les familles ou les sociétés comme je le soutiens à la section 3.2.1). À mon avis, au moins deux types de mécanismes rendent concret ce principe de normativité du vivant. Les premiers sont les mécanismes homéostatiques qui maintiennent un facteur à un niveau optimal dans l'organisme. Chez les êtres humains, les mécanismes responsables de la régulation de la glycémie, de la température, du sommeil et de la plasticité homéostatique illustrent bien cela. Les deuxièmes types de mécanismes que j'ai en tête sont ceux impliqués dans la guérison du corps humain. Par exemple, les mécanismes de cicatrisation semblent opérer pour ramener les tissus cutanés dans un état normal. La réorganisation neurale impliquée dans le rétablissement des fonctions cognitives qui a lieu à la suite d'un accident vasculaire cérébral semble aussi être un mécanisme de ce type.

un trait T a une fonction F lorsque la probabilité conditionnelle que T fasse F à un moment donné, étant donné qu'il contribue à la survie et à la reproduction à ce moment précis, est très élevée. (Garson, 2016, p. 70, mes *italiques*)

Cette définition de la typicalité permet d'identifier une fonction normale d'un trait qui opère *rarement*. Par exemple, elle permet d'affirmer qu'un spermatozoïde a la fonction de fertiliser l'ovule. En effet, bien qu'il soit très rare qu'un spermatozoïde fertilise un ovule, quand ce dernier fait quelque chose pour la survie et la reproduction de l'organisme, c'est ce qu'il fait typiquement (ou le plus souvent).

Comme je l'ai mentionné précédemment, cette conception de la typicalité inhérente à la TBS est incapable de saisir les fonctions atypiques. Garson articule la structure de l'argument qui soutient cette thèse de la manière suivante :

Supposons qu'un trait, T, ait deux fonctions différentes, F1 et F2. Supposons que F1 est exécutée beaucoup plus fréquemment que F2. Ainsi, la probabilité que T fasse F1, étant donné qu'il contribue au succès reproducteur [fitness], est élevée, mais la probabilité que T fasse F2, étant donné qu'il contribue au succès reproducteur [fitness], est faible. Dans ce cas, F2 ne serait pas une fonction de T. (Garson, 2016, p. 70)

Ainsi, si F2 n'est pas une fonction, alors elle n'est pas saisie par la TBS et, en conséquence, elle ne peut pas être dysfonctionnelle. Pour rendre concrète sa critique, Garson donne deux exemples. Le premier va comme suit :

Prenons l'exemple du bulbe rachidien du cerveau, lequel remplit un certain nombre de fonctions vitales. L'une de ses fonctions (F1) consiste à contribuer à la respiration, ce qu'il fait très souvent. Une autre fonction (F2) est de contribuer au réflexe pharyngé, et il l'exerce très rarement (relativement à F1). La [conception de la typicalité mentionnée précédemment] implique donc que le bulbe rachidien n'a pas pour fonction de déclencher le réflexe nauséeux [...]. (Garson, 2016, p. 70)

Garson (2019b, p. 135) donne également l'exemple de l'urètre chez le mâle qui a deux fonctions, à savoir celle de transporter l'urine (F1) et celle de transporter le sperme (et les spermatozoïdes que contient ce liquide) (F2). Parce que, selon la TBS, la fonction d'un trait, dans un système vivant, est ce que fait *typiquement* ce trait *quand il fait quelque chose* pour aider au succès reproducteur de ce système à un moment précis, on est obligé de dire que la fonction de l'urètre est uniquement de transporter l'urine (c'est ce qu'elle fait le plus souvent, *relativement* à F2, en effet). La TBS est ainsi aveugle à l'autre fonction de l'urètre qui est de transporter le sperme (ce n'est pas ce qu'elle fait le plus souvent, *relativement* à F1, en effet). En conséquence, la TBS ne peut pas saisir une dysfonction relative au transport du sperme que pourrait présenter l'urètre.

Un troisième problème de la TBS est qu'elle permet qu'une fonction d'un trait opérant anormalement (selon nos intuitions médicales) dans un système vivant qui se trouve dans une situation qui lui est préjudiciable soit jugée comme opérant normalement (Kingma, 2010). Ce problème advient parce que, comme on l'a vu à la section 1.5.3, la TBS tient compte du fait empirique incontesté qu'une fonction d'un trait opère toujours relativement à une situation spécifique. Par exemple, la fonction des plaquettes sanguines est d'être actives lorsque nos vaisseaux sanguins sont endommagés. Pour illustrer cette faille dans la TBS, Garson (2019b, p. 135) reprend l'exemple de Kingma (2010) d'un empoisonnement d'un individu au paracétamol. Dans cette situation spécifique d'empoisonnement, l'estomac de l'individu fonctionnera à un taux très bas (par exemple, il se videra de 50 % après 4 heures plutôt qu'après une ou deux heures). Notre intuition médicale dit que l'estomac de l'individu est dysfonctionnel. Cependant, selon la TBS, cet estomac n'est pas dysfonctionnel. De fait, dans une situation d'empoisonnement au paracétamol, il est statistiquement fréquent ou typique que l'estomac d'un individu fonctionne plus lentement que lorsqu'il n'est pas empoisonné au paracétamol. Cela est conforme à la TBS selon laquelle la fonction d'un trait, relativement à une situation spécifique, est sa contribution typique à la survie et au succès reproducteur [fitness] de l'organisme dans cette situation (Garson, 2019b, p. 135).

Hausman (2011) a tenté de résoudre ce problème en affirmant que, lorsque l'on veut juger du fonctionnement d'une fonction, il ne suffit pas de se demander si cette fonction contribue de manière typique au succès reproducteur de l'organisme dans une situation spécifique. En effet, il faut aussi se demander si cette fonction est capable d'opérer de manière typique pour le succès reproducteur de l'organisme dans toutes les autres situations où elle pourrait potentiellement être opérante. Si ce n'est pas le cas, alors cette fonction opère anormalement. Suivant cela, on pourrait dire qu'un estomac empoisonné au paracétamol fonctionne anormalement (ici, trop lentement) parce que s'il fonctionnait de cette manière dans d'autres situations, alors il ne contribuerait pas au succès reproducteur de l'organisme.

À première vue, ce contre-argument semble bon. Mais ce n'est pas le cas parce que si on adopte la position de Hausman, pense Garson (2019b, p. 137), «l'estomac de presque tout le monde est dysfonctionnel tout le temps ». En effet, Garson (2019b, p. 137) observe qu'il existe des situations possibles où certaines fonctions d'un individu opèreront moins bien qu'elles opèrent chez la plupart des individus, et ce, même si ces fonctions n'ont pas été conçues pour opérer dans ces situations et même si ces fonctions ne seront jamais dans ces situations. Pour illustrer cela, Garson prend l'exemple du mal de mer. Il suppose qu'il fait partie d'une minorité de personnes qui, sur les bateaux qui tanguent, ont la nausée. Pour cette raison, son estomac ne peut pas digérer les aliments à un rythme normal alors que la plupart des estomacs des autres personnes le peuvent. Cela signifie qu'il existe une situation où l'estomac de Garson ne peut pas contribuer à son succès reproducteur. En conséquence, selon les idées de Hausman, son estomac est dysfonctionnel ici et maintenant, et ce, même si Garson ne se trouve pas sur un

bateau qui tangue et ne s'y trouvera jamais (il peut, en effet, décider de ne jamais prendre des bateaux qui tanguent pour se déplacer).

Un quatrième problème de la TBS est qu'elle permet parfois d'évaluer des traits normaux statistiquement rares comme étant des dysfonctions. On peut utiliser l'exemple de Stegenga (2018b, p. 28-29) d'un homme dont la hauteur est de cinq pieds pour illustrer cet écueil que rencontre la conception de Boorse. Être de cette grandeur, pour un homme, est statistiquement peu fréquent et on peut faire l'hypothèse que cela pourrait lui rendre plus difficile la tâche de se trouver une partenaire pour se reproduire. En conséquence, selon la TBS, mesurer cinq pieds devrait être considéré comme une pathologie. Mais cette conclusion est absurde. Être de cette taille n'est pas une pathologie.

On peut aussi illustrer ce problème de la TBS avec un exemple de Cooper (2002, p. 272) modifié par Stegenga (2018b, p. 31) qui met en scène des individus aux cheveux roux vivant dans une société hostile aux individus roux. Cette couleur de cheveux est « due à un gène récessif qui entraine des niveaux élevés de pigments de phéomélanine et des niveaux faibles de pigments d'eumélanine » (Stegenga, 2018b, p. 30). Avoir les cheveux roux est statistiquement rare. De plus, ce trait, dans le contexte de cette société particulière qui y est hostile, nuit à la capacité de l'individu roux à se reproduire et à survivre. En effet, les personnes rousses pourraient subir de la violence dans cette société. En outre, comme ces dernières sont rejetées, elles ont moins de facilité à rencontrer un partenaire pour se reproduire. Donc, selon la TBS, avoir les cheveux roux, dans cette situation, est une pathologie. Mais cette conclusion est insensée. Être rousse ou roux n'est pas un état pathologique. Bref, ces deux exemples montrent que la TBS a de la difficulté à saisir (et par le fait même à distinguer) la normalité ou l'anormalité d'un trait ou d'un mécanisme.

En somme, comme on peut le constater, en raison de ces quatre problèmes que pose sa composition conceptuelle interne, la TBS échoue à saisir la normativité dans les systèmes vivants. Elle ne peut donc pas être utilisée pour distinguer le normal du pathologique et, par conséquent, il serait mal avisé qu'elle soit au fondement des RDoC.

2.4 La dysfonction préjudiciable (DP) et la dysfonction correctement dévalorisée (DCD) : des alternatives à la TBS?

On a vu que la TBS a de la difficulté, en raison des limites que lui imposent ses composantes conceptuelles, à saisir la normativité inhérente aux traits ou aux mécanismes des systèmes vivants et, conséquemment, à identifier un trait ou un mécanisme dysfonctionnel chez ces derniers (chose qu'un concept de trouble mental devrait être capable de faire) (section 2.3). Est-ce que les tenants des RDoC pourraient apporter une solution à ce problème en adoptant une conception du trouble mental dont la composition conceptuelle est différente de celle de la TBS? Peut-être pourraient-ils adopter la dysfonction préjudiciable⁸⁹ (DP) de Wakefield (1992, 2006a; 2019) ou la dysfonction correctement dévalorisée⁹⁰ (DCD) de Powell et Scarffe (2019b, 2019a) qui satisfont chacune ce critère? Elles sont, en effet, des conceptions du trouble mental qui ne contiennent pas la même conception de la fonction que la TBS ni le même concept de classe de référence. Malheureusement, après avoir présenté ces dernières (sections 2.4.1 et 2.4.2), je montre qu'elles rencontrent elles aussi des problèmes qui font en sorte qu'elles ne peuvent pas être adoptées comme solution de rechange à la TBS (section 2.4.3). Pour cette raison, au troisième chapitre, je fais connaître la CÉOST

⁸⁹ Il s'agit de la traduction de l'expression « harmful dysfunction ».

⁹⁰ Je traduis l'expression anglaise « properly disvalued » par l'expression française « correctement dévalorisée ».

qui, à mon avis, peut nous sortir de cette impasse (voir la section 3.2 et, particulièrement, la section 3.2.2).

2.4.1 La DP de Wakefield

Wakefield a développé le concept de *dysfonction préjudiciable* (DP) (voir, entre autres, Horwitz et Wakefield, 2010; Wakefield, 1992, 2006a, 2011; Wakefield et Conrad, 2019). Il s'agit d'une conception générale du trouble⁹¹ (Wakefield, 2006a, p. 39). En effet, la DP s'applique en principe autant aux troubles physiques (comme le cancer) qu'aux troubles mentaux (comme la dépression majeure) (ibidem, 2006a, p. 39 et 46). La DP est aussi une conception du trouble mental qui tente d'intégrer le meilleur des approches biologiques (Boorse, 1975, 1976a, 1997, 2014; Kendell, 1975) et normatives (par exemple, Scheff, 1966; Szasz, 1961)⁹². La DP contient, en effet, une composante factuelle que l'on peut associer à l'approche biologique (comme celle de Boorse) et une composante évaluative que l'on peut associer à l'approche normative.

La première composante est le concept « dysfonction ». Celui-ci réfère au monde objectif des faits scientifiques. Il faut comprendre le concept « dysfonction » dans le cadre de la théorie de l'évolution et à la lumière du concept « fonction naturelle ». Dans un organisme, une fonction naturelle est conçue comme un effet d'un mécanisme (ou d'un organe) qui, au cours de l'histoire évolutive, a contribué positivement (relativement aux organismes qui ne possédaient pas ce mécanisme ou qui en possédaient un fonctionnant à un degré différent de celui-ci) à la survie et à la reproduction des membres de l'espèce à laquelle appartient cet organisme et qui, pour cette raison, a été sélectionné. Par exemple, dans un organisme vivant, la fonction

91 Le mot « trouble » est une traduction française du mot anglais « disorder ».

⁹² Wakefield tente, en effet, de mettre fin à une querelle sur la conception des troubles mentaux ayant eu lieu entre les défenseurs de ces deux approches dans les années 60-70 (voir Bolton, 2008, p. 103-116 pour un résumé de ce débat).

naturelle du cœur (un organe) est de pomper le sang (et non de faire du bruit. Le bruit que fait le cœur, de fait, n'a pas contribué à la capacité de reproduction et de survie des ancêtres de cet organisme et, par voie de conséquence, n'a pas été sélectionné). De la même façon, certains font l'hypothèse qu'une des fonctions naturelles du mécanisme de la peur est de faire fuir l'organisme quand il y a du danger. Enfin, on suppose que la fonction des systèmes perceptuels (visuel, olfactif, sensoriel, etc.) est de transmettre à l'organisme des informations sur son milieu intérieur et sur l'environnement extérieur.

Wakefield adopte ainsi, comme le disent Roe et Murphy (2011, p. 219), une conception sélectionniste de la fonction. Suivant cette dernière, Wakefield dira qu'« [u]ne "dysfonction" existe [chez un organisme vivant] lorsqu'un mécanisme [ou un organe] interne [de cet organisme] est incapable d'accomplir sa fonction naturelle » (Wakefield, 2006a, p. 46, mes italiques). Par exemple, si le cœur ne pompe pas le sang adéquatement, il y a présence d'une dysfonction cardiaque. De la même façon, si le mécanisme de la peur ne fait pas fuir l'organisme quand il y a un danger, il y a une dysfonction de ce mécanisme. Enfin, si le système perceptuel cause constamment des hallucinations auditives et visuelles chez l'organisme, dans un environnement que l'on pourrait qualifier de normal, ce système est dysfonctionnel.

On remarquera ici que cette conception de la fonction à la base de cette conception de la dysfonction n'est pas celle que dit utiliser Boorse. En effet, alors que Boorse dit adopter ⁹³ une conception *anhistorique* de la fonction qu'il qualifie de normale (section 1.5.3), Wakefield adopte plutôt une conception *historique* selon laquelle la fonction d'un trait chez un organisme est un *effet* sélectionné, dans le *passé*, chez les

⁹³ J'écris « dit adopter » parce que, comme je le soutiens à la section 1.5.3, la conception de la fonction normale de Boorse ne semble pas, dans les faits, être une conception anhistorique. En effet, elle semble plutôt être une conception historique qui contient un aspect statistique.

ancêtres de cet organisme en raison de sa contribution au succès reproducteur de ces derniers. Bref, selon Wakefield, la fonction d'un trait est déterminée par l'utilité qu'elle *a eue* chez les ancêtres de l'organisme et pas pour son utilité adaptative dans le présent.

Maintenant que cela est précisé, on peut aborder la deuxième composante de la DP, à savoir le concept « préjudiciable ». Celui-ci réfère aux valeurs et aux constructions sociales contenues dans le monde symbolique humain. Le concept « préjudiciable », explique Wakefield (2006a, p. 39), « est un terme évaluatif se rapportant aux conditions jugées négatives selon les standards socioculturels ». Il est important de mentionner, comme l'ont récemment précisé Wakefield et Conrad, que les standards socioculturels

[...] ne sont pas synonymes des attitudes ou des opinions qui prédominent à un moment donné. Ils font plutôt partie d'un système de valeurs culturelles qui a une structure complexe à plusieurs niveaux et qui peut faire l'objet d'un examen critique et d'une révision [...] (Wakefield et Conrad, 2019, p. 1)

Voici deux exemples qui permettent de mieux comprendre le concept « préjudiciable ». Le premier est celui de quelqu'un qui présente une dépression majeure. Cette condition est préjudiciable parce qu'elle est jugée négativement selon les standards socioculturels de l'Amérique du Nord parmi lesquels figurent le bonheur individuel, l'indépendance et la réussite professionnelle. Les préjudices résultant d'une condition jugée négativement comme la dépression majeure peuvent être de la souffrance personnelle, une augmentation du risque de mortalité, l'incapacité de travailler et de la difficulté à entrer en relation avec les autres (McNally, 2011, p. 71). Le deuxième exemple est celui d'être sans emploi. Ce statut est aussi une condition préjudiciable. De fait, en regard des standards socioculturels en Amérique du Nord susmentionnés, il est préjudiciable de ne pas avoir de travail rémunéré. Les préjudices causés par le fait d'être sans emploi sont, par exemple, une baisse de l'estime de soi et une perte d'indépendance.

Il est important de mentionner que ces composantes factuelle et évaluative sont indissociables. Pour cette raison, un trouble physique ou mental ne peut pas être de nature essentiellement biologique (factuel) et ne peut pas être de nature essentiellement sociale. Il doit être nécessairement à la fois biologique et social. C'est ce que Wakefield (1992, p. 373) affirme : «l'ordre qui est perturbé lorsqu'on est atteint d'un trouble est [...] à la fois biologique et social; ni l'un ni l'autre ne suffit à lui seul à justifier l'étiquette de trouble. » Dans cette perspective, le cancer est un trouble parce qu'il résulte d'une dysfonction du développement cellulaire qui est jugée préjudiciable selon les standards sociaux (par exemple, le bonheur individuel et l'indépendance). De fait, avoir le cancer cause des préjudices à l'individu comme de la tristesse, de la détresse, une diminution de sa qualité de vie due à sa souffrance et, peut-être, sa mort. On peut aussi dire que la dépression majeure est un trouble mental. En effet, celle-ci résulte, entre autres, de plusieurs dysfonctions psychobiologiques (comme une perturbation de l'activité des circuits neuraux responsables de la régulation des émotions) qui causent des préjudices à l'individu (comme de la détresse et une baisse de l'estime de soi). À l'inverse, suivant la DP, les angiomes bénins, lesquels résultent d'une dysfonction de la croissance des vaisseaux sanguins, ne sont pas considérés comme des troubles parce qu'ils ne sont pas jugés préjudiciables (Wakefield, 2006a, p. 41). Selon cette même logique, la drapétomanie, un diagnostic inventé par le médecin américain Samuel Cartwright en 1851 et que l'on attribuait aux esclaves noirs qui tentaient de fuir leur captivité, n'était pas un trouble mental. En effet, cette croyance selon laquelle les esclaves ont pour fonction naturelle de servir est tout simplement fausse (Wakefield, 1992, p. 386). De fait, il n'y a pas de dysfonction que l'on peut associée à la drapétomanie. Celle-ci, en effet, n'était pas un trouble mental, mais essentiellement une catégorie pour classifier les personnes noires qui adoptaient un comportement jugé de manière négative dans le cadre de la culture raciste de l'époque. Dans le même ordre d'idée, la condition « sans emploi » dont je parlais précédemment n'est pas un trouble puisqu'aucune dysfonction n'est en principe sous-jacente à cette condition.

En résumé, pour Wakefield, un trouble mental est (1) une dysfonction qui est (2) jugée préjudiciable selon les standards socioculturels⁹⁴.

2.4.2 La DCD de Powell et Scarffe

Powell et Scarffe (2019b, 2019a) ont développé récemment la conception selon laquelle « un état biomédical n'est une maladie que s'il implique un dysfonctionnement biologique qui est, ou serait, correctement dévalorisé» (Powell et Scarffe, 2019b, p. 582). Formulé de manière plus succincte, selon ces deux philosophes, un trouble mental ou physique est une « dysfonction correctement dévalorisée » (DCD).

Comme on peut le constater, la conception du trouble de Powell et Scarffe est, comme celle de Wakefield, une conception hybride. Elle contient une composante factuelle (l'expression « dysfonction biologique ») et une composante normative (l'expression « correctement dévalorisée »). Parce que la conception de la dysfonction qu'adoptent Powell et Scarffe est identique à celle qu'adopte Wakefield que j'ai présentée à la section 2.4.1, je ne m'y attarderai pas ici.

La composante normative qu'adoptent Powell et Scarffe, cependant, n'est pas celle qu'adopte Wakefield. En effet, plutôt que de faire appel aux standards socioculturels pour évaluer les dysfonctions (section 2.4.1), ceux-ci utilisent ce qui semble être le cadre de la bioéthique (Powell et Scarffe, 2019a, p. 1175). Dans cette perspective, la préjudiciabilité des dysfonctions est évaluée selon des justifications rationnelles

même raison, je ne m'intéresse pas non plus à la thèse de Wakefield selon laquelle le concept de trouble doit correspondre aux intuitions qu'ont les gens de ce qu'est un trouble (à ce sujet, voir Wakefield, 2006a, p. 40, 2007, p. 150).

⁹⁴ Pour être plus exact, il faudrait dire qu'un trouble mental est une dysfonction mentale préjudiciable (Wakefield, 1992, p. 385; voir aussi, 2006b pour plus de détails à ce sujet). Dans le cadre de cette exposition de la conception du trouble de Wakefield, cependant, je ne m'intéresse pas à cette subtilité parce que la critique que je fais, à la section 2.4.3, de la DP ne requiert pas que je m'y attarde. Pour la

morales [rational moral justification] (Powell et Scarffe, 2019b, p. 582). Ce type de justifications implique deux éléments. Le premier est présupposé par le terme « moral » et dit qu'il existe des normes (ou des valeurs) morales objectivement justifiables. Le deuxième élément est sous-jacent à l'expression « justification rationnelle ». Cet élément suppose qu'une justification rationnelle (ici, morale) doit contenir des raisons empiriquement fondées et pertinentes. Cela implique que ces raisons ne relèvent pas de ce qui est populaire ou de la tradition. Comme le disent Powell et Scarffe,

[...] ces raisons ne doivent pas être arbitraires ou fanatiques, elles ne doivent pas reposer sur de fausses affirmations empiriques ou sur des croyances religieuses impénétrables et elles doivent faire l'objet d'un examen et d'une révision critiques. (Powell et Scarffe, 2019b, p. 582)

Selon Powell et Scarffe, faire une justification morale rationnelle a pour but d'évaluer si les dysfonctions sous-jacentes à une condition contreviennent, chez l'*individu*⁹⁵, à l'atteinte de normes (ou de valeurs) morales (par exemple, la norme selon laquelle le bien-être de l'individu est nécessaire à une vie réussie). Si elles ne contreviennent pas à cet objectif, il y a, certes, une dysfonction, mais pas de trouble. En revanche, si elles nuisent à cet objectif, on est alors en présence d'un trouble.

Powell et Scarffe appliquent leur DCD au trouble de stress post-traumatique (TSPT) (Powell et Scarffe, 2019b, p. 584). D'une part, ils montrent que la composante factuelle de cette dernière saisit le fait empirique que les personnes présentant cette condition ont des dysfonctions neurobiologiques associées à des dysfonctions cognitives quant à la régulation des émotions. Parmi ces mécanismes dysfonctionnels, on retrouve une hyperactivité de l'amygdale et du cortex insulaire qui sont responsables (entre autres)

_

⁹⁵ Je mets « individu » en *italique* parce que cette évaluation « ne doit pas tenir compte des coûts indirects pour la *société* (qu'ils soient sociaux, économiques, esthétiques ou autres) » (Powell et Scarffe, 2019b, p. 583).

de la réponse de peur et du réflexe de sursaut. On retrouve également une hypoactivation des régions du cerveau responsables du traitement des émotions négatives, laquelle expliquerait la réactivité émotionnelle plus faible, lors des épisodes de dissociation et de gel émotif, que vivent les personnes présentant un TSPT. D'autre part, ils montrent que la composante normative de leur conception saisit le fait que les conséquences psychologiques (ex. : événement traumatique revenant par le biais de pensées intrusives, insomnie, cauchemars, hypervigilance et hypersensibilité aux stimuli ambiants, détachement émotionnel), sociales (ex. : difficulté à avoir des relations stables et à entretenir des relations intimes) et physiques (ex. : dégradation de la santé en raison d'une réponse au stress exacerbée et à l'apparition de comportements destructeurs) des dysfonctions biologiques du TSPT nuisent à l'autonomie de la personne et à son épanouissement dans le monde d'aujourd'hui.

Suivant ce qui précède, selon Powell et Scarffe, le TSPT est donc un trouble mental puisque les deux conditions de la DCD sont remplies par cette dernière. En effet, (1) des dysfonctions biologiques sont présentes dans l'individu et (2) on peut dévaloriser ces dysfonctions parce qu'elles nuisent à l'atteinte d'un certain nombre de choses auxquelles il est moralement justifié d'aspirer (ex. : l'autonomie, la formation de liens sociaux, l'épanouissement, la réussite de sa vie).

2.4.3 Critique de la DP et de la DCD

La DP (section 2.4.1) et la DCD (section 2.4.2) sont intéressantes. Néanmoins, à l'instar de la conception du trouble issue de la TBS, elles sont problématiques. En effet, elles prêtent le flanc à trois critiques visant leur composante factuelle, à savoir la conception étiologique de la fonction. D'abord, cette dernière, comme le fait remarquer Stegenga (2018a, p. 24), à l'instar de celle de Boorse, rencontre le problème du choix de la classe de référence. Cependant, parce que, dans le cadre de la conception étiologique, la classe de référence n'est pas définie de la même manière que dans la

TBS, Stegenga dit qu'il s'agit d'une *version différente* du problème de la classe de référence de la version présentée à la section 2.3.

Quelle est cette nouvelle définition de la classe de référence ? Selon la conception étiologique de la fonction, la classe de référence renvoie aux caractéristiques d'un moment précis de l'histoire évolutive des ancêtres du système vivant présentant le trait porteur de la fonction. Suivant cela, le problème de la détermination de la classe de référence prend la forme de cette question : considérant que les pressions sélectives changent avec le temps, quel moment de l'histoire évolutive des ancêtres de l'organisme à l'étude est-il le plus approprié pour évaluer le fonctionnement normal d'un trait ou d'un mécanisme mental? Comme Stegenga le note,

[d]evrions-nous évaluer la fonction normale en tenant compte des pressions sélectives auxquelles l'entité ou le processus a été confronté lorsque cette entité ou ce processus est apparu la première fois ou devrions-nous tenir compte des pressions sélectives plus récentes? Ou des pressions sélectives en ce moment? Le verdict relativement à l'attribution des maladies peut différer en fonction de la référence temporelle que chacun [utilise] pour évaluer les pressions sélectives. Par exemple, dans les sociétés préagricoles, courir lentement peut avoir été un désavantage sélectif, mais pas dans la société moderne [...]. (Stegenga, 2018a, p. 24)

On peut penser que choisir cette classe de référence — c'est-à-dire ce moment dans l'histoire évolutive de l'espèce à laquelle appartient le système vivant considéré — exige que l'on fasse appel à des considérations non biologiques de nature normative. Ainsi, la DP et la DCD qui contiennent la conception étiologique, contrairement à ce qu'elles supposent, et à l'instar de la TBS, ne peuvent pas appréhender de manière objective la normativité inhérente aux systèmes vivants.

On peut maintenant passer à la seconde critique. Si on accepte, pour les besoins de l'argumentaire, que la conception étiologique puisse rendre compte des fonctions et des dysfonctions des systèmes vivants de manière objective, cela sauve-t-il la DP et la

DCD? On peut penser que non parce que *la présence d'une dysfonction*, contrairement à ce que supposent la DP et la DCD, *n'est pas nécessaire* pour qu'il y ait présence d'un trouble mental. En effet, il est *théoriquement plausible* de penser qu'une fonction adaptative *saisie par la conception étiologique* puisse être sous-jacente à certaines conditions que l'on conçoit aujourd'hui comme des pathologies mentales (Cooper, 2007, p. 33-34; Garson, 2015, p. 177-178).

Pour illustrer cette affirmation, on peut utiliser deux hypothèses psychoévolutionnistes qui supposent que la classe de référence pour évaluer le fonctionnement d'un trait soit l'époque du pléistocène. La première hypothèse est que « [...] certains [organismes] [...] [furent] conçus par l'évolution pour être des psychopathes (McGuire et Troisi, 1998, p. 191-193) » (Garson, 2015, p. 177). Pour justifier cette hypothèse, Garson⁹⁶ affirme que « [l]a psychopathie [...] *pourrait aider à survivre* dans des environnements difficiles en émoussant la capacité naturelle de sympathie envers les autres » (Garson, 2015, p. 177, mes *italiques*).

La deuxième hypothèse qu'énonce Garson est que « la dépression a pu aider [...] nos ancêtres primates [à survivre et à se reproduire] en les dissuadant de poursuivre des objectifs déraisonnables ». Pour donner de la crédibilité à cette conjecture, Garson donne l'exemple suivant :

Supposons [...] qu'un primate mâle perde une bataille pour la nourriture au profit d'un mâle plus grand. Il pourrait envisager de se battre à nouveau, même si les chances de succès sont grandement contre lui. La dépression s'installe, et il se trouve privé de sa volonté, *ce qui lui permet de vivre*

 $^{^{96}}$ À noter que Garson n'est pas un partisan ni un défenseur de la psychologie évolutionniste. En fait, il est plutôt critique à son endroit (voir Garson, 2015, p. 65-66).

longtemps (Price et al., 1994; Rottenberg, 2014). (Garson, 2015, p. 177, mes italiques)

Si ces deux hypothèses sont vraies, alors deux choses s'en suivent. La première est qu'une fonction adaptative est sous-jacente à des conditions mentales qu'on juge aujourd'hui, selon des intuitions médicales, être des pathologies. La deuxième est que la DP et la DCD ne peuvent pas saisir que la psychopathie et la dépression sont des troubles mentaux. En effet, elles exigent toutes deux qu'une dysfonction soit sous-jacente à un trouble mental (ou physique). La conception étiologique de la fonction inhérente à la DP et à la DCD pose donc un problème.

Une troisième critique qui peut être faite à l'endroit de la DP et de la DCD est que les fonctions et les dysfonctions de l'esprit qu'elles saisissent avec la conception étiologique sont tirées de spéculations sans fondements empiriques. En effet, on ne pourra pas vérifier si les hypothèses psychoévolutionnistes susmentionnées sont vraies. De fait, il semble impossible d'avoir accès à ces moments de l'histoire évolutive où sont potentiellement apparues les fonctions psychobiologiques des êtres humains. En effet, comme le dit Zachar (2014b, p. 85) suivant Richardson (2007) :

[...] il n'y a presque pas de données empiriques en biologie et en paléontologie sur l'évolution du cerveau humain, et encore moins sur les pressions de sélection responsables du développement de capacités psychologiques spécifiques [...] [Ainsi], les fonctions psychologiques naturelles ne sont plus du tout des faits observables (dans un sens évolutionniste). (Voir aussi Faucher, 2012, p. 43 et, 2021, p. 62-63 pour un avis semblable)

Bref, compte tenu de la conception étiologique de la fonction qu'elles contiennent, la DP et la DCD sont problématiques pour trois raisons. Premièrement, la détermination du moment de l'histoire évolutive des ancêtres de l'organisme qui servira de classe de référence à la lumière de laquelle on évaluera si un trait est fonctionnel ou dysfonctionnel implique des considérations normatives. En conséquence, la DP et la

DCD sont incapables d'appréhender objectivement la normativité inhérente aux systèmes vivants. Deuxièmement, la DP et la DCD ne saisissent pas certaines conditions pathologiques qui ont théoriquement pour fondement une fonction adaptative. Troisièmement, les fonctions et dysfonctions de l'esprit que la DP et la DCD saisissent sont tirées d'hypothèses psychoévolutionnistes hautement spéculatives à propos de notre lointain passé.

2.5 Conclusion du deuxième chapitre

Au chapitre 2, j'ai examiné la conception du trouble mental des RDoC et j'ai conclu qu'elle était problématique pour quatre raisons. Premièrement, elle suppose une conception de l'esprit qui n'est pas en harmonie avec ce qu'ont dit plusieurs études empiriques sur l'incrustation de l'esprit dans l'environnement et sur la constitution extraneurale de la cognition (voir sections 2.1.1-2.1.4). Deuxièmement, en conséquence, cette conception de l'esprit défavorise fortement ou empêche que soient menées des études sur les mécanismes environnementaux qui pourraient favoriser le développement des troubles mentaux et sur les mécanismes corporels, multipersonnels et cognitivo-insitutionnels qui constituent potentiellement les troubles mentaux (voir sections 2.1.1-2.1.4). Troisièmement, le principe épistémoquantitatif sous-jacent à cette conception des troubles mentaux défavorise fortement ou empêche que des études phénoménologiques et qualitatives soient menées en santé mentale et, en conséquence, empêche que soient découvertes des connaissances présentant une pertinence épistémique, diagnostique et thérapeutique en santé mentale (sections 2.2.1 et 2.2.2). Quatrièmement, elle contient la TBS de Boorse (1975, 1976a, 1977, 1997, 2014) qui a de la difficulté à saisir la normativité et, en conséquence, les dysfonctions dans les systèmes vivants (section 2.3). On aurait pu penser que la DP de Wakefield (2010; 1992, 2006a; 2019) (section 2.4.1) et la DCD de Powell et Scarffe (2019b, 2019a) (section 2.4.2), parce qu'elles contiennent une conception étiologique de la fonction et, implicitement, un concept différent de classe de référence, auraient pu éviter ce problème. Malheureusement, ce n'est pas le cas (section 2.4.3). Que faire pour remédier à cette situation problématique? Au troisième chapitre, je propose une solution à toutes ces difficultés auxquelles se bute la conception du trouble mental des RDoC.

CHAPITRE III

UNE CONCEPTION ÉNACTIVE ET OBJECTIVE DE LA SANTÉ ET DU TROUBLE (CÉOST) : UNE SOLUTION À QUATRE PROBLÈMES DE LA CONCEPTION DU TROUBLE MENTAL DES RDOC

Au chapitre 2, j'ai identifié quatre problèmes que rencontre la conception du trouble mental des RDoC (voir la section 2.5 pour un résumé de ces derniers). Devant ce constat, je soutiens qu'il faut développer une manière différente d'appréhender les troubles mentaux qui résout ces quatre difficultés. Je m'attelle donc à cette tâche dans ce troisième chapitre où je présente la conception énactive et objective de la santé et du trouble (CÉOST) (section 3.2). La CÉOST suppose la conception énactive autopoïétique de la cognition (CÉAC) qui résout les quatre problèmes susmentionnés. Puisque la CÉOST a pour base la CÉAC, je présente cette dernière à la section 3.1. Ensuite, à la section 3.1.1, je montre que la CÉAC résout les trois premiers problèmes de la conception du trouble mental des RDoC. Enfin, après avoir présenté la CÉOST (section 3.2.1), je montre que cette dernière apporte une solution au quatrième problème (section 3.2.2).

3.1 La conception énactive autopoïétique de la cognition (CÉAC)

The Embodied Mind: Cognitive Science and Human Experience (dorénavant TEM) (Varela et al., 1991) est l'ouvrage canonique où a d'abord été développé ce que l'on

pourrait appeler le « programme de recherche » en sciences cognitives énactives. Selon Ward, Silverman et Villalobos (2017), dans la foulée de l'énactivisme originel que l'on trouve dans TEM, trois types d'énactivisme furent élaborés et le sont encore aujourd'hui. Ceux-ci sont l'énactivisme autopoïétique (Di Paolo, 2005; Thompson, 2007; Varela, 1997; Weber et Varela, 2002), l'énactivisme sensorimoteur (Noë, 2004, 2012; O'Regan et Noë, 2001) et l'énactivisme radical (Hutto et Myin, 2013, 2017).

Pour construire la conception énactive et objective de la santé et du trouble (CÉOST) (section 3.2.1), je choisis d'utiliser principalement⁹⁷ l'énactivisme autopoïétique ou, plus précisément, la conception énactive autopoïétique de la cognition (CÉAC). Principalement, trois raisons motivent ce choix. La première est que la CÉAC contient une définition de la normativité du vivant qui me permet de développer les notions de santé et de trouble que contient la CÉOST (voir section 3.2.1). La deuxième raison est que la notion d'institution cognitive de Gallagher que je mobilise pour construire la CÉOST suppose la CÉAC, mais pas les énactivismes sensorimoteur et radical (voir les sections 2.1.4 et 3.2.1). La troisième raison est que la CÉAC, contrairement à l'énactivisme sensorimoteur et à l'énactivisme radical, suppose une conception énactive de l'évolution que j'utilise dans la CÉOST et qui me permet de distinguer cette dernière de la théorie biostatistique (TBS) de la santé (sous-jacente au concept de trouble mental des RDoC) qui contient une conception adaptationniste ⁹⁸ de l'évolution⁹⁹ (voir section 1.5.3).

⁹⁷ J'écris « principalement » parce que, pour construire la CÉOST (section 3.2.1), je mobilise aussi certaines notions provenant des travaux en biologie théorique de Moreno, Mossio et Saborido (Mossio *et al.*, 2009, 2010 ; Mossio et Moreno, 2015 ; Saborido et Moreno, 2015 ; Saborido *et al.*, 2016).

⁹⁸ À noter que la téléosémiotique de l'énactivisme radical suppose aussi la conception sélectionniste ou adaptationniste de l'évolution (Thompson, 2018).

⁹⁹ Une quatrième raison qui me porte à adopter la CÉAC est qu'elle contient un projet de naturalisation de la phénoménologie (Petitot *et al.*, 1999; Thompson et Lutz, 2003; Varela, 1996) qui est pertinent pour la psychiatrie. Dans le cadre de cette thèse, cependant, je n'utilise pas ces thèses énactives sur la

Maintenant que j'ai dit pourquoi j'adoptais la CÉAC, je vais l'examiner de plus près. Celle-ci est issue à l'origine d'une réflexion fondamentale qu'ont menée Maturana et Varela (1980) sur les systèmes vivants. Cette réflexion a été poursuivie par d'autres philosophes et d'autres scientifiques (Di Paolo, 2005, 2009; Di Paolo et Thompson, 2014; Thompson, 2007, 2017; Varela, 1997; Varela et al., 1993; Weber et Varela, 2002). Il ressort des travaux de ces derniers que les systèmes vivants (et par conséquent les êtres humains) présentent, selon mon analyse, au moins dix propriétés. Ils sont des systèmes autopoïétiques (1) opérationnellement clos et (2) thermodynamiquement ouverts. Ils sont également (3) précaires, (4) adaptatifs, (5) autonomes et (6) nécessairement cognitifs. Il faut ajouter à cela (7) qu'ils ont une expérience subjective (plus ou moins complexe) du monde et d'eux-mêmes et (8) qu'ils sont, étant donné qu'ils sont des systèmes cognitifs, des constructeurs de mondes de significations plus ou moins sophistiquées qu'ils partagent (selon divers degrés) entre eux. Enfin, (9) ils sont des systèmes dynamiques et (10) forment eux-mêmes un système dynamique avec leur environnement (autant physique, social que culturel). Je présente dans ce qui suit ces dix propriétés et la manière dont elles sont reliées entre elles.

J'ai dit précédemment que les systèmes vivants sont des systèmes autopoïétiques ¹⁰⁰. Selon Varela et Weber, un système autopoïétique

[...] est organisé (défini comme unité) comme un réseau de processus de production (synthèse et destruction) de composants tels que ces composants :

naturalisation de la phénoménologie. Toutefois, comme je l'indique dans la conclusion, il serait pertinent d'explorer ces dernières dans des recherches futures.

_

¹⁰⁰ À noter, cependant, que tous les systèmes autopoïétiques ne sont pas nécessairement des systèmes vivants (pour plus de détails à ce sujet, voir Bitbol et Luisi, 2004).

- (1) régénèrent continuellement le réseau qui les produit, et
- (2) constituent le système en tant qu'unité distincte dans l'espace où ils [(les composants)] existent. (Weber et Varela, 2002, p. 115. Voir Varela, 1989, p. 45 pour une formulation légèrement différente)

L'idée exprimée dans la première clause est que les systèmes vivants sont constitués de processus qui se produisent eux-mêmes (voir Varela *et al.*, 1993, p. 199-200). Pour être capables de s'autoproduire, ces systèmes doivent, entre autres, perpétuellement se fabriquer une membrane semi-perméable. En effet, celle-ci met à l'abri de l'environnement les réseaux de processus métaboliques internes qui produisent les composants formant le système vivant et, conséquemment, la membrane semi-perméable qui fait elle-même partie de ce système et qui permet l'existence de ces réseaux de processus. Il est à noter que, selon le type de système, cette frontière qui délimite l'intérieur et l'extérieur du système n'est pas toujours une membrane cellulaire (comme en est une, entre autres, la peau chez un humain). Par exemple, « [1]es membres d'une colonie d'insectes [...] forment un réseau social autonome, mais la frontière est sociale et territoriale [...] » (Thompson, 2007, p. 44).

La deuxième clause introduit l'idée selon laquelle le système vivant émerge dans un espace — j'utiliserai dorénavant le terme « environnement » — duquel il se distingue en tant qu'unité dynamique concrète tout au long de son existence. En introduisant cette idée, la deuxième clause montre qu'on ne peut pas concevoir un système vivant sans tenir compte de l'environnement où il évolue. On peut expliquer cela par au moins deux raisons. Premièrement, pour pouvoir exister en tant qu'unité dynamique distincte, il faut nécessairement que ce système ne se dissolve pas dans l'espace où il se trouve. Il faut, au contraire, qu'il s'en distingue. Deuxièmement, pour être en mesure de rester en vie (et de se distinguer de l'environnement), le système vivant a besoin de consommer les sources d'énergie de l'environnement et d'y évacuer ses déchets. Il a donc besoin d'être en relation constante avec lui. C'est la membrane semi-perméable

du système vivant qui concrètement rend possible cette relation où ont lieu ces échanges d'énergie¹⁰¹. Selon une perspective plus théorique et abstraite, on peut dire aussi que ce système vivant est en mesure d'entretenir de telles transactions énergétiques avec l'environnement parce qu'en plus d'être (1) *opérationnellement clos* (au sens où il est constitué de processus qui se produisent eux-mêmes et le constituent en tant qu'unité dynamique distincte), il est (2) *thermodynamiquement ouvert*.

Le schéma de Varela suivant illustre bien ce qui vient d'être présenté, à savoir, d'une part, la dynamique métabolique interne ayant lieu dans le système vivant membrané qui s'autoproduit en tant qu'unité dynamique distincte et, d'autre part, la relation que ce dernier entretient avec l'environnement :

¹⁰¹ L'énergie échangée est contenue essentiellement dans les entités matérielles. Et comment déterminer si une entité est matérielle? Une entité est matérielle si et seulement si elle change (Bunge, 2006, p. 12). Et parce que le mot « énergie » est la notion théorique qui saisit la « changeabilité », alors on peut dire qu'une entité est matérielle si et seulement si elle contient de l'énergie (Bunge, 2006, p. 12) (voir aussi section 1.5.5). L'énergie peut donc se trouver autant dans le sucre, dans une plante, dans un système vivant ou dans la lumière. En effet, toutes ces entités changent d'état et, en conséquence, sont matérielles.

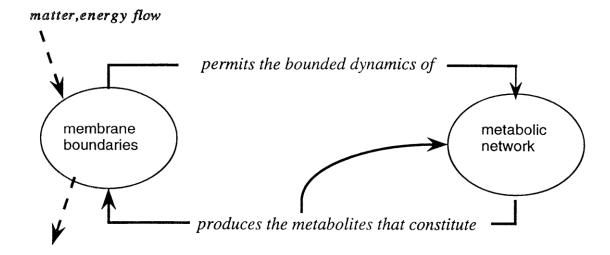


Figure 3.1 Mécanisme autopoïétique nécessaire à l'existence d'un système vivant (Varela, 1997, p. 75)

Récemment, Di Paolo (2009) et Di Paolo et Thompson (2014) ont développé la thèse selon laquelle un système vivant, en plus d'être opérationnellement clos et thermodynamiquement ouvert, a aussi (3) la propriété d'être constitué de processus *précaires*¹⁰². On qualifie ces derniers de la sorte parce que, s'ils n'opèrent pas dans le système vivant dont ils font partie, leurs activités diminuent et, éventuellement, s'arrêtent (2014, p. 72). Conséquemment, ces processus cessent d'exister. Ces derniers ne peuvent donc exister que dans le réseau des processus eux-mêmes précaires qui constituent le système dont ils font partie. Les processus précaires constituant les systèmes vivants entretiennent donc une relation de dépendance les uns envers les autres. Compte tenu de cela, on peut dire que le système vivant « dans son ensemble est lui-même la condition [d'existence] de ses parties » (Fuchs, 2018, p. 85).

¹⁰² Avant Di Paolo et Thompson, Weber et Varela (2002, p. 113 et 120), suivant Hans Jonas (1973), ont soulevé cette idée que les systèmes vivants sont des systèmes précaires, mais ils ne l'ont pas beaucoup développée.

Froese (2017) pense que c'est parce que les systèmes vivants sont précaires et ultimement mortels qu'ils agissent continuellement dans le but de se conserver et qu'en conséquence (4) ils font preuve activement d'*adaptabilité*. Bien que cette notion théorique soit implicite dans certains écrits de Varela (1997), c'est Di Paolo (2005, 2009) qui, dans le cadre énactif, est le premier à l'avoir clairement explicitée. Il définit l'adaptabilité comme

[l]a capacité d'un système, dans certaines circonstances, à réguler ses états et sa relation avec l'environnement, de sorte que, si les états [du système] sont suffisamment proches de la limite de sa viabilité,

- (1) [1]es tendances sont distinguées et traitées selon que les états s'approchent ou s'éloignent de cette limite et, par conséquent,
- (2) [l]es tendances du premier type sont rapprochées des ou transformées en tendances du deuxième [type] et ainsi les états futurs [du système] sont empêchés d'atteindre cette limite trop rapidement. (Di Paolo, 2005, p. 438)

En d'autres termes, la notion d'adaptabilité réfère au fait que les systèmes vivants présentent la propriété « de s'autoréguler en fonction des limites de [...][leur] propre viabilité » (Di Paolo, 2005, p. 430). Cette viabilité est relative à une *norme interne* au système vivant, à savoir celle qui consiste pour ce système à tendre vers (a) le maintien de son autopoïèse (et plus généralement de son existence en tant qu'unité dynamique) et, conséquemment, à tendre vers (b) le maintien d'un couplage structurel ou, autrement dit, d'un état homéostatique avec l'environnement, et ce, malgré les perturbations de ce dernier au fil du temps. Bref, s'il réussit à fonctionner selon cette norme interne, le système vivant reste en vie.

Avant d'aller plus loin, il importe de noter deux choses. La première est qu'un système qui s'autoproduit en tant qu'unité dynamique distincte (au sens autopoïétique vu précédemment), qui fait activement preuve d'adaptabilité et qui, en conséquence, opère dans le but de s'automaintenir selon une norme interne ou selon « ses propres lois de

fonctionnement » (de Haan, 2020b, p. 82), a la propriété d'être (5) *autonome*. De Haan illustre cette propriété (l'autonomie) en faisant une comparaison éclairante :

Contrairement [...] à un robot qui se comporte selon les règles que j'ai programmées, un organisme vivant n'est pas déterminé de l'extérieur de cette façon : c'est sa propre organisation qui présente certaines lois auxquelles il doit se soumettre. (de Haan, 2020b, p. 82)

La deuxième chose importante à noter est que, sans cette notion d'adaptabilité, il ne serait pas possible de rendre compte de phénomènes comme ceux de stress, de fatigue, de maladaptation et de pathologie (Di Paolo, 2005, p. 440). En effet, on peut penser que si un système vivant n'a pas une capacité adaptative suffisante pour se maintenir dans sa zone de viabilité, il est susceptible de développer des troubles physiques et, de surcroit, pour les systèmes vivants doués d'un esprit suffisamment sophistiqué, des troubles mentaux. Un système vivant qui ne réussit pas à se maintenir dans sa zone de viabilité perdra graduellement de son autonomie et, éventuellement, cessera d'exister (j'aborderai tout cela plus en détail quand je présenterai, à la section 3.2.1, la CÉOST).

Maintenant que ces deux précisions ont été faites, on peut présenter la thèse énactiviste selon laquelle, pour rester adapté à son environnement, le système vivant peut mettre à profit *deux niveaux de processus métaboliques* (qui sont aussi des *processus cognitifs*, comme on le verra plus loin dans cette section).

Quand l'environnement n'est pas perturbé, le système utilise le premier niveau de processus métabolique. Celui-ci

prend en charge des composés « familiers », à savoir des métabolites qui ont accompagné la vie de l'être vivant et de sa progéniture au cours des générations, et qui se sont intégrés depuis longtemps à ses cycles récursifs d'interaction. Il sous-tend l'équilibre dynamique ordinaire de l'être vivant, autrement dit son homéostasie. (Bitbol et Cohen-Varela, 2017, p. 20)

Par exemple, si un système vivant quelconque sent qu'il va atteindre les limites de sa viabilité en termes de niveau d'énergie, alors il essaiera (consciemment ou non), dans l'environnement où il se trouve, de repérer et d'ingérer une source d'énergie qu'il connait et consomme régulièrement. S'il réussit, ce système maintiendra son couplage structurel avec l'environnement, ce qui contribuera à le garder en vie (c'est-à-dire à maintenir son autopoïèse). On dira alors que ce système vivant est resté adapté dans son environnement habituel.

Le second niveau de processus métabolique est « [...] suscité par la rencontre [...] [du système vivant¹⁰³] avec des composés nouveaux. Cette rencontre ouvre la possibilité de changements temporaires ou permanents dans les itinéraires métaboliques [...] [du système vivant] » (Bitbol et Cohen-Varela, 2017, p. 20-21).

Les changements dont il est question ici touchent la *structure* du système vivant. Un changement de structure est un changement de « l'ensemble des relations *actuelles* qui unissent les composants [...] [dont le système vivant est fait] » (Bitbol et Cohen-Varela, 2017, p. 21). Autrement dit, un changement de structure est un changement de « l'agencement des processus qui permettent [1]a réalisation physique [du système vivant] [...] » (Bitbol et Cohen-Varela, 2017, p. 21). Il est à noter qu'un changement de structure est associé, chez les systèmes vivants, à des « processus d'adaptation et d'évolution » (Bitbol et Cohen-Varela, 2017, p. 21).

¹⁰³ Bitbol et Varela-Cohen utilisent l'expression « unité autopoïétique ». Pour être cohérent avec le reste de mon texte, j'utilise pour ma part l'expression « système vivant ». Cela ne pose pas de problèmes parce que, selon le cadre énactif, tous les systèmes vivants sont des systèmes autopoïétiques (mais tous les systèmes autopoïétiques ne sont pas nécessairement des systèmes vivants. Voir Bitbol et Luisi, 2004 à ce sujet).

Avant de définir le concept d'évolution tel qu'il est compris dans le cadre énactif, il faut distinguer la notion de structure de celle d'*organisation*. Cette dernière est conçue comme «l'ensemble moins contraignant des relations *possibles* entre les types de composants [du système vivant] » (Bitbol et Cohen-Varela, 2017, p. 21). Il faut au moins retenir trois choses de cette notion. La première est que l'organisation d'un système vivant définit ce dernier comme un «élément d'une classe déterminée : la classe de tous les membres viables d'une espèce » (Bitbol et Cohen-Varela, 2017, p. 21). La deuxième chose à savoir est que l'organisation d'un système vivant ne peut pas changer¹⁰⁴. En effet, un changement de cette dernière entrainerait une modification trop profonde du système et, pour cette raison, sa mort (Escobar, 2012, p. 61-62). La troisième chose est que l'organisation d'un système vivant détermine en partie la capacité de ce dernier à s'adapter à des situations nouvelles. En effet, il faut que le changement de structure qui rend un système vivant quelconque viable pour une situation nouvelle soit permis par l'organisation de ce système.

Pour illustrer ce troisième point, on peut utiliser le cas fictif d'une colonie de systèmes vivants A évoluant dans un environnement E1. Ce cas suppose que les systèmes vivants A mangent toute la substance X de l'environnement où ils se trouvent. Cet environnement métamorphosé, qu'on peut nommer E2, aura alors à son tour un impact sur la structure des systèmes vivants A. Afin de démontrer cela, on peut poser que l'environnement E2 contient aussi une substance Y. En outre, on peut faire l'hypothèse que cette dernière se trouve maintenant en plus grande quantité étant donné qu'elle servait de nourriture aux systèmes vivants B dont la lignée s'est éteinte en raison d'une attaque des systèmes vivants C (lesquels ne s'intéressent pas aux systèmes vivants A et

 $^{^{104}}$ Cela ne signifie pas, cependant, qu'un système vivant ne peut pas présenter des structures très différentes. Par exemple, la structure que présente une chenille à un temps t_1 et la structure qu'elle présente à un temps t_2 une fois qu'elle s'est transformée en un papillon sont très différentes. Toutefois, selon le cadre énactif, ces structures très différentes sont permises par la même organisation.

ne mangent pas la substance Y). On peut penser que, dans cet environnement E2 où de nouvelles pressions sélectives sont opérantes, certains systèmes vivants A pourraient avoir un changement de structure que permet leur organisation et qui serait favorable au maintien de leur viabilité dans l'environnement E2. Par exemple, il pourrait avoir une mutation génétique qui leur permettrait de générer un module capable de produire des enzymes qui pourraient dégrader cette substance Y. Ce faisant, les systèmes vivants A établiraient une nouvelle forme de couplage structurel avec l'environnement E2 et pourraient ainsi maintenir leur autopoïèse. Bref, un changement de structure chez les systèmes vivants A permettrait le maintien de leur adaptation dans l'environnement E2 et, conséquemment, le maintien de leur organisation, ce qui serait essentiel au maintien de leur vie.

Quand un changement de structure advient chez les systèmes vivants A, on dit de ces derniers, non seulement qu'ils se sont adaptés à leur nouvel environnement (ou à cette perturbation de l'environnement), mais aussi qu'ils ont évolué. Mais attention, dans le cadre énactif, l'évolution d'un système vivant n'est pas conçue, comme c'est le cas selon l'adaptationisme empirique (voir section 1.5.3), comme une optimisation de l'adaptation par l'intermédiaire de la sélection naturelle. L'évolution est plutôt conçue comme la conservation de l'adaptation (Thompson, 2007, p. 204). Autrement dit, dans le cadre énactif, un système vivant adapté qui a évolué n'est pas conçu comme un système qui présente des traits optimaux (retenus par le mécanisme de la sélection naturelle) pour l'environnement où il se trouve, mais plutôt comme un système qui présente une structure viable (et pas nécessairement optimale) qui lui permet de rester en vie. Sous-jacente à cette conception du système vivant adapté qui a évolué est une nouvelle conception de la sélection. Dans le cadre énactif, cette dernière est définie comme un mécanisme qui opère en écartant les propriétés des systèmes vivants qui ne sont pas viables (et non en sélectionnant essentiellement les variations génétiques optimales comme le stipule l'adaptationnisme). Les tenants de l'énactivisme

interprètent donc le processus évolutif selon une logique proscriptive. Ils renoncent, en effet, « à l'idée suivant laquelle la sélection est un processus prescriptif qui guide et commande la fonction d'amélioration de la valeur adaptative » (Varela et al., 1993, p. 264). Ils soutiennent plutôt qu'au cours de l'histoire des systèmes vivants évoluant dans des environnements divers, « la sélection écarte [dans ces systèmes] ce qui n'est pas compatible avec la survie et la reproduction » (Varela et al., 1993, p. 264). Dans le cadre énactif, la sélection opère donc « comme un large filtre de survie admettant toute structure dotée d'une robustesse suffisante pour persister » (Varela et al., 1993, p. 265). La robustesse est ici définie comme la capacité d'un système vivant « de s'automaintenir, de s'autoajuster et de s'auto-organiser face à l'adversité » (Thompson, 2007, p. 195). Quant à elle, la persistance est définie, dans ce cadre évolutionniste particulier, comme une valeur adaptative qui « mesure la probabilité de la permanence reproductive au cours du temps » (Varela et al., 1993, p. 252). Cependant, il est à noter que « [c]e qui est optimisé ici [quand il y a persistance] n'est pas la quantité de descendances, mais la probabilité d'éviter l'extinction » (Varela et al., 1993, p. 252). En somme, selon Varela, Thompson et Rosch, dans le cadre énactif, il faut analyser

le processus évolutif en termes de *satisficing* (c'est-à-dire d'adoption d'une solution sous-optimale, mais néanmoins acceptable, en ce sens qu'elle dépasse un seuil limite de satisfaction) plutôt que d'optimisation [...] (Varela *et al.*, 1993, p. 265)

Maintenant que cinq propriétés essentielles des systèmes vivants ont été présentées, à savoir qu'ils sont des systèmes autopoïétiques (1) opérationnellement clos, (2) thermodynamiquement ouverts, (3) précaires, (4) autonomes et (5) capables d'adaptabilité (une notion qui suppose les concepts de niveau de processus métaboliques/cognitifs, de structure, d'organisation et d'évolution que l'on a aussi vus), on peut aborder une sixième propriété des systèmes vivants, à savoir celle selon laquelle (6) ces derniers sont nécessairement des *systèmes cognitifs* ou, autrement dit, des systèmes qui génèrent de la signification. Un système vivant peut être caractérisé

de la sorte, parce qu'en même temps qu'il s'autoproduit et s'automaintient en tant qu'unité dynamique distincte, émerge chez lui une identité et, au même moment, une forme minimale d'activité cognitive (Weber et Varela, 2002, p. 116). Pour un tel système, cette dernière consiste à distinguer, grâce à ses capacités adaptatives, les choses de son environnement qui sont nécessaires ou qui nuisent au maintien de son existence en tant qu'identité (Di Paolo, 2005). Cette activité cognitive minimale entraine l'émergence d'un monde qui est signifiant pour le système vivant au sens où ce monde se met à contenir des choses que le système vivant identifie (consciemment ou non) comme ayant plus ou moins de valeur *pour* lui. Autrement dit, en raison de l'activité cognitive du système vivant, des valences positives, neutres et négatives apparaissent et se « collent » (si on peut s'exprimer ainsi) aux différentes substances qui sont dans l'environnement. C'est précisément ce que soutiennent Varela et Weber quand ils affirment qu'« avoir un monde [de significations] pour un organisme signifie avant tout avoir des valeurs qui apparaissent au cours du processus d'individuation » (Weber et Varela, 2002, p. 118).

Ce qu'affirment Varela et Weber est saisi par la formulation de la CÉAC de Sanneke de Haan selon laquelle « fabriquer de la signification [sense-making], cette activité d'un organisme en interaction avec son environnement, est la cognition » (de Haan, 2020b, p. 54, mes italiques). Cette conception de la cognition est inhérente à la thèse de la continuité vie-esprit [life-mind continuity thesis] au fondement de l'énactivisme autopoïétique (de Haan, 2020b, p. 55) et qui stipule que « quand il y a de la vie il y a de l'esprit et que l'esprit dans sa forme la plus articulée appartient à la vie » (Thompson, 2007, p. ix). Cette conception de la cognition est inhérente à cette thèse parce que si tous les systèmes vivants doivent, afin de s'adapter et pour rester en vie, distinguer dans leur environnement ce qui est bon et mauvais pour eux, alors il s'en suit que « fabriquer de la signification [c'est-à-dire s'adonner à une activité cognitive] [...] est inhérent à tous les êtres vivants » (de Haan, 2020b, p. 55) et, en conséquence, autant

aux plantes, aux bactéries et aux humains. Enfin, il suit de cette thèse que les processus mentaux et les processus biologiques sont des faces différentes d'un même phénomène. Cela signifie qu'*ils ne sont pas deux types de processus différents*: l'esprit est inhérent aux processus du vivant et les processus du vivant sont inhérents au processus de l'esprit. C'est ce qui explique que les deux niveaux de processus métaboliques dont il a été question précédemment sont aussi considérés, dans le cadre énactif, comme des processus cognitifs.

L'exemple souvent utilisé pour illustrer cette conception de la cognition chez un système vivant non humain met en scène des bactéries, lesquelles sont conçues comme des systèmes autopoïétiques vivants « minimaux » (Varela, 1997, p. 75-76 et surtout 79). Ces dernières sont appréhendées de cette manière parce qu'elles sont des systèmes vivants qui présentent une membrane semi-perméable qui leur permet de se différencier de l'environnement et d'exister en tant qu'unités distinctes. Cette membrane permet aussi aux bactéries d'être en relation avec l'environnement. Cette interaction qu'elles entretiennent avec ce dernier leur permet de rester en vie puisqu'elles se nourrissent de l'énergie qui s'y trouve et y évacuent les déchets qu'elles produisent. L'énergie de l'environnement qu'elles consomment se trouve dans des composés comme le sucre. Pour consommer cette substance, elles doivent être capables de la distinguer d'autres substances de l'environnement qui pourraient nuire ou, à tout le moins, ne pas contribuer au maintien de leur existence. C'est lors de cette activité cognitive élémentaire qu'apparait dans l'environnement — lui-même composé de substances neutres — un monde de significations (ou de valences positives, négatives et neutres) pour chacune des bactéries. Dans le monde de ces dernières, le sucre a une valeur ou une valence positive parce qu'il est de la nourriture *pour* elles. À l'inverse, les virus bactériophages ont une valeur négative pour les bactéries parce qu'ils nuisent au maintien de leur capacité d'autopoïèse. Il suit de cela que le sucre, en tant que bonne source d'énergie, et que les virus bactériophages, en tant qu'agents dangereux, sont des

propriétés relationnelles liées au métabolisme des bactéries évoluant dans cet environnement (Thompson, 2007, p. 158)¹⁰⁵.

Toute cette activité cognitive des bactéries suppose (7) qu'elles ont une *expérience* subjective du monde et d'elles-mêmes. En effet, pour être en mesure de distinguer ce qui est bon ou mauvais pour elles et, conséquemment, d'augmenter leurs chances de conserver leur capacité adaptative et de rester en vie, les bactéries doivent à tout le moins être capables de *percevoir* le monde autour d'elles et de se sentir et se percevoir elles-mêmes.

Il peut sembler qu'il soit exagéré et même étrange de penser que les bactéries ont une expérience subjective du monde et d'elles-mêmes. C'est pourquoi, dans le cas de ce type de systèmes vivants, je propose de parler d'expérience subjective minimale. Bien qu'elle ne l'utilise pas, cette idée d'expérience subjective minimale est bien saisie par la philosophe de Haan quand elle réfléchit au type d'expériences que peuvent avoir les tiques, les papillons ou les plantes. En effet, elle écrit que

[...] [bien qu'il soit] très peu probable que l'expérience de ces organismes soit une expérience consciente d'elle-même et donc aussi riche et complexe que nos expériences [humaines] [...], [ces] êtres vivants ont des sens : ils

105 Il est intéressant de noter que toutes ces considérations énactives sur la bactérie résonnent avec les propos suivants du médecin, biologiste et prix Nobel de physiologie ou médecine François Jacob :

Les êtres vivants ne peuvent survivre, croître et se multiplier que grâce à un flux incessant de matière, d'énergie et d'information. C'est donc une nécessité absolue pour un organisme de percevoir son milieu, ou du moins les aspects de son milieu liés à ses exigences vitales. Le plus simple organisme, la plus humble bactérie doit « savoir » le type de nourriture qui est à sa disposition et ajuster son métabolisme en conséquence. Chez les microorganismes, perception et réaction sont rigoureusement déterminées par les gènes. Elles se réduisent chacune à une alternative, à oui ou non. Tout ce qu'une bactérie peut percevoir, c'est ce que son programme génétique lui permet de déceler au moyen de quelques protéines dont chacune « reconnait » spécifiquement un composé particulier. Pour une bactérie, le monde extérieur se réduit à quelques substances en solution. (Jacob, 1981, p. 98-99)

sont sensibles à certains aspects ou événements de leur environnement, certains aspects ou événements les affectent de certaines manières, et ils y réagissent de certaines manières aussi. *Cette perception est une forme très fondamentale d'expérience et elle a une structure « subjective »*: elle implique que les organismes, en tant qu'unités de perception, incarnent une perspective sur le monde. (de Haan, 2020b, p. 85)

Je pense aussi, suivant de Haan, que cette expérience subjective minimale ne suppose pas nécessairement

[...] qu'il existe un homuncule ou un « organisme » interne qui « fait la détection » ou qui enregistre et évalue le résultat de l'influence de l'environnement. Il n'est pas nécessaire de supposer un tel doublement puisque la capacité de détection est implicite dans la structure autoorganisée, autopoïétique et adaptative qui caractérise les êtres vivants. (de Haan, 2020b, p. 85)

Cette hypothèse selon laquelle tous les systèmes vivants ont à tout le moins une expérience subjective minimale conforte la thèse énactive autopoïétique vue précédemment selon laquelle *la vie est par définition consciente*. En fait, cette thèse est une formulation différente de la thèse de la continuité vie-esprit vue précédemment. Cette dernière est très semblable à la thèse d'Aristote selon laquelle l'âme et le corps sont les deux côtés du processus de vivre (Thompson, 2007, p. 226)¹⁰⁶.

Je vais maintenant porter mon attention sur les humains et noter deux choses que suppose la CÉAC à leur égard. La première est qu'à l'instar des bactéries, et comme tous les systèmes vivants, ceux-ci ont une expérience subjective du monde et d'eux-mêmes. Mais cette dernière n'est généralement pas seulement une expérience subjective minimale parce qu'à l'inverse des bactéries, les humains sont conscients

_

¹⁰⁶ À noter que cette thèse s'oppose à au moins deux thèses cartésiennes : (1) l'esprit et le corps sont des substances et (2) l'esprit peut exister sans un corps (et vice versa) (Thompson, 2007, p. 226).

qu'ils ont une expérience subjective du monde et d'eux-mêmes. Il en est ainsi parce qu'ils sont capables et font preuve de *réflexivité* (de Haan, 2020b, p. 146). Ils possèdent cette capacité parce qu'ils ont un cerveau très développé capable de produire, par l'intermédiaire du langage, une grande complexité conceptuelle. Qu'ils aient cette capacité leur permet, par exemple, de remettre en question le sens de leurs actions ou leurs manières de penser. Cela permet aussi à chacun des humains d'avoir une identité narrative. C'est d'ailleurs ce que dit Varela, un fondateur de l'énactivisme autopoïétique :

On peut dire que ce que nous appelons « je », nous-mêmes, naît des capacités linguistiques récursives de l'homme et de sa capacité unique d'autodescription et de narration. (Varela, 2004, p. 99)

La deuxième chose qu'il faut noter est que, comme les bactéries, tous les êtres humains habitent un monde de significations qu'ils créent ensemble et partagent entre eux. On peut ainsi considérer que (8) ils sont des constructeurs de monde. Ce monde, comme le sucre qui fait partie de celui des bactéries, contribue à leur autoproduction, à la conservation de leur capacité adaptative, à leur automaintenance et, en principe (pour les humains, pas pour les bactéries), à leur épanouissement. Mais le monde de significations des humains, en raison de la sophistication de leurs cerveaux et de la capacité de réflexivité qu'il permet, présente une plus grande complexité conceptuelle que celui des bactéries. C'est ce qui explique qu'ils vivent, à l'inverse de ces dernières, dans un monde de significations qui contient, entre autres choses, des normes sociales qui régulent leurs interactions, des systèmes de valeurs, des systèmes de croyances religieuses, des idéologies politiques, des théories philosophiques, des théories scientifiques et des manières de vivre. Tous ces systèmes d'idées et de pratiques sont des institutions cognitives (voir section 2.1.4 et annexe A). Ce monde de significations des humains contient aussi des artéfacts culturels ou des mécanismes sur lesquels les humains accolent des significations, avec lesquels ils associent des émotions et

auxquels ils attribuent des fonctions. Par exemple, ce film rappelle des moments de l'enfance de tel individu. Quant à cette pièce de musique (un mécanisme sonore), elle fait vivre à ce dernier de la mélancolie. Enfin, les ustensiles permettent de manger sans se salir les doigts et s'en servir d'une certaine manière est un signe, pour tel individu vivant dans la société occidentale, de civilité.

Enfin, il est à noter que, tandis qu'ils fabriquent de la signification, les systèmes vivants — qu'ils soient des bactéries ou des humains — modifient nécessairement leur environnement (physique, social, culturel) et qu'à son tour, au fil du temps, l'environnement modifie les systèmes vivants qui s'y trouvent. Plus précisément, cela signifie que, d'une part, l'activité métabolique/cognitive des systèmes vivants s'accompagne nécessairement de changements dans l'environnement où ces systèmes se développent et que, d'autre part, cet environnement exerce des pressions sur ces systèmes vivants, lesquelles causent éventuellement des changements dans la structure de ces derniers. Ainsi, au fil du temps qui passe, les systèmes vivants et l'environnement se spécifient l'un l'autre (Varela et al., 1993, p. 266) et, pour cette raison, on peut dire qu'ils sont toujours structurellement couplés (Thompson, 2007, p. 45)¹⁰⁷. En effet, «[d]eux ou plusieurs systèmes sont couplés lorsque la conduite de chacun est fonction de la conduite de l'autre » (Thompson, 2007, p. 45). Et quand l'on qualifie ce couplage de « structurel », on « fait référence à l'histoire des interactions récurrentes entre deux ou plusieurs systèmes qui aboutissent à une congruence structurelle entre eux » (Maturana, 1975; Maturana et Varela, 1992, p. 75; Thompson, 2007, p. 45)¹⁰⁸. À noter que les systèmes vivants dont il est question ici ont (9) la

-

 $^{^{107}}$ Thompson (2007, p. 45) écrit en effet : « Un système autonome est toujours structurellement couplé à son environnement. »

¹⁰⁸ La notion de couplage est aussi présentée aux sections 2.1.3 et 2.1.4 et à l'annexe A.

propriété d'être *dynamiques* au sens où ils changent au fil du temps (Shapiro, 2019, p. 146; Thompson, 2007, p. 38-39)¹⁰⁹.

En dernière instance, on peut dire que (10) le système vivant dynamique forme avec son environnement un autre *système dynamique* plus grand. Pour illustrer cette dixième propriété des systèmes vivants, on peut prendre le cas des colonies de bactéries qui vivent dans l'intestin d'une personne, lequel est pour ces dernières un environnement. Ces bactéries influent sur l'état vital d'une personne parce qu'elles opèrent des fonctions essentielles pour elle (cela signifie que sans elles cette personne mourrait). En effet, les communautés de bactéries qui habitent notre tube digestif décomposent les aliments non digestibles, alimentent l'intestin en énergie, fabriquent des vitamines, désagrègent des toxines et des médicaments et entrainent notre système immunitaire (Enders, 2015, p. 190). À son tour, notre intestin influe sur l'état des bactéries qui évoluent en lui. Il permet notamment à plusieurs d'entre elles de rester en vie. En fait,

[p]lus de la moitié des bactéries de notre tube digestif sont tellement habituées à nous qu'elles ne peuvent pas survivre ailleurs. Notre intestin est leur foyer. Elles y sont à l'abri de l'oxygène, elles aiment sa moiteur tiède et apprécient la cuisine prémâchée qu'on y sert. (Enders, 2015, p. 192)

Bien que les bactéries se sentent bien dans nos intestins, elles sont néanmoins soumises à des pressions environnementales diverses. Ces dernières peuvent être ce que l'on mange ou ingère (des médicaments antibiotiques), le stress que l'on vit, le stade développemental que l'on traverse (puberté, vieillesse, etc.) ou une maladie quelconque (Enders, 2015, p. 213). Pour rester en vie, les bactéries dans notre intestin doivent s'adapter à ces pressions environnementales par le biais d'un changement structurel que permet leur organisation. Ces changements de structures entrainent une variation

¹⁰⁹ Voir l'annexe B pour plus d'informations sur les systèmes dynamiques.

de la composition des colonies bactériennes et cette nouvelle composition a un impact sur les êtres humains (immunitaire ou cognitif, par exemple). En d'autres termes, les bactéries et les êtres humains qui sont leurs hôtes se spécifient l'un l'autre au fil du temps. Comme le remarque Enders, «[...] [n]ous modifions parfois notre flore intestinale, et parfois, c'est elle [la flore intestinale ou, et c'est la même chose, les bactéries dans notre intestin] qui nous modifie» (Enders, 2015, p. 213). Les êtres humains sont donc couplés avec ces colonies de bactéries. En effet, on dit que « [d]eux ou plusieurs systèmes sont couplés lorsque la conduite de chacun est fonction de la conduite de l'autre » (Thompson, 2007, p. 45). Bref, les bactéries et les êtres humains sont indissociables et forment un système dynamique.

Les humains forment aussi, à l'instar des bactéries, un système dynamique avec toutes sortes d'entités dynamiques. Par exemple, les humains forment un système dynamique avec leurs institutions cognitives. On a vu, par exemple, que les êtres humains formaient un système dynamique avec le système légal (section 2.1.4). Les humains et le système légal, en effet, se spécifient l'un l'autre au fil du temps. Cela signifie que, selon les réalités et problématiques légales que rencontrent les humains d'une époque, le droit évolue et modifie, en conséquence, la manière dont les humains réfléchissent à ce type d'enjeux. Par exemple, à une époque, il n'existait pas de droit à l'avortement. Maintenant, après que des femmes et des hommes aient montré l'importance de son existence, ce droit d'avoir de l'aide pour mettre un terme à une grossesse est inscrit dans le système légal canadien. On peut penser que ce système légal, comme le devraient toutes les institutions cognitives, contribue au maintien et à l'épanouissement de l'existence de tous les êtres humains liés à lui.

3.1.1 La CÉAC résout les trois premiers problèmes de la conception du trouble mental des RDoC

Maintenant que j'ai présenté les dix propriétés de la CÉAC (laquelle est, je le rappelle, sous-jacente à la CÉOST), j'aimerais expliquer pourquoi la CÉAC résout trois des quatre problèmes (vus au chapitre 2 et énumérés à la section 2.5) que rencontre la conception du trouble mental des RDoC (la CÉOST résout le quatrième problème, comme je le montrerai à la section 3.2.2).

La CÉAC résout le premier problème parce que, contrairement à l'hypothèse de l'identité psychoneurale de Bunge, elle suppose une conception de l'esprit qui est en harmonie avec ce que disent les études empiriques et théoriques sur l'incrustation de l'esprit dans l'environnement et sur sa constitution extraneurale, c'est-à-dire sur sa constitution corporelle, multipersonnelle et cognitivo-institutionnelle (voir les sections 2.1.1-2.1.4 pour prendre connaissance d'un échantillon de ces études). Premièrement, la CÉAC suppose que tout système vivant entretient une très forte interaction avec l'environnement (physique et socioculturel) où il se trouve. En effet, c'est grâce à cette interaction qu'il peut s'automaintenir dans le temps en tant qu'unité dynamique distincte. Ainsi, on ne peut pas concevoir un système vivant sans tenir compte de l'environnement où il évolue. La bactérie, par exemple, entretient une interaction constante avec son environnement avec lequel elle échange de l'énergie. Quant à lui, l'environnement a une influence sur la bactérie. S'il manque de sucre dans ce dernier, par exemple, celle-ci doit s'adapter à cette situation. Comme elle est un système vivant, il en va de même pour la personne humaine. En effet, celle-ci est, tout comme la bactérie, fortement en interaction avec l'environnement qui a une influence sur sa cognition. On a vu, par exemple, à la section 2.1.1, que l'activité cognitive d'une personne était influencée par la température et les bruits ambiants de même que par les jugements des personnes qui l'entourent. On a également vu, à la fin de la section 2.1.1, que le fait de vivre dans un quartier pauvre où le taux d'homicide est élevé faisait

augmenter la prévalence de la dépression majeure chez les personnes âgées. Bref, tous ces résultats empiriques sont prévus par la CÉAC (et ne le sont pas par l'hypothèse de l'identité psychoneurale de Bunge qui, sans être incompatible avec l'idée selon laquelle un système vivant est incrusté dans l'environnement, *ne contient néanmoins pas cette idée*).

Deuxièmement, la CÉAC suppose que l'activité cognitive est constituée de mécanismes corporels. Elle suppose, en d'autres termes, que la cognition est *incarnée*. Selon la CÉAC, la cognition est constituée des actions que pose le système vivant dans l'environnement où il se trouve. Par exemple, le déplacement qu'opère la bactérie dans l'espace pour trouver du sucre est, ultimement, une activité cognitive. C'est pour cela qu'au sein de la CÉAC on définit l'activité cognitive d'un système vivant comme une *énaction*. Il faut ajouter à cela que les actions cognitives que pose le système vivant dans son environnement sont déterminées par les propriétés physiques (ou biologiques ou corporelles) qui caractérisent ce système. Par exemple,

[t]out ce qu'une bactérie peut percevoir, c'est ce que son programme génétique lui permet de déceler au moyen de quelques protéines dont chacune « reconnait » spécifiquement un composé particulier. Pour une bactérie, le monde extérieur se réduit à quelques substances en solution. (Jacob, 1981, p. 98-99)

Tout comme la cognition d'une bactérie, celle d'une personne est indissociable des propriétés de son corps et des actions qu'elle pose. On a vu, par exemple, que les propriétés des mains de Paganini ont influencé sa manière de composer de la musique (section 2.1.2). On a vu également que la manière de penser qu'adoptait Hawking dans le domaine de la physique théorique était influencée par le fait qu'il présentait la sclérose latérale amyotrophique (SLA) (section 2.1.2). On a vu aussi, suivant les résultats d'études empiriques, que l'état du corps d'un système vivant constituait ses états cognitifs. On peut penser, par exemple, à l'influence qu'a le microbiote d'une

personne sur son état mental (section 2.1.2). Bref, comme on peut le constater, à l'inverse de l'hypothèse de l'identité psychoneurale de Bunge, la CÉAC contient cette idée que le corps participe à la cognition des personnes (et des systèmes vivants en général).

Troisièmement, la CÉAC suppose que l'activité cognitive d'un système vivant peut être constituée de mécanismes multipersonnels. Il en est ainsi parce que, selon la CÉAC, le système vivant forme avec son environnement (physique et, éventuellement, social et culturel) un système dynamique. Ce fait s'explique parce que le système vivant et l'environnement où il évolue sont couplés l'un à l'autre et s'autodéterminent mutuellement. On a vu, par exemple, que l'état des bactéries dans l'estomac d'un être humain était influencé par ce qu'il faisait et que son propre état était influencé par l'état des bactéries en lui. Il s'agit ici d'un exemple mettant en scène un environnement biologique. Mais le même principe dynamique s'applique quand les systèmes vivants qui s'autodéterminent sont des personnes. On dira alors que celles-ci forment un système dynamique plus grand que l'on pourrait qualifier de social. Comme on l'a vu à la section 2.1.3, des mécanismes cognitifs comme la régulation des émotions du nourrisson ou la récupération des souvenirs peuvent émerger de ce système dynamique social. Bref, contrairement à l'hypothèse de l'identité psychoneurale de Bunge, la CÉAC saisit aussi dans le monde le fait que l'activité cognitive puisse être constituée de mécanismes multipersonnels.

Quatrièmement, la CÉAC suppose que l'activité cognitive d'un système vivant peut être constituée de mécanismes cognitivo-institutionnels. C'est le cas parce que, selon la CÉAC, le système vivant et l'environnement (physique et, éventuellement, social et culturel) où ce dernier évolue sont couplés l'un à l'autre, s'autodéterminent mutuellement, et forment ainsi tous les deux un système dynamique. Quand le système vivant — en l'occurrence ici une personne — forme un système dynamique avec une

institution cognitive, on dira alors, selon la CÉAC, qu'il forme avec cette dernière un système dynamique que l'on pourrait qualifier de culturel. Par exemple, on a vu que les personnes et le système juridique formaient ensemble un tel type de système (section 2.1.4). Comme plusieurs institutions cognitives existent, plusieurs grands systèmes dynamiques culturels existent et participent à la vie cognitive des personnes. Bref, contrairement à l'hypothèse de l'identité psychoneurale de Bunge, la CÉAC saisit le fait que l'activité cognitive des personnes est constituée par des mécanismes cognitivo-institutionnels (ou par des institutions cognitives pour reprendre l'expression utilisée par Gallagher).

La CÉAC règle le deuxième problème que pose la conception du trouble des RDoC. Contrairement à cette dernière, en effet, elle incite fortement à ce que soit menées des études sur les mécanismes environnementaux (physiques et, éventuellement, sociaux et culturels) favorisant le développement des troubles mentaux et sur les mécanismes corporels, multipersonnels et cognitivo-insitutionnels qui constituent ces derniers. En effet, à l'inverse de l'hypothèse de l'identité psychoneurale de Bunge, comme on vient de le voir, la CÉAC saisit le fait que la cognition est incrustée dans l'environnement et qu'elle est constituée par les mécanismes extraneuraux susmentionnés. La CÉAC motive donc des études empiriques qui examineraient la conception incarnée du TDAH comme celle due à Maiese que l'on a vue à la section 2.1.2. La CÉAC favorise également des projets de recherche qui chercheraient à comprendre dans quelles circonstances un trouble mental chez une personne est potentiellement constitué de mécanismes cognitifs qui n'ont pas pour base cette personne, mais une autre personne (comme dans le cas où une dépression majeure chez une personne serait causée par un mécanisme dysfonctionnel responsable de la régulation émotionnelle dans une autre personne [voir section 2.1.3]). Enfin, la CÉAC ouvre la porte à des recherches sur les différentes institutions cognitives qui pourraient être des constituantes de certains troubles mentaux et qui, ce faisant, causeraient l'augmentation de leur prévalence. Par

exemple, une recherche sur l'institution du genre que Merritt pense être une des constituantes principales du trouble de la dysfonction sexuelle féminine (voir les sections 2.1.4 et 3.2.1) trouverait un milieu favorable au sein de la CÉAC.

Enfin, la CÉAC résout le troisième problème dû au principe epistémoquantitatif sous-jacent à la conception du trouble des RDoC (voir sections 1.5.4, 2.2, 2.2.1 et 2.2.2). En effet, contrairement à ce dernier, la CÉAC favorise que soient menées des études phénoménologiques et qualitatives en santé mentale. En effet, en vertu de la thèse de la continuité vie-esprit inhérente à la CÉAC, l'activité cognitive est invariablement accompagnée d'une expérience subjective du monde et d'une fabrication de significations. Donc, selon la CÉAC, cette conscience phénoménale et cette fabrication de sens sont des facettes de la cognition et ont ainsi une valeur épistémique indéniable. Si on ne les étudie pas, il manquera ainsi des données importantes qui pourraient permettre de contribuer aux connaissances qui aideraient à comprendre l'esprit et ses troubles.

Maintenant que j'ai expliqué pourquoi la CÉAC résout trois problèmes de la conception du trouble mental des RDoC, je présente, à la section 3.2.1, la CÉOST. Comme on le verra à la section 3.2.2, cette dernière résout le quatrième problème (constitué de plusieurs petits problèmes) que rencontre la TBS de Boorse sous-jacente à la conception du trouble des RDoC (voir section 2.3).

3.2 La CÉOST (basée sur la CÉAC) et la solution au quatrième problème de la conception du trouble mental des RDoC

Dans cette partie de ma thèse, je fais deux choses. Premièrement, à la section 3.2.1, je présente la conception énactive et objective de la santé et du trouble (CÉOST). Je construis cette dernière suivant la conception énactive autopoïétique de la cognition (CÉAC) (section 3.1) et en puisant des idées dans les travaux de biologie théorique de

Moreno, Mossio et Saborido (Mossio *et al.*, 2009, 2010; Mossio et Moreno, 2015; Saborido et Moreno, 2015; Saborido *et al.*, 2016). On verra, lors de cette exposition de la CÉOST, que cette dernière permet trois actions. D'abord, elle permet de saisir si une personne est en santé ou si elle présente un trouble physique ou mental. Ensuite, elle permet d'identifier les cas de troubles mentaux ou de dysfonctions mentales où le mécanisme dysfonctionnel se trouve *physiquement à l'extérieur* de l'enveloppe corporelle de la personne malade. Enfin, elle permet de déterminer si un système dynamique autre qu'une personne (comme un couple, une organisation ou une société) est en santé ou malade.

Deuxièmement, à la section 3.2.2, je montre que la CÉOST ne rencontre pas les problèmes de la TBS (sous-jacente au concept de trouble mental des RDoC) (section 2.3) et les problèmes de la DP de Wakefield et de la DCD de Powell et Scarffe (section 2.4.3).

3.2.1 Présentation de la CÉOST

Suivant la conception énactive autopoïétique de la cognition (CÉAC) présentée à la section 3.1, on peut qualifier les personnes (et, également les couples de personnes, les familles, les organisations et les sociétés, comme on le verra à la fin de cette section) de systèmes dynamiques *BioCognitifsSocioCulturels* (ou, pour faire plus court, de systèmes dynamiques BCSC). Deux raisons me semblent pouvoir être avancées pour défendre cette idée. La première est que, comme tous les systèmes dynamiques, les personnes changent d'états avec le temps. La deuxième raison est que, conformément à la thèse de la continuité vie-esprit, les personnes sont *constituées* de systèmes intégrés

de mécanismes biocognitifs¹¹⁰ (ou d'entités dynamiques biocognitives), de certains mécanismes sociaux et de certains mécanismes culturels (ou d'entités dynamiques sociales et culturelles)¹¹¹. En un mot, la CÉAC implique que les personnes se modifient au fil du temps et qu'elles sont *constituées* de divers mécanismes intégrés les uns aux autres et appartenant à la classe des mécanismes BCSC¹¹². Les personnes présentent une composition biocognitive-socioculturelle parce qu'elles sont *couplées* avec des systèmes dynamiques sociaux (comme une personne ou une famille) et des systèmes dynamiques culturels (c.-à-d., des institutions cognitives¹¹³) avec lesquels elles forment de grands systèmes dynamiques (pour la notion de couplage, voir les sections 2.1.3, 2.1.4 et 3.1 de même que l'annexe A).

Comme tous ces mécanismes ou toutes ces entités dynamiques constituant les personnes présentent des *fonctions* et que les personnes elles-mêmes ont des capacités fonctionnelles, je propose, suivant les travaux en biologie théorique de Mossio, Saborido et Moreno (2010, p. 166), qu'une des trois *propriétés essentielles* des

¹¹⁰ Ces mécanismes ou entités biocognitives supposent eux-mêmes ou elles-mêmes des mécanismes ou entités physiques (comme des influx nerveux et des ions). Pour des raisons de lisibilité, toutefois, je ne les intègre pas dans l'acronyme « BCSC ».

¹¹¹ J'utilise les notions de mécanismes et d'entités dynamiques parce que chacune d'elle est utile pour saisir différentes facettes de ce qui est, selon moi, un *même* phénomène dans le monde. Par exemple, les poumons sont des *entités* dynamiques, mais la ventilation pulmonaire est un *mécanisme*.

en même temps biocognitifs, sociaux et culturels. Certains mécanismes qui constituent une personne peuvent être seulement de nature biocognitive. Le mécanisme de la ventilation pulmonaire dont la fonction est de permettre l'échange des gaz vitaux entre une personne et le monde est, par exemple, essentiellement de cette nature. Il peut sembler bizarre de dire que ce mécanisme est à la fois biologique et cognitif (ou biocognitif), mais, dans le cadre de la CÉAC, c'est de cette manière qu'il faut le qualifier. En effet, comme le stipule la thèse de la continuité vie-esprit, « quand il y a de la vie il y a de l'esprit et [...] l'esprit dans sa forme la plus articulée appartient à la vie » (Thompson, 2007, p. ix) (voir section 3.1). Selon cette thèse, par exemple, on pourrait dire que les poumons (des entités dynamiques biocognitives) «perçoivent» l'environnement intérieur (par exemple, la présence d'oxygène dans l'air) et y réagissent.

¹¹³ À noter qu'une institution cognitive émerge d'un système social. L'existence de l'institution cognitive dépend donc de l'existence du système social duquel elle émerge.

personnes soit de présenter (1) une différenciation organisationnelle. Cette dernière réfère au fait que les systèmes dynamiques BCSC, que sont les personnes, contiennent nécessairement (et, plus précisément, produisent et automaintiennent, comme on le verra très bientôt) un ensemble d'entités dynamiques différentes et localisables et un ensemble de mécanismes ¹¹⁴ différents et localisables qui contribuent de manière spécifique et concertée, par l'intégration de leurs différentes fonctions, à l'automaintien de leur organisation et, conséquemment, à leurs capacités fonctionnelles générales ¹¹⁵.

La notion d'organisation correspond ici, conformément à la CÉAC, à toutes les relations *possibles* qui peuvent exister entre les types de composants (comme les cellules) de la personne (Bitbol et Cohen-Varela, 2017, p. 21) (section 3.1). L'existence de cet automaintien de l'organisation de la personne implique que toutes les personnes (comme tous les systèmes vivants) présentent une clôture organisationnelle (ou opérationnelle) (on verra cette notion très bientôt. Voir aussi section 3.1). Enfin, l'automaintien de l'organisation est nécessaire au maintien de l'existence de la personne. En effet, un changement de cette organisation entrainerait une modification trop profonde de la personne et, pour cette raison, sa mort (Escobar, 2012, p. 61-62) (section 3.1).

En plus de présenter la propriété essentielle d'être (1) organisationnellement différenciées, les personnes ont également les propriétés essentielles de (2) s'autoproduire et de (3) s'automaintenir au fil du temps par l'intermédiaire de la

 114 Pour alléger le texte, plutôt que de toujours spécifier que l'on a affaire à des entités dynamiques et à des mécanismes, j'utiliserai dorénavant, la plupart du temps, seulement l'expression « mécanisme ».

¹¹⁵ À l'instar de Saborido et ses collègues (2016, p. 112, note 7), je mets à profit la thèse de Garson (2013, 2017a) selon laquelle les mécanismes servent des fonctions. Celle-ci est utile parce qu'elle implique que les mécanismes opèrent selon une normativité qui leur est inhérente. Cette thèse permet ainsi d'affirmer qu'un mécanisme fonctionne bien ou mal. Enfin, cette thèse exclut qu'il puisse exister des mécanismes pour des désordres. Par exemple, il n'y a pas de mécanismes pour l'arythmie cardiaque. Tout ce que l'on peut dire est qu'un cœur qui fait de l'arythmie est un cœur qui dysfonctionne.

conservation de leur capacité d'adaptation (Di Paolo, 2005, voir aussi la section 3.1 de cette thèse). Par exemple, les personnes consomment des aliments dont elles tirent l'énergie qui leur sert à autoproduire toutes les entités dynamiques fonctionnelles (par exemple, les poumons) et les mécanismes biocognitifs fonctionnels (comme la ventilation pulmonaire) dont elles sont composées. Aussi, s'il vient à manquer de nourriture dans leur environnement, elles se déplacent pour aller s'en procurer (par exemple, à l'épicerie) afin de s'automaintenir. Les personnes se « nourrissent » également des relations sociales qu'elles entretiennent avec d'autres personnes et de tout ce monde culturel symbolique dont elles sont, quelque part, les créatrices et/ou les consommatrices. Cette « nourriture » socioculturelle leur permet, à l'instar des denrées matérielles comme les bananes, le beurre d'arachide et le pain, de s'autoproduire, de s'automaintenir et de persister en tant que systèmes dynamiques autonomes BCSC organisationnellement différenciés.

On peut formuler deux observations qui font la lumière sur les relations dynamiques qu'entretiennent les propriétés (2) et (3). La première observation est que (3) implique (2). En effet, si elles ne s'autoproduisent pas, alors les personnes ne peuvent pas exister en tant que systèmes dynamiques BCSC et, en conséquence, elles ne peuvent pas s'automaintenir et persister dans le temps. Autrement dit, pour s'automaintenir, les personnes doivent d'abord exister en tant que systèmes dynamiques BCSC organisationnellement différenciés.

La deuxième observation est que (2) n'implique pas (3). Il convient d'utiliser le cas des cellules de convection de Bénard pour illustrer cette affirmation. Ces dernières sont des structures dissipatives qui émergent (entre autres) dans l'eau. Pour que ces cellules prennent forme, au moins deux conditions doivent être respectées. La première est que l'eau doit reposer sur une surface chauffante plane qui maintient l'eau à une température constante (par exemple, le fond d'une casserole qui repose sur un rond de

poêle qui est allumé). La deuxième condition est que l'eau reposant sur cette surface doit être située dans un endroit où l'air ambiant est plus froid que l'eau (par exemple, dans une cuisine où la température ambiante est de 20 degrés Celsius).

L'émergence de ces cellules de convection résulte des mécanismes physiques suivants qu'explique très bien le philosophe des sciences (et, entre autres, de la physique) Niall Shanks (voir aussi figure 3.2) :

L'eau réchauffée par le bas monte; en montant, la chaleur [de l'eau qui s'élève] se dissipe, et l'eau se refroidit et commence à redescendre au fond pour être réchauffée, et ainsi peut se répéter le processus [...]. L'eau ne peut pas monter et descendre au même endroit, c'est pourquoi les régions où l'eau monte se différencient des régions où elle descend. C'est cette différenciation qui donne naissance aux cellules. (Shanks, 2004, p. 125, mes italiques)

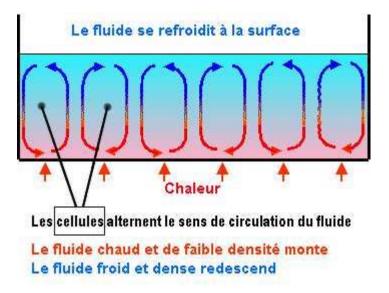


Figure 3.2 Cellules de Bénard vues de côté (Cellules de Bénard, 2021)

Shanks précise également que

[l]es cellules ont un aspect de fossettes, puisque l'eau remonte le long des « parois » de la cellule et s'écoule vers la « fossette » centrale pour redescendre, complétant ainsi la circulation convective. (Shanks, 2004, p. 125)

Quand on les observe d'en haut, ces fossettes ressemblent à des puits miniatures juxtaposés les uns sur les autres dans lesquels on perçoit un mouvement circulaire qui va de bas en haut et de haut en bas et ainsi de suite (voir figure 3.3 à gauche).

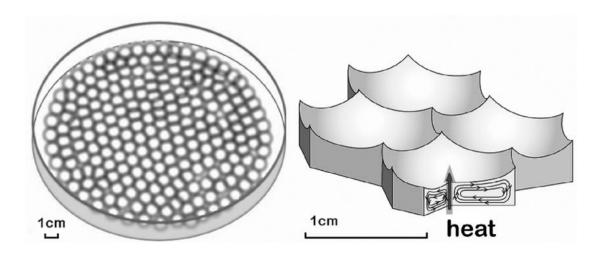


Figure 3.3 Cellules de Bénard vues d'en haut (Deacon et Koutroufinis, 2014, p. 414)

Comme on peut le constater, ces cellules de convection de Bénard faites d'eau s'autoproduisent et, en conséquence, existent en tant que systèmes dynamiques. Mais elles ne s'automaintiennent pas, en ce sens qu'elles n'adaptent pas leurs activités afin de se conserver en tant que systèmes dynamiques en fonction des perturbations externes qu'elles pourraient subir (parmi lesquelles, par exemple, pourrait figurer une baisse de la chaleur émise par la plaque chauffante). De telles cellules de convection, en raison, entre autres, de leur passivité à l'égard de ces perturbations et parce qu'elles ne sont

pas organisationnellement différenciées ¹¹⁶, ne sont pas des personnes (ou, plus généralement, des systèmes dynamiques BCSC).

On peut déduire de l'existence des propriétés (1), (2) et (3) ce que j'appellerais une norme téléologique naturelle (NTN) opérante chez toutes les personnes que je formulerais de cette manière :

NTN: si toutes les personnes ont comme *propriétés essentielles* de faire (2) et (3), alors toutes les personnes *sont censées* être composées de *mécanismes BCSC* qui (a) servent (2) et (3) et, conséquemment, (1) et qui, ainsi, (b) contribuent à la persistance¹¹⁷ des personnes en tant que systèmes dynamiques autonomes¹¹⁸ organisationnellement différenciés¹¹⁹.

les cellules de Bénard (ici faites d'eau) ne sont pas organisationnellement différenciées ou, autrement dit, sont organisationnellement *in* différenciées parce que les constituants microscopiques matériels qui les composent — à savoir les milliards de molécules d'eau — n'ont pas de fonctions. On peut dire que ces molécules d'eau n'ont pas de fonctions parce qu'*elles font toutes la même chose*, à savoir contribuer à générer la cellule de Bénard par l'intermédiaire du mouvement rotatif qu'elles adoptent et qui dépend de ce qu'elles se réchauffent ou de ce qu'elles se refroidissent. Il est intéressant de noter qu'ultimement, les molécules d'eau sont soumises à la gravité. En effet, parce qu'elles deviennent moins lourdes quand elles sont chauffées, elles montent et, parce qu'elles s'alourdissent lorsqu'elles se refroidissent, elles redescendent.

¹¹⁷ Dans le cadre de la NTN, il faut appréhender le concept de persistance comme une notion ontologique. Toutefois, je crois que cette notion de persistance n'est pas contenue dans les théories de la persistance que sont celle de la perdurance et celle de l'endurance (pour un aperçu de ces dernières, voir Baron et Miller, 2019, chapitre 7). Celles-ci, en effet, supposent une ontologie de choses [thing ontology] (Meincke, 2019), alors que la CEOST suppose une ontologie des processus. Dans le cadre de cette dernière, contrairement à ce que prévoit une ontologie de choses, la persistance de l'identité d'un système requiert que s'opèrent en lui des changements (Meincke, 2019, p. 24). C'est justement ce que prévoit la NTN. En effet, c'est parce que des mécanismes sont actifs dans une personne qu'elle peut maintenir son organisation et, par là même, son identité (dynamique). Il faut noter, par ailleurs, que cette notion ontologique de persistance contenue dans la NTN est compatible avec la notion énactive et évolutionniste de la persistance (voir section 3.1). En effet, l'acte de se reproduire est un événement dans une chaîne d'événements qui permet aux systèmes vivants de conserver leur organisation en la transmettant à d'autres systèmes vivants qui différeront d'eux quant à leur structure (voir section 3.2.2). ¹¹⁸ Je précise qu'il s'agit d'un système autonome (au sens énactif). Mais j'aurais pu ne pas le préciser parce qu'un système qui s'automaintient est, par définition, dans le cadre énactif autopoïétique, un système autonome (voir section 3.1).

l'19 À noter que la NTN, que je tire des points (1), (2) et (3), lesquels sont eux-mêmes issus de la CÉAC et des idées de Moreno, Mossio et Saborido, n'est pas originale. En effet, elle s'accorde avec l'idée de Schrödinger (rapportée par Lane) selon laquelle « la vie résiste en quelque sorte à la tendance universelle

Comme on peut le constater, les propriétés (2) et (3) produisent la propriété (1), à savoir la *différenciation organisationnelle* de la personne. En effet, c'est parce qu'elle s'autoproduit et s'automaintient que la personne peut rester organisationnellement différenciée. Cependant, à son tour, (1) rend possible (2) et (3). En effet, c'est parce qu'elle est organisationnellement différenciée qu'elle peut s'autoproduire et s'automaintenir. On est donc ici en présence d'une sorte de boucle causale rétroactive entre ces trois propriétés qui forment une sorte de système dynamique. Cette boucle causale permet l'émergence de ce que Varela appelle la clôture organisationnelle (ou opérationnelle) qui est ici comprise en termes de processus ou de mécanismes BCSC¹²⁰¹²¹ (Varela, 1989, p. 45; Weber et Varela, 2002, p. 115, voir aussi section 3.1 de cette thèse).

C'est à l'aide de la NTN que l'on peut construire une conception énactive et objective de la santé et du trouble (CÉOST) applicable à toutes les personnes (ce n'est que plus tard, dans cette section, que je dirai ce qu'est spécifiquement un trouble de santé mentale et ce que pourrait potentiellement être un trouble d'un couple, d'une organisation ou d'une société).

Il convient, d'abord, de prendre connaissance de la *conception énactive et objective de la santé* (CÉOS) :

à la décomposition, à l'entropie croissante ([au] désordre) qui est stipulée par la deuxième loi de la thermodynamique » (Lane, 2015, p. 22 suivant Schrödinger, 1944) La thèse selon laquelle il existe une normativité dans le monde du vivant allant à l'encontre de l'entropie est bien exposée et défendue par Bolton et Gillett (2019, chapitre 2).

¹²⁰ Varela utilise le terme « processus ». J'utilise pour ma part le terme « mécanisme ». Je ne fais pas de différence conceptuelle entre ces termes.

¹²¹ Il est intéressant de noter que cette notion ontologique qu'est la clôture organisationnelle est d'autant plus plausible qu'elle est en harmonie avec la thèse métaphysique selon laquelle « [c]haque propriété se joint à d'autres propriétés [ou], [e]n d'autres termes, [à la thèse métaphysique] selon laquelle les propriétés sont regroupées en lots ou en systèmes» (Bunge, 2006, p. 12).

CÉOS: une personne *en santé s'automaintient*, grâce aux mécanismes BCSC qui la *composent* et qui servent la NTN, *dans sa zone de viabilité* et, en conséquence, *persiste* en tant que système dynamique autonome organisationnellement différencié.

Cette conception énactive de la santé est *objective* parce qu'elle est relative à la NTN (vue précédemment) qui est *inhérente* à toutes les personnes¹²². En ce sens, la santé d'une personne ne dépend pas d'une évaluation faite, suivant des critères socioculturels, par une observatrice ou un observateur qui lui est extérieur. La santé est plutôt un *état* (ou un ensemble de propriétés) d'une personne qui se maintient dans le temps et que l'on ne fait qu'observer¹²³. Cet état résulte des activités concertées et spécifiques des mécanismes BCSC qui fonctionnent bien et qui composent une personne.

Conformément à la CÉOS, on dira qu'un mécanisme BCSC fonctionne bien s'il sert la NTN et, en conséquence, contribue à l'automaintien et à la persistance de la personne. Par exemple, les deux reins d'une personne qui synthétisent certaines hormones, éliminent les toxines du sang et régulent la pression sanguine opèrent une telle contribution. Un autre exemple est celui des parents qui, dans le système dynamique « nourrisson-parent », contribuent à réguler les émotions de leur enfant (voir section 2.1.3 pour plus de détails à ce sujet).

Comme le stipule la CÉOS, pour contribuer à cette capacité d'automaintenance et de persistance, le mécanisme (ou l'entité dynamique) BCSC en question doit être une

¹²² À noter que la NTN se manifeste *différemment* dans chacune des personnes. Il en est ainsi parce que les personnes sont à la fois semblables et différentes. Par exemple, certaines personnes ont besoin de moins de sommeil que d'autres. D'autres métabolisent mieux le glucose ou les lipides. Mais toutes ces personnes peuvent être en santé au sens où elles persistent dans le temps en tant que systèmes organisationnellement différenciés.

¹²³ À noter, cependant, que la santé, en plus d'être un état, est également une *capacité* naturelle parce qu'elle est un état qu'un système BCSC est *disposé* à automaintenir dans des situations qui ne sont pas présentes. Bref, en plus d'être un état, la santé est également une propriété dispositionnelle.

composante fonctionnelle de la personne (comme la paire de reins de la personne ou comme les parents dans le système dynamique « nourrisson-parents »).

Pour déterminer si un mécanisme ou une entité dynamique est une composante fonctionnelle d'une personne, on a besoin d'une théorie des attributions fonctionnelles. Parmi celles que l'on retrouve dans la littérature aujourd'hui ¹²⁴, l'approche organisationnelle des fonctions (AOF) de Mossio, Moreno et Saborido (Mossio *et al.*, 2009, 2010; Mossio et Moreno, 2015, chapitre 3) me semble être celle qui conviendrait le mieux pour la CÉOST. En effet, à l'instar des tenants de la CÉAC (que suppose la CÉOST), ces philosophes construisent l'AOF dans un cadre théorique où la vie est conçue de manière métabolique et où les systèmes vivants sont appréhendés comme des systèmes dynamiques qui s'automaintiennent et qui sont, conséquemment, autonomes (pour une synthèse des travaux en biologie théorique de ces auteurs, voir Mossio et Moreno, 2015). Voici l'AOF:

Dans AOF, un type de trait T a une fonction si et seulement s'il est soumis à [une] clôture organisationnelle C dans un système automaintenu différencié S. Cette définition implique de satisfaire trois conditions différentes. Ainsi, un trait T a une fonction si et seulement si :

- C1: T contribue au maintien de l'organisation O de S;
- C2 : T est produit et maintenu sous quelques contraintes exercées par O;
- C3 : S est organisationnellement différencié. (Mossio et al., 2010, p. 168)

Toutefois, pour construire la CÉOST, il faut *modifier* l'AOF. En effet, la version originale de l'AOF ne suppose pas la CÉAC (présentée à la section 3.1). La CÉOST contient aussi les concepts « système dynamique BCSC », « personne » et « mécanisme

¹²⁴ Pour une excellente revue critique de cette littérature, voir le livre *A Critical Overview of Biological Functions* (Garson, 2016).

BCSC » que ne contient pas l'AOF. Cela précisé, voici l'approche organisationnelle des fonctions *modifiée* (AOFm) :

Dans l'AOFm, un mécanisme BCSC (ou une entité dynamique) quelconque M a une fonction si et seulement s'il est soumis à [une] clôture organisationnelle C dans une personne automaintenue différenciée quelconque P (laquelle est en dernière instance un système dynamique BCSC). Cette définition implique de satisfaire trois conditions différentes. Ainsi, un mécanisme M a une fonction si et seulement si :

C1 : M contribue au maintien de l'organisation O de P;

C2 : M est [...] maintenu¹²⁵ par l'activité de divers mécanismes BCSC de l'O de P¹²⁶;

C3 : P est organisationnellement différenciée.

(Formulation modifiée, selon la CÉAC, de l'AOF de Mossio *et al.*, 2010, p. 168)

Voici des exemples qui illustrent la manière dont s'applique l'AOFm :

(1) Le cœur (M) qui pompe le sang d'une personne (P) est une composante fonctionnelle de P parce que (C1) M contribue à l'automaintien de P (si le cœur ne pompe pas le sang de la personne, alors cette dernière mourra) (C2) M est maintenu par cette personne (s'il était retiré d'elle et posé au sol, le cœur

¹²⁵ Dans les formulations de 2009 et de 2010 de l'AOF, C2 commande que le trait (et ici dans l'AOFm le mécanisme ou l'entité dynamique) soit produit *et* maintenu par les contraintes exercées par O. Dans la formulation de 2015 de l'AOF, C2 mentionne *seulement* que le trait doit être maintenu (Mossio et Moreno, 2015, p. 73). J'adopte la formulation de 2015 parce que je trouve les formulations de 2009 et de 2010 trop restrictives. En effet, comme l'a bien vu Rioux (2014), ces formulations de l'AOF ne *saisissent pas* le fait avéré que les bactéries dans l'estomac des systèmes vivants (comme les personnes) ont des fonctions. En effet, si les bactéries sont maintenues par l'organisme, elles ne sont en revanche pas produites par lui. En conséquence, la condition C2 de l'AOF n'est pas respectée.

¹²⁶ Dans l'AOF, C2 dit que «T est produit et maintenu sous quelques *contraintes* exercées par O» (mes *italiques*). Pour ma part, je ne dis pas que M (ou T) est maintenu par des *contraintes* de O parce que la notion de clôture, dans le cadre de la CÉAC, est conçue en termes de *mécanismes* ou de *processus* et non en termes de *contraintes* comme dans l'AOF (pour la notion de clôture en termes de contraintes, voir Mossio et Moreno, 2015, chapitre 1).

s'arrêterait de battre et, éventuellement, cesserait d'exister en tant qu'entité dynamique biocognitive) et (C3) P est organisationnellement différenciée. En effet, P est constituée d'un ensemble organisé de mécanismes BCSC qui présentent des fonctions spécifiques intégrées les unes aux autres qui soutiennent l'O de P et qui permettent ainsi à P de persister en tant que système dynamique autonome BCSC organisationnellement différencié.

- (2) Le système de croyances religieuses (M) (une institution cognitive), qui détermine le sens qu'a la vie pour une personne (P) couplée à M, est une composante fonctionnelle de cette personne parce que (C1) M contribue à l'automaintien de P en donnant un sens à sa vie (peut-être que sans ce sens que lui procure M, P développerait une dépression majeure sévère et se suiciderait), (C2) M est (entre autres) maintenu par l'activité de P (par exemple, par l'intermédiaire des rituels religieux auxquels participe P ou par l'intermédiaire du sens que donne P à ses actions quotidiennes) et (C3) P est organisationnellement différenciée. En effet, P est constituée d'un ensemble organisé de mécanismes BCSC (dont ceux dans M) qui présentent des fonctions spécifiques qui automaintiennent l'O de P et qui, conséquemment, permettent P persister tant que système dynamique autonome organisationnellement différencié.
- (3) Les mécanismes mentaux (M) de Jean, qui a une tendance dépressive, sont une composante fonctionnelle des mécanismes mentaux de Jeanne (P), qui a des problèmes de mémoire, parce que (C1) les M de Jean contribuent au maintien de P en contribuant au maintien (de la performance) des mécanismes mentaux (mnémoniques) de P (cela permet, par exemple, à P d'être plus autonome), (C2) les mécanismes mentaux M (responsables de l'humeur) de Jean sont maintenus par les mécanismes mentaux de P (sans la présence de P, en effet, Jean

tomberait en dépression majeure) et (C3) P est organisationnellement différenciée. Une des choses qui permet à Jeanne de persister en tant que système dynamique autonome BCSC organisationnellement différencié est qu'elle et Jean forment un couple qu'autant elle que Jean automaintiennent à travers le temps¹²⁷.

(4) L'air ambiant (M) qui permet la ventilation pulmonaire chez une personne P *n'est pas* une composante fonctionnelle de P parce que M (et l'oxygène que contient M), bien qu'il contribue au maintien de P, n'est pas maintenu par P. Cependant, cela ne signifie pas que M n'a pas une incidence *causale* et vitale sur P. Au contraire, il en a assurément une. Par exemple, si l'air est pollué, alors le risque pour P de développer des maladies respiratoires est augmenté¹²⁸.

Garson affirme que, dans le cadre de l'AOF, ce qu'il appelle le « système panique » peut être conçu *en lui-même* (c'est-à-dire appréhendé indépendamment de tout système) comme un système organisationnellement différencié qui s'automaintient dans le temps. Il pense qu'il est permis de soutenir cela parce qu'on peut dire, dans ce cadre, que les différentes parties du système panique comme l'hypervigilance aux sensations physiques, les fausses croyances et la crise de panique elle-même (lesquelles sont toutes corrélées à des mécanismes neuraux formant un système) ont différentes fonctions qui contribuent à l'automaintien du système panique (lequel peut être lui-même conçu comme ayant des fonctions, comme celle de causer une crise de panique). Garson pense que le fait que l'AOF permet de dire que les parties du système panique et le système panique lui-même ont des fonctions est problématique parce que les parties de ce système et ce système lui-même, dans les faits, n'ont pas de fonctions. En effet, ce système et ses parties sont clairement dysfonctionnels. C'est sur la base de ce raisonnement que Garson juge que l'AOF doit être abandonnée.

À première vue, j'ai l'impression que l'AOFm n'est pas affectée par cette critique parce qu'elle s'applique aux parties des systèmes BCSC et que le système panique n'est pas un système BCSC. En

¹²⁷ On peut penser que les personnes ne sont pas le seul type de systèmes dynamiques BCSC. En effet, comme je le suggérerai à la fin de cette section, les couples, les familles, les organisations et les sociétés peuvent aussi être conçus comme des systèmes dynamiques BCSC.

¹²⁸ Au terme de cette exposition de l'AOFm, j'aimerais au passage montrer que cette dernière échappe aux critiques que Garson (2017b) a formulées à l'endroit de l'AOF. Je souhaiterais également montrer que son argument est problématique dans le cadre de la conception énactive autopoïétique de la cognition (CÉAC) dans laquelle s'inscrit la CÉOST. Avant d'expliquer pourquoi je soutiens cela, il convient de résumer sa critique de l'AOF.

À noter, pour finir cette présentation de la CÉOS, qu'une personne en santé *présente* une bonne robustesse, c'est-à-dire qu'elle présente une bonne capacité à s'automaintenir et à persister malgré les perturbations internes et externes qu'elle rencontre ou qu'elle pourrait rencontrer. Par exemple, une personne non diabétique aurait plus de chance de survivre à la maladie à coronavirus (COVID-19) (une perturbation interne) qu'une personne âgée et diabétique. Un autre exemple est qu'une personne en bonne santé mentale a plus de chance de bien gérer le stress au travail (perturbation externe) qu'une personne qui présente un profil psychologique très anxieux¹²⁹.

effet, le système panique n'est pas constitué de mécanismes BCSC, mais seulement de mécanismes biocognitifs (voir la note de bas de page numéro 112 pour cette notion). En conséquence, l'AOFm ne capture pas les parties du système panique dans le cas où celui-ci est appréhendé en lui-même (comme le fait Garson). L'AOFm saisit plutôt le système panique (et ses parties) dans le cas où on le considère comme une partie d'une personne. Cela s'accorde avec le fait que la CÉOST (qui contient l'AOFm) est une conception de la santé et du trouble des personnes (et plus généralement des systèmes dynamiques BCSC) et pas des systèmes dynamiques composant ces derniers : ces systèmes dynamiques sont fonctionnels ou dysfonctionnels (comme le serait le système biocognitif causant des crises de paniques), mais on ne peut pas dire d'eux qu'ils sont en santé ou qu'ils présentent un trouble : seulement les personnes (et plus généralement les systèmes dynamiques BCSC) peuvent être en santé ou présenter un trouble.

Par ailleurs, même si, pour les besoins de l'argumentaire, on modifiait l'AOFm (et donc la CÉOST) afin d'étendre son domaine d'application pour qu'elle capture le système panique en lui-même, j'ai l'intuition que l'argument de Garson serait aussi problématique. En effet, il me semble qu'au sein de la CÉAC dans laquelle s'inscrit la CÉOST (et l'AOFm), on ne peut pas affirmer que le système panique est un système organisationnellement différencié qui s'automaintient. Il en est ainsi parce que le système cérébral qui cause des crises de panique détruit continuellement ses propres parties au fil du temps. Par exemple, en raison du taux élevé de cortisol qu'ont les gens présentant un trouble panique, l'apoptose neuronale dans l'hippocampe de ces derniers est très importante et, en conséquence, cette dernière rapetisse (Gao et al., 2014). Il me semble que ce fait va à l'encontre de l'esprit général de la CÉAC qui suppose qu'un système contribue à l'automaintien et à la persistance des systèmes qui le composent et contribuent eux-mêmes à leur tour à l'automaintien de ce système (voir section 3.1). Bref, suivant la CÉAC, parce qu'il s'autodétruit pour exister, le système panique est (même si on le conçoit en lui-même) un système dysfonctionnel.

¹²⁹ Cette thèse selon laquelle les personnes en santé sont plus robustes semble aussi s'appliquer à d'autres types de systèmes dynamiques BCSC. Par exemple, un couple dont sont élevés les niveaux d'intimité, d'engagement et de désir qu'entretiennent entre elles les personnes qui le composent a plus de chance de sortir indemne d'une importante épreuve de la vie (comme une maladie chronique grave que pourrait

Maintenant que l'on connait la CÉOS, on peut aborder la *conception énactive et objective du trouble* (CÉOT). Cette conception du *trouble* est (a) *formellement* presque identique à la conception de la *malfonction biologique organisationnelle* (CMBO) de Saborido et Moreno (2015, p. 93), mais (b) n'a pas le même *contenu* que cette dernière. En effet, comme pour l'AOF, il faut modifier la CMBO afin de l'intégrer dans la CÉOST. Avant de présenter la CÉOT, il faut prendre connaissance de la CMBO :

[La CMBO] implique quatre conditions pour l'imputation des malfonctionnements organisationnels aux traits biologiques :

C1 : Un trait biologique T a la fonction F, c'est-à-dire que l'organisation O d'un organisme S « présuppose »¹³⁰ que T exerce l'effet spécifique F (selon une gamme d'activités).

C2: Dans une circonstance donnée, certaines conditions internes ou externes perturbent l'intégration fonctionnelle de T dans O, c'est-à-dire que T ne peut pas effectuer F selon la gamme d'activités présupposées par l'organisme dans son ensemble.

C3 : En conséquence de cette défaillance, une réaction adaptative est déclenchée par le (sous-)système régulateur de S.

C4 : Toutefois, cette défaillance n'est pas entièrement compensée par la réaction adaptative. (Saborido et Moreno, 2015, p. 93)

Voici maintenant la CÉOT:

La CÉOT implique quatre conditions pour l'imputation d'un *trouble* à une personne P (laquelle est ultimement un système dynamique BCSC) :

développer l'enfant de ce couple). Dans le même ordre d'idées, un pays où l'État de droit est fort a plus de chance qu'un pays où ce dernier est faible de résister à la montée en son territoire de mouvements extrémistes qui visent à le faire tomber.

¹³⁰ Saborido et Moreno utilisent la conjugaison anglaise «presupposes». C'est pourquoi j'utilise la traduction «présuppose» dans la citation. Je spécifie cela parce que l'utilisation de ce verbe peut sembler curieuse.

C1: Un mécanisme (BCSC) M a la fonction F, c'est-à-dire que l'organisation O d'une personne P « présuppose » que M exerce l'effet spécifique F (selon une gamme d'activités).

En d'autres termes, C1 dit que M est une composante fonctionnelle de P¹³¹.

C2 : Dans une circonstance donnée, certaines conditions internes ou externes perturbent l'intégration fonctionnelle de M dans l'O de P, c'est-à-dire que M ne peut pas effectuer F selon la gamme d'activités présupposées [pour le maintien de l'O] de P dans son ensemble.

On peut aussi dire ici que M fonctionne mal parce qu'il est incapable d'opérer selon la NTN¹³² inhérente à P et nécessaire au maintien de l'O de P.¹³³.

C3 : En conséquence de cette défaillance ou de ce dysfonctionnement de M, une réaction adaptative est déclenchée par des mécanismes BCSC de régulation de P.

Cette réaction adaptative est déclenchée afin que les mécanismes BCSC de P servent la NTN et, conséquemment, jugulent les dommages que cause (ou causera) le dysfonctionnement de M à l'O de P.

C4 : Toutefois, cette défaillance n'est pas entièrement compensée par la réaction adaptative.

Quand les conditions C1, C2, C3 et C4 sont remplies, l'O de P cesse graduellement (ou très rapidement ou même subitement) d'être maintenue. Il suit de cela que M *empêche* la persistance de P en tant que système dynamique autonome BCSC organisationnellement différencié. En conséquence, une désorganisation systémique ou une augmentation continue de l'entropie, lente ou rapide selon la nature du trouble, s'opère en P. Par exemple, on peut penser qu'un lymphome non hodgkinien (LNH)

¹³² La NTN est la *norme téléologique naturelle* vue précédemment. Elle dit ceci : si toutes les personnes ont comme *propriétés essentielles* de faire (2) et (3), alors toutes les personnes *sont censées* être composées de *mécanismes BCSC* qui (a) servent (2) et (3) et, conséquemment, (1) et qui, ainsi, (b) contribuent à la persistance des personnes en tant que systèmes dynamiques autonomes organisationnellement différenciés.

¹³¹ M satisfait donc toutes les conditions de l'AOFm que l'on a vues précédemment.

¹³³ Par exemple, chez une personne, une glande thyroïde qui ne produit pas assez d'hormones thyroïdiennes fonctionne mal puisqu'elle nuit au maintien de l'O de P. Un autre exemple est celui d'une personne dont les idéations culpabilisantes et suicidaires (un mécanisme biocognitif) sont envahissantes.

indolent (qui évolue lentement) non traité cause une augmentation de l'entropie beaucoup moins importante qu'un cancer agressif du pancréas non traité.

Généralement, la personne qui développe un trouble présente éventuellement des *symptômes* qui sont causés par cette désorganisation ou cette augmentation de l'entropie sous-jacente et incessante en elle (qui, si elle n'est pas traitée ou est intraitable, mène à sa dissolution). Par exemple, les plaques amyloïdes et les amas de protéines Tau qui se forment dans le cerveau d'une personne et le détruisent sont parmi les mécanismes dysfonctionnels sous-jacents au *trouble* d'Alzheimer *et* sont accompagnés de symptômes cognitifs comme la perte de mémoire, des difficultés d'apprentissage et des difficultés à planifier. Un autre exemple serait que l'institution cognitive de l'autonomie personnelle¹³⁴, le fait de vivre dans une zone de guerre et le dysfonctionnement des systèmes monoaminergiques figurent hypothétiquement tous parmi les mécanismes dysfonctionnels BCSC sous-jacents à un *trouble* de dépression majeure létale *et* sont accompagnés de symptômes comme une humeur dépressive, une insomnie ou une hypersomnie, une agitation psychomotrice, une difficulté à se concentrer, un sentiment de dévalorisation ou de culpabilité excessive ou inappropriée et des pensées suicidaires¹³⁵.

Il est à noter, en outre, qu'un trouble, chez une personne, s'accompagne généralement par une baisse *significative* et *continue* ¹³⁶ de sa robustesse, c'est-à-dire par une

¹³⁴ L'institution cognitive de l'autonomie personnelle des sociétés libérales réfère à la conception kantienne de l'autonomie qui suppose que les humains sont des agents parfaitement rationnels qui s'autosuffisent et qui opèrent dans un vide social. Cette conception de l'autonomie est très différente de la conception relationnelle de l'autonomie qui suppose que les humains sont des « agents socialement et historiquement situés, métaphysiquement non isolés et [...] façonnés par des facteurs comme la race et les classes [sociales] » (Stoljar, 2018, section 1).

¹³⁵ Ces symptômes sont listés dans le DSM-5 (American Psychiatric Association, 2013, p. 161-162).

¹³⁶ Je précise que cette baisse de robustesse qui accompagne un trouble est *significative* et *continue* pour la distinguer de la baisse de robustesse qu'accompagne une dysfonction qui est compensée par des

diminution significative et continue de sa capacité à s'automaintenir et à persister malgré des perturbations internes et externes qu'elle pourrait rencontrer. Par exemple, une personne alitée en phase terminale d'un cancer du côlon est plus susceptible de mourir d'une diarrhée nosocomiale, due à *Clostridium difficile*, qu'une personne en bonne santé ou qu'une personne présentant une dysfonction quelconque qui peut être compensée par des mécanismes de régulation BCSC.

Si on résume, on pourrait dire que, selon la CÉOT, une personne présente un *trouble* si son *état* fait en sorte qu'elle *est incapable de s'automaintenir dans sa zone de viabilité et, conséquemment, de persister* en tant que système dynamique autonome organisationnellement différencié (quand elle présente un tel état, la personne tend donc inexorablement vers une *in*différenciation organisationnelle). Cet état de la personne qui engendre cette incapacité est attribuable (a) au mauvais fonctionnement des mécanismes BCSC qui composent cette dernière et (b) à l'échec des mécanismes régulateurs qui s'activent en elle pour contrer ce dysfonctionnement.

Cette conception énactive du trouble est *objective*. En effet, ultimement, c'est parce que les mécanismes BCSC ne servent pas la NTN inhérente à la personne que cette dernière présente un trouble. En ce sens, l'existence d'un trouble ne dépend pas d'une évaluation faite, suivant des critères socioculturels, par une observatrice ou un observateur extérieur à cette personne. Un trouble est plutôt un *état naturel* d'une personne qui se dégrade que l'on ne fait que constater.

Pour finir, il est intéressant d'observer que, selon la CÉOT, le fait qu'un mécanisme fonctionne mal dans un système BCSC *n'implique pas* nécessairement le

mécanismes BCSC. La présence d'un trouble chez une personne indique donc que cette dernière est

incapable de développer un nouveau couplage structurel avec l'environnement afin de rester viable.

développement d'un trouble (j'emprunte cette idée à Saborido et Moreno, 2015, p. 88-89 qui l'appliquent, quant à eux, aux systèmes biologiques et qui diraient plutôt, selon leur terminologie, qu'une défaillance fonctionnelle n'implique pas nécessairement une malfonction biologique organisationnelle). En effet, il se peut très bien que la personne (consciemment ou non) et/ou l'entourage de cette personne mettent en œuvre des mécanismes compensatoires de régulation qui (a) permettent à cette dernière de se maintenir dans une zone de viabilité plus restreinte et qui (b) préviennent conséquemment l'apparition d'un trouble chez elle. Un cas qui illustre bien cela est celui d'une personne qui fait de l'hypothyroïdie, qui prend chaque jour des hormones de remplacement et qui continue de vivre normalement. Un autre exemple pertinent est celui de quelqu'un qui perd ses deux bras lors d'un accident, mais qui, grâce à une réorganisation de son environnement, à du support social et à l'utilisation de prothèses, réussit néanmoins à mener une vie épanouissante. Enfin, on peut penser à une personne anxieuse qui réussit à conserver son autonomie en structurant sa vie autour de personnes qui l'acceptent comme elle est et autour d'activités qui ne lui causent pas trop de stress.

Il est intéressant de souligner que ces cas où des mécanismes de régulations compensatoires sont mis en place par ces personnes et donnent lieu à un nouveau régime de couplage structurel entre elles et leur environnement peuvent être conçus comme des *cas d'adaptation*. En effet, selon la conception énactive de l'évolution, une personne (et plus généralement un système vivant) adaptée est *une personne qui présente une structure viable* (et *pas nécessairement optimale*) qui lui permet de rester en vie (voir section 3.1).

Comme on peut le constater, cette interprétation énactive de l'adaptation est différente de celle contenue dans l'interprétation adaptationniste de la théorie biostatistique (TBS) de la santé (sous-jacente au concept de trouble mental des RDoC) qui stipule qu'une

personne (et plus généralement un système vivant) adaptée est *une personne qui* présente des traits optimaux (retenus par le mécanisme de la sélection naturelle) pour l'environnement où elle se trouve et, en dernière instance, pour l'environnement dans lequel ses ancêtres ont évolué (à sujet, voir aussi les sections 1.5.3 et 3.1).

Si la présence d'une dysfonction chez une personne n'implique pas nécessairement la présence d'un trouble et peut être l'occasion d'une adaptation (au sens énactif), il faut toutefois noter deux choses. La première est qu'il est important, si l'on veut prévenir l'apparition d'un trouble (et maintenir l'existence de la personne), d'intervenir sur les dysfonctions qui affectent cette dernière (c'est d'ailleurs ce que font les personnes mises en scène dans les exemples vus précédemment). La deuxième est que la présence d'une dysfonction dans une personne, si elle ne cause pas toujours un trouble, s'accompagne généralement d'une baisse de sa robustesse, laquelle est une diminution de sa capacité de s'automaintenir et de persister malgré des perturbations internes et externes qu'elles pourraient rencontrer. Par exemple, si une personne qui fait de l'hypothyroïdie se trouve dans le désert et y perd ses hormones de remplacement, elle pourrait mourir.

Maintenant que la CÉOS et la CÉOT ont été présentées, on peut définir et circonscrire de manière plus pointue ce qu'on appelle aujourd'hui un trouble *mental*. Pour ce faire, il faut préciser deux choses.

La première chose à souligner est que plusieurs des conditions que l'on appelle en ce moment des *troubles* mentaux ne sont pas, selon la CÉOT, des *troubles*. Elles sont plutôt *seulement* des *dysfonctions* mentales. Il en est ainsi parce qu'à l'heure actuelle, plusieurs des conditions que l'on qualifie de troubles mentaux n'entrainent pas systématiquement une dissolution de la personne comme un cancer foudroyant du pancréas peut en entrainer une. Par exemple, la phobie sociale, la kleptomanie ou le

trouble de thésaurisation ne semblent pas létaux. Cela ne signifie pas, cependant, que l'on ne doit pas intervenir sur ces dysfonctions mentales. Au contraire, il est important que la personne elle-même et/ou les différents systèmes dynamiques sociaux où évolue cette personne et/ou les professionnelles et professionnels de la santé mentale et/ou la société (par l'intermédiaire de mesures de santé publique ou de changements idéologiques) interviennent sur ces dysfonctions afin que soit établi un couplage nouveau entre cette personne et son environnement qui permettrait à cette dernière de s'automaintenir dans une nouvelle zone de viabilité.

La deuxième chose à mentionner est qu'un trouble mental et une dysfonction *mentale* ne sont pas définis comme un trouble et une dysfonction *physique*. Graham (2013) a développé un critère *épistémique* servant à distinguer ces deux types de troubles ou ces deux types de dysfonctions (j'ajoute cette distinction que Graham ne fait pas entre « trouble » et « dysfonction » parce que la CÉOT la fait). Selon lui, quand le trouble (ou la dysfonction) est de type mental, les symptômes sont *expliqués* par un *mixte de causes* (ou, selon la CÉOT, par un mixte de mécanismes ou d'entités dynamiques), lesquelles pourraient être qualifiées, pour être conformes avec ce que j'ai dit dans cette section, de biocognitives, de sociales et de culturelles. Par exemple,

si une dépression est causée, en partie, par l'expérience stressante d'une perte d'emploi [(un mécanisme social)], une perception de déchéance conjugale [(un mécanisme biocognitif)] ou par une situation d'aliénation professionnelle [(un mécanisme social)], et si, plus généralement, [l'explication de cette dépression] exige [...] [de faire] référence partielle[ment] à des [...][mécanismes] [...] psychologiques [...] ou [relevant de l'intentionnalité] ¹³⁷ [...], alors [...] il s'agit d'un trouble mental [ou d'une dysfonction mentale]. (Graham, 2013, p. 36)

¹³⁷ Je pense que la notion d'intentionnalité qu'évoque ici Graham n'est pas incompatible avec l'énactivisme autopoïétique. En effet, selon ce que je comprends des propos de Gallagher, bien que leur

Quand le trouble ou la dysfonction est de type physique, les symptômes sont expliqués essentiellement par des mécanismes corporels ou neuraux ou, selon la CÉOT, par des mécanismes biocognitifs ou des entités dynamiques biocognitives qui ont lieu essentiellement à l'intérieur de la personne. Par exemple, si une dépression majeure sévère est causée par un accident vasculaire cérébral ou par la maladie de Huntington, cette dépression est alors un aspect d'un trouble physique ou, plus précisément, d'un trouble neurologique (qu'on peut aussi qualifier, selon la CÉOT, de trouble biocognitif) (Graham, 2013, p. 36). Pour les mêmes raisons, le diabète de type 2 est également défini comme un trouble physique. En effet, si une personne présente cette pathologie, alors les symptômes¹³⁸ de faim, de somnolence, d'augmentation du volume et de la fréquence des urines, de soif intense, de bouche sèche, de vision embrouillée, de faim exagérée et d'irritabilité qu'elle pourrait présenter sont tous essentiellement expliqués (si on fait l'hypothèse que cette personne n'a aucune autre pathologie qui pourrait expliquer ces symptômes) par les mécanismes biocognitifs de la régulation de la glycémie qui ne fonctionnent pas bien en elle. Un tel dysfonctionnement peut être initié lorsque les récepteurs cellulaires membranaires à l'insuline deviennent insensibles à cette hormone et qu'en conséquence le glucose n'entre plus assez dans les cellules et s'accumule dans le sang. On est bel et bien ici face à un dysfonctionnement parce qu'avoir trop de glucose dans le sang cause, entre autres, chez la personne une inflammation systémique. En outre, le fait que le glucose n'entre pas assez dans les cellules a pour résultat que ces dernières sont sous-alimentées. Pour contrer ce dysfonctionnement qui nuit à la viabilité de la personne, le pancréas de cette dernière déclenche une hyperinsulinémie (c'est-à-dire une plus grande production d'insuline)

_

contenu soit dérivé et *soit dépendant des activités* qu'opère le système vivant dans son environnement (physique, social et culturel), les états mentaux (comme les croyances) *existent* néanmoins (Gallagher, 2013, p. 11; voir aussi Gallagher et Miyahara, 2012, p. 139). En conséquence, les états mentaux «*sont inexplicables à moins de faire référence à de telles activités»* (Gallagher, 2013, p. 11, mes *italiques*). ¹³⁸ Ces symptômes sont listés dans le site Internet *Diabète Québec* (2021).

laquelle peut être conçue comme une réaction adaptative à cette situation. Mais, éventuellement, si l'on n'intervient pas ou si l'on ne peut tout simplement plus intervenir, le pancréas de la personne s'épuise à force de fonctionner à plein régime et devient incapable de produire suffisamment d'insuline, ce qui cause éventuellement une hyperglycémie et les complications systémiques et possiblement létales qu'elle peut engendrer.

Je montre maintenant que la CÉOST (et, plus particulièrement, la CÉOT) peut rendre compte de cas de troubles mentaux ou de dysfonctions mentales où le mécanisme dysfonctionnel se trouve *physiquement à l'extérieur* de l'enveloppe corporelle de la personne *tout en faisant partie de cette dernière* lorsqu'on la conçoit comme un système dynamique BCSC.

Pour démontrer cela, je vais mettre à profit le cas de figure, vu à la fin de la section 2.1.3, où, m'inspirant de Bray (2008), je fais l'hypothèse qu'un des mécanismes dysfonctionnels de la régulation des émotions (M) causant une dépression majeure chez une personne P1 ne se trouve pas *physiquement* dans P1, mais dans une autre personne P2 qui présente un trouble de la personnalité limite.

Ce cas suppose que P1 et P2 sont couplées (au sens énactif) et forment un système dynamique social. Elles sont couplées parce que « la conduite de [l'une] est fonction de la conduite de l'autre » (Thompson, 2007, p. 45). Autrement dit, P1 et P2 s'autodéterminent l'une et l'autre. Par ailleurs, elles forment un système dynamique parce que l'état du couple que forment P1 et P2 change avec le temps.

Ce cas suppose, en outre, que M de P2 qui est couplée à P1 pourrait, *en principe*, dans un monde possible, bien fonctionner et être une composante fonctionnelle de P1. En effet, pour être en droit d'affirmer que M de P2 est une composante *dysfonctionnelle* de P1, il faut que l'on puisse concevoir que M *pourrait* être une composante

fonctionnelle de P1. Cette hypothèse me semble censée. En effet, il est plausible de penser que les membres d'un couple s'entraident à réguler mutuellement leurs émotions. Dans le présent exemple, cela signifie plus spécifiquement que P1 et P2 qui forment un couple (et un système dynamique) peuvent en principe s'entraider dans la régulation mutuelle de leurs émotions. Ainsi, en principe, P2, grâce à M, peut aider P1 à réguler ses émotions.

Cependant, dans notre exemple, M dysfonctionne chez P2 qui est couplé à P1. Et qui plus est, ce dysfonctionnement de M est si important que P2 se sert des mécanismes mentaux de P1 pour gérer ses propres émotions, et ce, sans tenir compte des émotions de P1. On peut penser que cette dynamique relationnelle peut causer potentiellement une détresse mentale chez P1 et, éventuellement, si elle perdure, une dépression majeure sévère de P1 menant à son suicide.

Dans cette histoire tragique, on serait donc en présence d'un trouble mental de P1 dont au moins une des composantes dysfonctionnelles M est physiquement à l'extérieur de l'enveloppe corporelle de P1 (en l'occurrence, dans celle de P2 qui forme avec P1 un système dynamique).

Je pense que l'on peut potentiellement rendre compte de ce cas, dans le cadre de la CÉOT, de la manière suivante :

C1 : Un mécanisme M de P2 a la fonction F de réguler à l'occasion les émotions de P1, c'est-à-dire que l'organisation O de P1 « présuppose » que M (de P2) exerce à l'occasion l'effet spécifique F (selon une gamme d'activités).

C2 : Non seulement M de P2 n'aide plus à l'occasion P1 à réguler ses émotions, mais il ne régule plus que très peu les émotions de P2. En conséquence, P2 se sert des mécanismes mentaux de P1 pour gérer ses propres émotions, et ce, sans

tenir compte des besoins (disons émotifs) de P1. Ainsi, dans ce cas de figure, une composante physiquement externe à l'enveloppe corporelle de P1 (à savoir M de P2), mais qui fait, néanmoins, partie du système BCSC qu'est P1 (P1 et P2 sont couplées, en effet), perturbe l'intégration fonctionnelle de M (de P2) dans l'O de P1. Autrement dit, non seulement M ne peut pas réguler les émotions, selon la gamme d'activités présupposées par P1 dans son ensemble, mais il empiète sur le bon fonctionnement des mécanismes mentaux de P1. Par exemple, P1 se sent peu à peu aliénée d'elle-même.

C3 : En conséquence de cette défaillance de M de P2 (lequel fait partie du système dynamique BCSC qu'est P1), une réaction adaptative est déclenchée par des mécanismes de régulation de P1. Cette dernière tente, par exemple, d'exprimer à P2 les limites de l'aide qu'elle peut lui apporter dans la gestion de ses émotions.

C4 : Toutefois, cette défaillance n'est pas entièrement compensée par la réaction adaptative de P1. En effet, P2 se met en colère, tient un discours culpabilisant et fait du chantage à P1. Bref, cette situation conjugale dégénère.

Parce que les conditions C1, C2, C3 et C4 sont respectées, M (de P2) *empêche* la persistance de P1 en tant que système dynamique autonome BCSC organisationnellement différencié. Par exemple, P1 qui était déjà biocognitivement vulnérable à la dépression majeure se met à avoir des idéations suicidaires et se tue.

Cette conclusion funeste, bien qu'elle soit improbable, n'est pas impossible. Toutefois, on peut penser qu'il est plus vraisemblable que P1 continuera à chercher, par l'intermédiaire de mécanismes de régulation compensatoires, une manière de retrouver

une nouvelle zone de viabilité. Par exemple, P1 pourrait quitter P2. Aussi, P2 aidée par P1 pourrait réussir à améliorer sa capacité à gérer ses émotions.

Pour montrer qu'une dysfonction d'un trouble mental peut se trouver *physiquement à l'extérieur* de l'enveloppe corporelle d'une personne, on peut également utiliser le cas de figure, vu à la fin de la section 2.1.4, où est présentée l'hypothèse de Merrit (2013) selon laquelle un mécanisme BCSC dysfonctionnel M de l'institution des genres (masculin et féminin) I est une des composantes de la dysfonction sexuelle féminine chez une femme P.

Cet exemple suppose que I et P sont couplées (au sens énactif) et forment un système dynamique. Elles sont couplées parce que l'état et les propriétés de l'une sont fonction de l'état et des propriétés de l'autre (Thompson, 2007, p. 45). Autrement dit, I et P s'autodéterminent l'un l'autre. Par ailleurs, I et P forment un système dynamique parce que l'état du système que forment I et P change avec le temps.

Cet exemple suppose également que le M de I pourrait, *en principe*, être un mécanisme (culturel) *fonctionnel* pour le maintien de l'organisation O de P. On peut faire l'hypothèse qu'un mécanisme culturel du genre féminin qui fonctionne bien est un mécanisme qui donne une représentation réaliste des femmes et de leur vie sexuelle. On peut supposer que cette représentation réaliste repose sur des recherches empiriques quantitatives et qualitatives sérieuses menées en psychologie et en sexologie, et non sur des stéréotypes que véhiculent certains médias et certaines idéologies du moment sur le genre féminin. Dans cette perspective, le M de I pourrait, en principe, être une composante *fonctionnelle* de P. M pourrait, par exemple, contribuer à l'épanouissement de la vie sexuelle de P.

Cela posé, on peut maintenant rendre compte de la dysfonction sexuelle féminine de P, dans le cadre de la CÉOT, de la manière suivante :

C1 : Un mécanisme culturel M de I a la fonction F de contribuer positivement à la vie sexuelle (autant dans ses aspects physiques que psychologiques) de P, c'est-à-dire que l'organisation O de P « présuppose » que M exerce l'effet spécifique F (selon une gamme d'activités).

C2 : En raison des images stéréotypées que les médias dominants véhiculent de la femme, l'intégration fonctionnelle de M de I dans l'O de P est perturbée. En effet, M de I ne peut pas effectuer F selon la gamme d'activités présupposées par la personne P dans son ensemble. Il peut, par exemple, résulter de ce dysfonctionnement de M de I de la tristesse chez P.

C3 : En conséquence de ce dysfonctionnement de M de I, une réaction adaptative est déclenchée par des mécanismes de régulation compensatoires de P. Par exemple, P tente de s'expliquer, d'accepter et de normaliser le fait qu'elle n'a pas une très grande libido ou qu'elle n'a pas d'orgasme lors de toutes ses relations sexuelles.

C4 : Toutefois, cette défaillance n'est pas entièrement compensée par la réaction adaptative. Par exemple, étant toujours envahie par les images stéréotypées du genre féminin véhiculées par I, P se met à se dévaloriser et développe une souffrance mentale importante.

Parce que les conditions C1, C2, C3 et C4 sont respectées, le M de I *empêche* la persistance de P en tant que système dynamique autonome BCSC organisationnellement différencié. Par exemple, P parce qu'elle se dévalorise trop finit par s'enlever la vie.

Comme on peut le constater, dans ce cas de figure funeste et improbable, mais possible, un M de I *constituant* le trouble de la dysfonction féminine sexuelle chez P est *physiquement à l'extérieur* de l'enveloppe corporelle de P.

Toutefois, suivant la CÉOT, on peut penser que la dysfonction sexuelle féminine, puisqu'elle n'est généralement pas létale, n'est habituellement pas un *trouble* mental (comme dans l'exemple précédent où la fin est tragique), mais plutôt une *dysfonction* mentale sur laquelle on peut intervenir. Une des interventions qui pourrait être effectuée est de changer cette institution cognitive malsaine du genre pour une institution saine.

Une question immédiate que fait poindre à l'esprit ce cas de figure est celle-ci : comment fait-on exactement pour déterminer de manière *objective* qu'une institution cognitive (ici celle du genre) et, plus généralement, la société dont émerge cette dernière fonctionnent bien ou mal? J'ai suggéré rapidement qu'on pourrait déterminer cela suivant les résultats d'études scientifiques pertinentes. Par exemple, j'ai proposé qu'une institution du genre qui est bonne soit une institution du genre qui véhicule, par l'intermédiaire de diverses représentations et pratiques, ce que la science dit du genre féminin.

Mais ce critère n'est pas suffisant parce qu'il n'est pas objectif. En effet, il repose sur des jugements de valeur comme celui selon lequel, pour les humains, il est *bon* de construire nos institutions cognitives suivant ce que dit la science¹³⁹.

¹³⁹ J'aimerais noter que cet énoncé normatif me semble, néanmoins, parfaitement censé.

La CÉOST peut mettre fin à cette situation. En effet, elle semble contenir des outils conceptuels qui permettent de déterminer objectivement et naturellement pourquoi une société (ou tout autre système dynamique de type BCSC) fonctionne bien ou mal. En effet, suivant la NTN (que contient la CÉOST), une société (et les institutions cognitives qui en émergent) fonctionne bien dans la mesure où elle contribue à l'automaintenance et à la persistance des systèmes dynamiques BCSC qui la composent et contribuent eux-mêmes à leur tour à l'automaintenance de cette société. Ainsi, suivant la NTN, une société qui véhicule une conception stéréotypée et erronée des femmes et de leur sexualité et qui, conséquemment, nuit à l'automaintenance de ces dernières, est une société (et plus précisément une institution cognitive) qui contient une dysfonction. Pour les mêmes raisons, une société où domine une idéologie fermée aux différentes orientations et identités sexuelles, et où, en conséquence, sont opprimées les personnes lesbiennes, gaies, bisexuelles, transgenres et queers (LGBTQ), fonctionne également mal.

Ces considérations normatives basées sur la NTN supposent l'hypothèse selon laquelle les couples de personnes, les organisations et les sociétés (et les institutions cognitives qui émergent de ces organisations et de ces sociétés) sont des systèmes dynamiques de types BCSC qui présentent les propriétés d'être (1) organisationnellement différenciés, de (2) s'autoproduire et de (3) s'automaintenir afin de persister dans le temps.

Sans être certaine, cette hypothèse me semble *intuitivement* plausible ¹⁴⁰. On peut prendre le cas d'une compagnie de biscuits pour le démontrer. Celle-ci est organisationnellement différenciée. En effet, elle contient des entités dynamiques

¹⁴⁰ Je ne suis pas le seul à faire ce genre d'hypothèse. En fait, cette dernière n'est pas neuve. Par exemple, déjà en 1975, Joël de Rosnay (1975, p. 48-60) comparait les villes et les entreprises à des organismes vivants.

comme des ouvrières et ouvriers et des départements divers (production, finance, ressources humaines, marketing, recherche et développement, etc.). Ces entités dynamiques contribuent de manière spécifique et concertée, par l'intégration de leurs différentes fonctions, à l'autoproduction et à l'automaintien de la compagnie de biscuits. Par exemple, cette dernière ajuste le salaire des employées et employés en fonction des besoins de ces derniers et en fonction du marché. Aussi, cette compagnie offre de bonnes conditions de travail à ses employées et employés afin qu'ils soient heureux et restent en emploi. En outre, cette compagnie peut réagir à la compétition que lui fait une autre compagnie en changeant la technologie qu'elle utilise pour faire des biscuits et en formant ses employées et employés afin qu'ils puissent opérer cette nouvelle technologie. En somme, par ces différentes fonctions, la compagnie s'autoproduit, s'automaintient et, en conséquence, persiste dans le temps en tant que système BCSC organisationnellement différencié.

Parce qu'ils semblent posséder les propriétés (1), (2) et (3), on peut penser que ces types de systèmes dynamiques BCSC comme les sociétés, les organisations ou les couples ont, à l'instar des personnes, une clôture organisationnelle. Il semble possible d'appliquer cette notion de clôture organisationnelle aux sociétés, aux organisations et aux couples parce que, comme l'a dit Thompson (2007, p. 44), la frontière produite par cette clôture et qui distingue l'extérieur et l'intérieur du système n'est pas nécessairement une membrane (comme celle d'une cellule), mais peut aussi prendre d'autres formes (voir section 3.1). Elle pourrait, entre autres, être faite de mécanismes sociaux. Thompson donne l'exemple d'une frontière qu'établit une colonie d'insectes entre elle et son environnement (Thompson, 2007, p. 44). Dans le même ordre d'idée, des mécanismes sociaux, culturels et économiques pourraient constituer une frontière entre un système dynamique BCSC (comme une société ou une famille) et ce qui lui est extérieur et avec lequel il entretient une relation de couplage (au sens énactif). Par exemple, une compagnie quelconque, en payant ces employées et employés, retient ces

derniers en elle et engendre (peut-être) chez eux un sentiment d'appartenance à son endroit. Ces employées et employés, quant à eux, entretiennent des liens professionnels. Il se peut même qu'ils partagent certaines valeurs de la compagnie. Pour tout cela, il se crée une frontière entre ceux qui travaillent pour cette compagnie et ceux qui n'y travaillent pas. Plus généralement, on pourrait dire qu'il se développe une distinction entre l'intérieur de la compagnie et ce qui lui est extérieur. Une fois cette distinction établie, une relation de couplage s'installe entre la compagnie et l'environnement où elle évolue.

Enfin, on pourrait *distinguer* ces différents *types* de systèmes BCSC selon leurs capacités fonctionnelles respectives. Par exemple, les mécanismes biocognitifs dans une personne peuvent digérer des aliments, mais les mécanismes sociaux et culturels d'une société sont incapables d'accomplir une telle tâche. En revanche, les mécanismes sociaux et culturels d'une société peuvent produire et automaintenir un système de croyances religieuses et les pratiques et rituels qu'elles génèrent, alors que les mécanismes d'une *seule* personne ne le peuvent pas.

Une critique que l'on pourrait faire à ce qui vient d'être dit au sujet des systèmes dynamiques BCSC est qu'il existe des sociétés qui perdurent malgré le fait qu'elles nuisent à l'automaintenance de beaucoup des personnes qui évoluent en elles. Les sociétés totalitaires sont, par exemple, des sociétés de ce type qui font souffrir les personnes qui y vivent.

On peut répondre à cette objection que si *toutes* les personnes qui constituent ce type de société fonctionnaient selon les normes et idéaux que suivent personnellement et qu'appliquent les dictateurs qui dirigent ce type de société, alors la société constituée

par ces personnes cesserait d'exister¹⁴¹. En effet, si tout un chacun avait pour objectif de soumettre par la force son prochain à des idéaux totalitaires homogénéisant, tant sur le plan de sa vie privée que sur celui de sa vie publique, et d'ainsi l'aliéner de lui-même, il est fort probable que l'organisation de cette société se dissoudrait. On peut donc penser qu'une telle société tend naturellement vers une *in*différenciation organisationnelle et qu'en conséquence elle est dysfonctionnelle. On pourrait dire, en effet, qu'ultimement, la *seule* fonction des personnes composant un système totalitaire est de maintenir le système totalitaire à leurs dépens (un peu comme la *seule* fonction des composantes matérielles d'une flamme est de produire cette flamme). Conséquemment, ce genre de société devrait cesser graduellement de s'autoproduire et de s'automaintenir en tant que système dynamique BCSC organisationnellement différencié.

On pourrait contre-argumenter que certaines sociétés totalitaires ont duré *très longtemps* et qu'il est difficile, pour cette raison, de concevoir qu'elles tendent naturellement vers une *in* différenciation organisationnelle qui les propulserait vers leur extinction.

On pourrait répondre à cela, suivant une métaphore approximative qui me semble, par ailleurs, avoir une certaine valeur heuristique, qu'une société totalitaire qui dure très longtemps est un peu comme une étoile : elle est une structure dissipative d'énergie qui est *morte-vivante*. En effet, bien qu'elle présente une complexité matérielle, l'étoile ne présente pas de différentions organisationnelles. De fait, toutes les composantes matérielles qui la constituent présentent la même fonction. Cependant, contrairement

¹⁴¹ Pour formuler cet argument, je m'inspire ici d'une idée formulée par Nielsen et Ward (2020a, p. 120). Ceux-ci disent qu'une société composée essentiellement de personnes présentant un trouble de la personnalité narcissique, un trouble de la personnalité limite ou un trouble de la personnalité antisociale ne pourrait pas protéger les intérêts et l'automaintien de ses membres.

aux cellules de Bénard d'eau (elles aussi des structures dissipatives) qui peuvent se former dans une casserole posée sur une plaque chauffante, une étoile perdure très longtemps.

Parce que les sociétés totalitaires peuvent être conçues de manière métaphorique comme des structures dissipatives mortes-vivantes, ce n'est pas un hasard si elles ont souvent vu émerger en leur sein des mouvements de révolte menés par des personnes porteuses d'idéaux humanistes, de justice sociale et d'ouverture à la diversité qui cherchaient à les faire tomber. Ce faisant, ces personnes tentaient de réinstaurer en elles une tendance *vivifiante* ayant pour objectif la résurgence d'une différenciation organisationnelle que l'on pourrait concevoir en termes *sociologiques* comme une résurgence de la *protection de la diversité intégrée des modes fonctionnels d'existence* des personnes.

On pourrait rétorquer à cela qu'il existe aussi des révoltes dans les systèmes qui, à première vue, fonctionnent bien et sont bons (ou, à tout le moins, ne sont pas mauvais) pour les personnes qui les composent. Pourquoi, par exemple, y a-t-il des groupes de personnes qui s'élèvent contre l'État de droit (lequel, on peut le penser, favorise la différenciation organisationnelle ou, en d'autres termes, comme cela est dit précédemment, la protection de la diversité des modes fonctionnels d'existences des personnes)?

On pourrait répondre à cela que, dans tous les systèmes dynamiques BCSC, en raison de perturbations internes ou externes inévitables, surgissent éventuellement toujours des mécanismes dysfonctionnels. Comme on l'a vu précédemment, lors de la présentation de la CÉOT, ces dysfonctions peuvent faire *grosso modo* deux choses. La première est qu'elles peuvent causer une augmentation de l'*in*différenciation organisationnelle dans un (bon) système dynamique BCSC (comme une société qui

contient un État de droit) et y engendrer, en conséquence, un trouble. La montée du nazisme en Allemagne au 20^e est un exemple de ce cas de figure.

La deuxième chose que ces dysfonctions peuvent faire est de causer une adaptation (au sens énactif vu précédemment) que l'on pourrait interpréter en termes sociologiques comme un changement social ou culturel positif. Il se peut, en effet, qu'un système qui est généralement bon, ou, à tout le moins, qui n'est pas mauvais pour les personnes qui le composent, puisse néanmoins changer de structure. Une réforme du droit qui vise à mieux appréhender les enjeux juridiques que posent les réalités d'une époque pourrait illustrer ce genre de cas de figure. En effet, cette réforme pourrait être conçue comme une adaptation (et même comme une évolution).

Cette idée prévue par la CÉOT et selon laquelle un changement social, et même une évolution sociale, peut être provoqué par des dysfonctions n'est pas nouvelle. Par exemple, en 1968, le sociologue Merton affirmait que «[l]e concept de dysfonctionnement, qui implique le concept de contrainte, de stress et de tension sur le plan structurel, fournit une approche analytique pour l'étude de la dynamique et du changement. » (Merton, 1968, p. 107, cité dans Koch 2020, p. 65)

Il est intéressant de noter que cette thèse semble toujours pertinente aujourd'hui. Par exemple, récemment, le philosophe Koch a affirmé que « [...] les dysfonctionnements peuvent, en général, être des facteurs pertinents pour expliquer le changement social et la déstabilisation des entités sociales ou même des sociétés dans leur ensemble » (Koch 2020, p. 45).

Cela termine mon exposition de la CÉOST et de la manière dont elle peut être appliquée aux personnes, mais aussi, selon certaines de mes intuitions qui viennent d'être exposées, potentiellement à d'autres types de systèmes dynamiques BCSC comme les

couples, les familles, les organisations et les sociétés. Comme on peut le constater, ces idées intuitives montrent qu'adopter la CÉOST implique de soutenir un certain nombre d'idées plutôt inattendues et moins communes en psychiatrie traditionnelle. Elle permet, par exemple, d'avancer la thèse surprenante selon laquelle une société quelconque présente un trouble. Cela dit, que l'on accepte ou non les exemples spéculatifs que j'ai donnés, il est toutefois clair que, si l'on adopte la CÉOST, la recherche en psychiatrie devra alors, non seulement s'intéresser aux aspects psychobiologiques et phénoménologiques des troubles mentaux, mais aussi, entre autres, aux structures sociales et aux institutions culturelles dans lesquelles évoluent les individus. Elle pourrait, par exemple, s'intéresser aux liens qui existent entre le niveau de bien-être des individus d'une société et les systèmes politiques qui y sont opérants.

3.2.2 La CÉOST résout le quatrième problème de la conception du trouble mental des RDoC

Aux sections 2.3 et 2.4.3, on a vu plusieurs problèmes qui montrent que la théorie biostatistique (TBS) de la santé (sous-jacente au concept de trouble des RDoC), la dysfonction préjudiciable (DP) et la dysfonction correctement dévalorisée (DCD) ne peuvent pas saisir la normativité inhérente aux traits (ou aux mécanismes ou aux entités dynamiques) dans le monde du vivant. Conséquemment, on a également vu que ces dernières ont de la difficulté à saisir les traits dysfonctionnels (ou les mécanismes dysfonctionnels ou les entités dynamiques dysfonctionnelles) dans une personne.

Dans cette section, je soutiens que la CÉOST règle tous ces problèmes et, en conséquence, qu'elle saisit les dysfonctions des traits, des mécanismes ou des entités dynamiques dans une personne (et aussi, comme on l'a vu à la fin de la section 3.2.1, dans les systèmes dynamiques BCSC autres qu'une personne comme une famille, une organisation et une société). Avant d'aborder les problèmes de la DP et de la DCD et

les réponses que leur offre la CÉOST, je vais commencer par examiner les problèmes que rencontre la TBS et la manière dont la CÉOST les résout.

Le premier problème de la TBS est celui de la détermination de la classe de référence. Ce problème peut être formulé ainsi : il n'est pas possible de déterminer de manière strictement empirique la classe de référence à la lumière de laquelle on juge si le fonctionnement d'un mécanisme d'un membre de cette classe est normal ou anormal. En effet, la détermination des classes de référence doit absolument faire appel à des considérations non biologiques. Plus précisément, elle requiert de faire appel à des considérations normatives non naturelles. Ce problème implique que la TBS ne peut pas appréhender de manière objective la normativité dans la nature et, en conséquence, qu'elle n'est pas une conception objective du trouble (contrairement à ce que prétend Boorse).

La CÉOST n'a pas ce problème. En effet, suivant cette dernière, le fonctionnement d'un mécanisme d'un système n'est pas évalué en le comparant au fonctionnement statistiquement typique du même mécanisme présent dans les systèmes d'une classe de référence à laquelle appartient ce système, mais plutôt selon une *norme téléologique naturelle* (NTN) inhérente à ce système et se manifestant de manière particulière en lui (voir section 3.2.1)¹⁴². Autrement dit, le fonctionnement d'un mécanisme d'un système est évalué à la lumière de la NTN qui est opérante *dans le système lui-même* et non en le comparant au fonctionnement d'un mécanisme que présentent d'*autres* systèmes qui lui ressemblent. La CÉOST permet donc de formuler une conception *objective* du trouble. En effet, comme cette norme téléologique est naturelle, le jugement opéré par

¹⁴² Il est à noter que, pour les mêmes raisons, la conception de la malfonction biologique organisationnelle dont je m'inspire fortement pour construire la CÉOST ne présente également pas ce problème (Saborido et Moreno, 2015, p. 94; Saborido et al., 2016, p. 115).

une personne selon lequel un mécanisme fonctionne bien ou non n'est à aucun instant influencé par des considérations normatives non naturelles (comme c'est le cas, à l'inverse, suivant le cadre de la TBS, lorsque des considérations normatives sont mobilisées par une personne pour déterminer une classe de référence).

Un deuxième problème que rencontre la TBS concerne les fonctions atypiques. Une fonction atypique est une fonction qui est rarement active chez un trait (ou un mécanisme) qui a plus d'une fonction. La fonction atypique pose un problème à la TBS parce que cette dernière est incapable de la saisir et, en conséquence, parce qu'elle est incapable de dire si elle fonctionne normalement ou anormalement. La TBS est incapable d'identifier une telle fonction parce que, pour elle, une fonction est ce que fait *typiquement* (ou le plus souvent) un trait *quand* il opère pour la survie et la reproduction d'un système vivant.

La CÉOST ne présente pas ce problème. En effet, pour elle, une fonction *est* tout ce que fait un trait pour servir l'automaintien et la persistance du système (en d'autres mots pour servir la NTN). Elle saisit alors *toutes* les fonctions d'un trait ou d'un mécanisme servant la NTN (et non pas, comme le fait la TBS, seulement la ou les fonctions qu'il fait le plus fréquemment). En conséquence, la CÉOST saisit les fonctions atypiques et leurs dysfonctions.

Pour illustrer concrètement cette supériorité de la CÉOST sur la TBS, il convient d'examiner à nouveau l'exemple de Garson (2019b, p. 135) concernant l'urètre chez l'homme (section 2.3). Cette dernière a pour fonction d'uriner (F1) et de transporter le sperme (F2). La TBS saisit seulement F1 parce que, quand l'urètre fait quelque chose pour la survie et la reproduction, c'est ce qu'elle fait le plus souvent (comparativement à F2). À l'inverse, la CÉOST saisit F1 et F2 (et, en conséquence, leurs potentiels dysfonctionnements). Il en est ainsi parce que les fonctions de l'urètre d'uriner et de

transporter le sperme se conforment aux exigences de la NTN inhérente à chacun des hommes. De fait, ces fonctions contribuent à la capacité d'automaintien et, conséquemment, de persistance de chacun de ces derniers. En effet, si un homme ne peut pas du tout uriner, il mourra. Quant à la fonction de transporter le sperme, elle sert elle aussi cette capacité d'automaintien et de persistance. En effet, le transport du sperme (et des spermatozoïdes que contient ce liquide) sert à la reproduction, laquelle est, selon la conception énactive autopoïétique de la cognition (CÉAC), un comportement qui permet au système dynamique BCSC (dans ce cas-ci un homme) de (1) transmettre son organisation à un autre système (plus précisément à sa progéniture qui différera de lui sur le plan de sa structure, mais pas en ce qui concerne son organisation) et ainsi (2) de s'automaintenir en conservant sa capacité d'adaptation et, conséquemment, de persistance par l'intermédiaire d'une histoire où il est lié aux systèmes qui le précèdent et aux systèmes qui le suivront le

Un troisième problème de la TBS qu'a soulevé Kingma (2010) est qu'elle permet qu'une fonction d'un trait opérant anormalement (selon nos intuitions médicales) dans un système vivant qui se trouve dans une situation ou un environnement qui lui sont préjudiciables soit jugée comme opérant normalement. La TBS prête le flanc à cette critique parce qu'elle stipule qu'une fonction d'un trait opère toujours *selon une*

 $^{^{143}}$ C'est d'ailleurs ce que soutient Thompson dans ces deux citations :

[[]L]a reproduction consiste à ce qu'une unité [ici un humain] soit à l'origine d'une autre unité de même classe [un garçon ou une fille], c'est-à-dire ayant la *même organisation*. La reproduction requiert donc deux conditions de base : une unité originale [un humain] et un processus qui reproduit cette unité. Dans le cas des êtres vivants [comme les humains], l'unité originale est une unité autopoïétique et le processus reproducteur aboutit sur la formation d'au moins une autre unité autopoïétique *distincte* [sur le plan de sa structure, mais pas de son organisation] de la première. (Thompson, 2007, p. 167)

La reproduction donne nécessairement naissance à des unités *historiquement liées* parce qu'elle se produit par la division et l'incorporation de l'unité originale dans sa progéniture. (Thompson, 2007, p. 169, mes *italiques*)

situation spécifique pour la survie et la reproduction du système vivant. Le contreexemple de Kingma pour mettre à jour la faille dans ce postulat de la TBS est celui-ci : comme il est statistiquement fréquent ou typique (relativement à une classe de référence) pour la survie et la reproduction que l'estomac d'un individu empoisonné au paracétamol fonctionne plus lentement, alors la TBS dira que, pour cette situation spécifique, l'estomac empoisonné fonctionne normalement. Même s'il est vrai que le fait que l'estomac d'un individu fonctionne plus lentement dans cette situation spécifique puisse effectivement contribuer à la survie de ce dernier en lui évitant une hépatite fulminante, il est contre-intuitif de dire que l'estomac de ce dernier fonctionne normalement.

En fonction de la façon dont le système gère cet empoisonnement, la CÉOST permet de résoudre ce problème que rencontre la TBS de deux manières. Premièrement, on peut dire qu'il y aurait un trouble physique dans les cas de figure où le système n'arriverait pas, par le biais des mécanismes compensatoires qu'il déclenche (le ralentissement de la digestion) pour amenuiser les dégâts au foie que pourrait causer le paracétamol, à maintenir en vie la personne. Deuxièmement, on peut dire qu'il y aurait une dysfonction, mais pas de trouble physique, dans les cas où le système déclencherait des mécanismes adaptatifs de compensation (comme le ralentissement de la digestion) pour pallier cet empoisonnement au paracétamol et réussirait grâce à eux à se maintenir en vie. Ce type de cas peut arriver si, par exemple, une personne tente de se suicider chez elle avec du paracétamol et échoue. Cela dit, dans une situation d'empoisonnement comme celle-là, que l'on soit en présence d'un trouble ou seulement d'une dysfonction, si l'on peut intervenir, il faut absolument le faire, car une surdose de paracétamol est extrêmement toxique pour le foie et potentiellement mortelle.

Le quatrième problème de la TBS est qu'elle permet parfois d'évaluer des traits normaux statistiquement rares comme étant des dysfonctions. Par exemple, comme on l'a vu à la section 2.3, elle permet de dire qu'un individu — disons un homme de 25 ans — qui a les cheveux roux présente une pathologie s'il se trouve dans une société hostile aux personnes qui ont les cheveux de cette couleur (Cooper, 2002, p. 272; Stegenga 2018b, p. 30-31). En effet, présenter ce trait rare dans cet environnement social implique que ce dernier a significativement moins de chances de survivre et de se reproduire que les hommes de 25 ans qui n'ont pas les cheveux roux. Que cela soit vrai ou non, être rousse ou roux n'est pas une pathologie. Cette conclusion est donc problématique pour la TBS.

La CÉOT fournit une solution à cet écueil. En effet, selon elle, ce ne sont pas les personnes rousses qui présentent une dysfonction, mais la société où vivent ces dernières. C'est le cas parce que (1) si selon la NTN une société doit s'autoproduire, s'automaintenir et persister tout en assurant l'automaintenance des personnes qui la composent et (2) si cette même société, en raison d'une institution cognitive discriminante qui émerge d'elle-même, méprise les personnes rousses, alors (3) cette société n'assure pas l'automaintenance de ces dernières et, en conséquence, (4) elle est dysfonctionnelle. Cette conclusion demande ainsi que l'on intervienne sur cette institution cognitive qui génère de l'hostilité à l'égard des personnes rousses. Une intervention possible serait de se débarrasser de cette institution malsaine.

Maintenant que l'on a pris connaissance des solutions qu'offre la CÉOST aux problèmes que rencontre la TBS, on peut étudier la manière dont elle répond à ceux de la DP et de la DCD.

Les problèmes de la DP et de la DCD sont au nombre de trois et ne concernent que leur composante factuelle respective qui est identique ¹⁴⁴ (voir sections 2.4.3). Cette composante est la notion de dysfonction qui doit être conçue suivant la conception étiologique. Selon cette dernière, une fonction d'un trait est un effet qui a été sélectionné au cours de l'histoire évolutive parce qu'il a favorisé la survie et la reproduction des ancêtres du système vivant qui le porte. Par exemple, selon cette conception, la fonction du cœur dans les systèmes vivants est de pomper le sang de manière adaptée aux situations diverses (plus rapidement ou plus lentement) parce que le fait que le cœur pompe le sang de cette manière a favorisé, dans le passé, la survie et la reproduction des systèmes vivants qui avaient cet organe. Dans cette perspective, un trait d'un système vivant a une dysfonction quand il ne fait pas ce pour quoi il a été sélectionné dans le passé. Par exemple, le cœur arythmique d'une personne est considéré comme dysfonctionnel parce qu'il ne pompe pas le sang de manière adaptée aux différentes situations où se trouve cette dernière et que cela nuit, en conséquence, à sa survie.

Le premier problème que présente cette conception de la fonction est celui de la détermination de la classe de référence. Parce qu'elle suppose une définition différente de la classe de référence (section 2.4.3), comme l'explique Stegenga (2018a, p. 24), la conception étiologique de la fonction suppose une *version différente* de ce problème que celle que rencontre la TBS (section 2.3). Quelle est cette nouvelle définition de la classe de référence? Selon la conception étiologique de la fonction, la classe de référence renvoie aux caractéristiques d'un moment précis de l'histoire évolutive de l'espèce du système vivant présentant le trait porteur de la fonction.

¹⁴⁴ Les composantes évaluatives de ces conceptions du trouble présentent aussi des problèmes, mais j'ai décidé, dans cette thèse, de ne m'intéresser qu'aux problèmes de leur composante factuelle respective.

Suivant cette nouvelle définition, la nouvelle version du problème de la classe de référence prend la forme de cette question : considérant que les pressions sélectives changent avec le temps, quel moment de l'histoire évolutive des ancêtres de l'organisme est-il le plus approprié pour évaluer le fonctionnement normal d'un trait ou d'un mécanisme ? (section 2.4.3)

Par exemple, dit Stegenga (2018a, p. 24), faut-il évaluer le trait « courir lentement » à la lumière des pressions sélectives opérantes à l'ère des sociétés préagricoles ou à la lumière de celles qui sont opérantes aujourd'hui? Comme on l'a vu à la section 2.4.3, on ne peut pas répondre à cette question sans faire appel à des considérations normatives. La composante factuelle de la DP et de la DCD, en conséquence, n'est pas objective.

La CÉOT ne rencontre pas ce problème parce qu'elle ne fait pas appel à la notion de classe de référence. Elle n'a donc pas à déterminer le moment de l'histoire évolutive de l'espèce à laquelle appartient ce système vivant et qui sert d'étalon pour établir ce qu'est le fonctionnement normal et le fonctionnement anormal d'un trait. En effet, comme on l'a vu, le fonctionnement d'un trait ou d'un mécanisme est évalué à la lumière de la NTN qui est inhérente à tous les systèmes vivants (comme les personnes), en tout lieu et à toutes les époques.

Le second problème que rencontrent la DP et la DCD est qu'elles ne saisissent pas les troubles mentaux qui ont pour base (selon la conception étiologique) une fonction adaptative. Par exemple, on peut émettre l'hypothèse psychoévolutionniste selon laquelle les traits caractéristiques des gens qui sont des psychopathes ont été sélectionnés parce qu'ils ont aidé celles ou ceux qui les avaient à survivre et à se reproduire dans les environnements difficiles du passé lointain où vivaient les chasseuses-cueilleuses et les chasseurs-cueilleurs (Garson, 2015, p. 177). Si cette

hypothèse est exacte, alors il faudrait dire que des traits qui fonctionnent normalement sont sous-jacents à la psychopathie, qui est un trouble mental. Comme la DP et la DCD disent qu'un trouble mental suppose une dysfonction, alors elles ne saisissent pas le trouble mental de la psychopathie.

La CÉOT ne permet pas ce genre de conclusion. En effet, être psychopathe, en vertu de la NTN, est un trouble (ou à tout le moins une dysfonction) parce que présenter ce profil psychologique nuit à la capacité de la personne de s'automaintenir et de persister dans le temps. En effet, à l'époque actuelle, en raison de leurs comportements risqués, l'espérance de vie des psychopathes est généralement moins longue (Nadelhoffer et Sinnott-Armstrong, 2013, p. 241).

La troisième critique que l'on peut faire de la conception étiologique de la fonction de la DP et de la DCD est que les traits qu'elles saisissent proviennent d'hypothèses psychoévolutionnistes très spéculatives. En effet, on en sait très peu sur les fonctions mentales des êtres humains ayant vécu dans un passé lointain (voir section 2.4.3). Il est donc difficile de dire que les fonctions mentales que l'on observe aujourd'hui sont celles qui étaient opérantes à l'époque du pléistocène.

La CÉOST ne rencontre pas ce problème parce que son concept de fonction n'est pas un concept historique. En effet, si on réfléchit en son cadre, il n'est pas nécessaire de savoir quel était le fonctionnement normal des traits et mécanismes dans le passé. On a seulement à vérifier si les traits et les mécanismes servent ou ne servent pas la NTN dans le moment présent.

3.3 Conclusion du troisième chapitre

Dans ce troisième chapitre, j'ai présenté la CÉAC (section 3.1) et, ensuite, la CÉOST (qui suppose la CÉAC) (section 3.2.1). On a vu que la CÉOST n'a pas les problèmes que rencontre la conception du trouble mental des RDoC (sections 3.1.1 et 3.2.2). En effet, contrairement à cette dernière, la CÉOST suppose que la cognition est fortement incrustée dans l'environnement et constituée de composantes extraneurales. En conséquence, la CÉOST favorise la recherche sur les facteurs environnementaux (physiques, sociaux et culturels) causant les troubles mentaux de même que celle portant sur leurs composantes corporelles, multipersonnelles et cognitivoinstitutionnelles. Par ailleurs, à l'inverse de la conception du trouble des RDoC, la CÉOST encourage que soient menées des recherches phénoménologiques en santé mentale. Vu tout cela, on peut dire que la CÉOST permet une diversification de la recherche psychiatrique, laquelle est plutôt limitée par les RDoC. Enfin, la CÉOST saisit la normativité dans le monde du vivant (sections 3.1, 3.2.1 et 3.2.2), une action que semble incapable de faire la TBS que contient la conception du trouble mental des RDoC (section 2.3) et la DP et la DCD (section 2.4.3). Bien que je n'en aie pas traité, on peut imaginer que cette capacité normative que présente la structure conceptuelle de la CÉOST a des implications pratiques. Par exemple, en posant la question « quels mécanismes de ce système dynamique X nuisent à la NTN? », on peut éventuellement identifier les composantes sur lesquelles on doit intervenir pour rétablir ce dernier. Une autre conséquence concrète qu'amène l'adoption de la CÉOST est d'entrainer une réduction du nombre de catégories de troubles mentaux que l'on trouve dans le DSM. En effet, plusieurs des troubles mentaux qui sont en ce moment listés dans cet ouvrage sont plutôt, selon la CÉOST, des dysfonctions mentales. Je laisse la lectrice ou le lecteur imaginer les autres possibilités pratiques que permet cet aspect normatif de la CÉOST et, plus généralement, l'entièreté de sa structure conceptuelle.

CONCLUSION

Dans cette thèse, j'ai soutenu l'idée selon laquelle la conception du trouble mental au cœur des RDoC est problématique pour la recherche et qu'elle doit, pour cette raison, être remplacée par une conception énactive et objective de la santé du trouble mental (CÉOST).

Mon argumentaire, pour faire valoir cette thèse, s'est déployé sur trois chapitres. Le contenu du premier d'entre eux est à la fois descriptif et analytique. En effet, les sections 1.1 à 1.4 contiennent une présentation des RDoC. Quant à elle, la section 1.5 contient principalement deux choses. La première est une définition précise de la conception du trouble mental des RDoC (section 1.5.1), à savoir la conception selon laquelle les troubles mentaux au sein des RDoC sont conçus comme des dysfonctions des grands domaines de fonctionnement de l'esprit/cerveau que l'on retrouve dans la matrice des RDoC. La seconde est mon exposition de ce qui me semble être les quatre bases philosophiques de la conception du trouble mental des RDoC. Ces dernières sont une certaine version à la fois pluraliste et réductionniste du néomécanisme de Bechtel, laquelle ressemble au réductionnisme fragmentaire de Schaffner (section 1.5.2), une interprétation adaptationniste de la théorie biostatistique (TBS) de la santé de Boorse (section 1.5.3), le principe épistémoquantitatif (qui suppose ce que je pense être une variante quantitative du réalisme des entités basé sur la robustesse d'Eronen) (section 1.5.4) et l'hypothèse de l'identité psychoneurale formulée par Bunge (section 1.5.5).

Le deuxième chapitre contient ma critique de la conception du trouble mental des RDoC. Plus précisément, je montre qu'elle présente quatre problèmes. Le premier d'entre eux est que l'hypothèse de l'identité psychoneurale de Bunge qui lui est sousjacente est inadéquate parce que, d'une part, elle tend à isoler la cognition d'une personne de son environnement (physique, social et culturel) et parce que, d'autre part, elle ne saisit pas les composantes extraneurales de la cognition, à savoir les composantes corporelles, multipersonnelles et cognitivo-institutionnelles (sections 2.1, 2.1.1, 2.1.2, 2.1.3 et 2.1.4). Le deuxième problème découle du premier. Ce second problème est que (a) la conception du trouble mental des RDoC empêche (ou du moins, n'encourage pas) que soient menées des recherches sur les mécanismes environnementaux (physiques, sociaux et culturels) favorisant le développement des troubles mentaux et que (b) elle empêche (ou du moins, n'encourage pas) que soient menées des enquêtes sur les composantes extraneurales susmentionnées (sections 2.1, 2.1.1, 2.1.2, 2.1.3 et 2.1.4). Cela est malheureux parce que les résultats de ces recherches enrichiraient notre compréhension des troubles mentaux. Cela est aussi regrettable parce que ces composantes extraneurales pourraient servir de cibles d'interventions thérapeutiques. Le troisième problème est que la conception du trouble mental des RDoC, en raison du principe épistémoquantitatif qui lui est sous-jacent, empêche que soient menées des recherches phénoménologiques et qualitatives sur les troubles mentaux (sections 2.2, 2.2.1 et 2.2.2). Cela est malheureux parce que ce type de recherches est pertinent épistémiquement et pour le développement d'interventions thérapeutiques. Finalement, le quatrième problème est que la théorie biostatistique (TBS) de la santé sous-jacente à la conception du trouble mental des RDoC ne saisit pas la normativité dans les systèmes vivants et, par conséquent, dans les personnes (section 2.3). Cela est problématique parce que, si elle ne peut pas faire cela, alors elle ne peut pas distinguer le normal du pathologique (chose qu'une conception du trouble devrait être capable de faire). On aurait pu penser que la dysfonction préjudiciable (DP) de Wakefield et la dysfonction correctement dévalorisée (DCD) de Powell et Scarffe,

qui ont une composition conceptuelle différente de celle de la TBS, auraient pu régler les problèmes de cette dernière (sections 2.4, 2.4.1 et 2.4.2). Cependant, ce n'est pas le cas (section 2.4.3). Ainsi, la TBS, la DP et la DCD ne sont pas des conceptions du trouble (mental et physique) adéquates.

Le troisième chapitre contient ma solution aux quatre problèmes susmentionnés. En effet, j'y soutiens que la CÉOST, que je présente à la section 3.2.1, résout ces quatre problèmes. Comme elle suppose la conception énactive autopoïétique de la cognition (CÉAT) (section 3.1), elle résout les trois premiers problèmes (section 3.1.1). Elle résout le premier d'entre eux parce que, contrairement à l'hypothèse de l'identité psychoneurale de Bunge, elle suppose que la cognition est incrustée dans l'environnement (physique, social et culturel) et constituée de composantes corporelles, multipersonnelles et cognitivo-institutionnelles. En conséquence, elle résout le deuxième problème. En effet, puisqu'elle saisit l'incrustation de la cognition dans l'environnement et ses constituantes extraneurales, elle permet (et encourage) que soient menées des recherches sur (a) les mécanismes environnementaux (physiques, sociaux et culturels) qui favorisent le développement de troubles mentaux et sur (b) les composantes corporelles, multipersonnelles et cognitivo-institutionnelles de ces derniers. Elle résout le troisième problème parce qu'elle permet et encourage que soient faites de la recherche phénoménologique et de la recherche qualitative. Il en est ainsi parce que la CÉAT suppose la thèse de la continuité vie-esprit qui tient pour fondamental l'expérience subjective et la vie symbolique et affective des humains. Enfin, la CÉOST résout le quatrième problème parce qu'elle contient les ressources conceptuelles qui lui permettent, contrairement à la TBS (sous-jacente à la conception du trouble mental des RDoC), à la DP et à la DCD, de saisir la normativité dans les systèmes vivants et donc dans les personnes (section 3.2.2).

Pour finir, j'aborde brièvement quelques pistes de recherche qui s'offrent à partir d'ici. Premièrement, on pourrait tenter de déterminer quelle ontologie conviendrait le mieux pour inscrire la CÉOST et ses aspects dynamiques et temporels. La *Théorie générale des processus* (TGP) développée par Seibt (2009, 2018) me semble appropriée à cette fin. Selon cette ontologie — et en cela elle entretient, me semble-t-il, des ressemblances avec l'énactivisme autopoïétique — les constituants fondamentaux du monde sont des *entités dynamiques* (Seibt, 2018, p. 114) ou des *processus généraux* qui :

(i) se produisent concrètement ou spatiotemporellement, mais pas nécessairement dans une région délimitée et déterminée, (ii) sont plus ou moins spécifiques ou déterminés, (iii) sont plus ou moins récurrents dans l'espace et/ou dans le temps, et (iv) sont plus ou moins complexes. (Seibt, 2009, p. 495-496)¹⁴⁵

D'un côté, Seibt pense que la TGP convient pour « toutes les différentes sortes d'entités dont on parle dans le sens commun et en science (y compris dans les sciences humaines) » (Seibt, 2018, p. 115). Par exemple, les personnes et les troubles mentaux (ou physiques) que présentent ces dernières pourraient être considérés comme des entités dynamiques, comme je l'ai soutenu dans cette thèse. D'un autre côté, Seibt soutient que s'inscrivent aussi tout naturellement dans cette ontologie les énoncés qui saisissent ces entités dynamiques se trouvant dans le monde comme les énoncés « le feu s'est arrêté, mais pas les radiations » (Seibt, 2018, p. 116), « les métastases envahissent le corps de cet homme » ou « cet individu a plus d'idéations suicidaires qu'hier ». Suivant ce qui précède, je crois que des structures conceptuelles comme la

¹⁴⁵ À noter au passage qu'en adoptant cette posture quant à l'ameublement du monde, Seibt s'oppose ainsi aux ontologies substantialistes selon lesquelles les constituants fondamentaux du monde sont des *entités statiques* (Seibt, 2018, p. 113) qui :

⁽i) ont chacune une localisation spatio-temporelle unique et lesquelles (ii) ont nécessairement cette localisation spatio-temporelle unique parce qu'elles sont individuées en termes de cette localisation. (Seibt, 2009, p. 484)

CÉOST gagneraient à être enchâssées dans la TGP. Ultimement, la science psychiatrique tout entière pourrait également y trouver une demeure.

Deuxièmement, on pourrait explorer le programme de naturalisation de la phénoménologie que contient l'énactivisme autopoïétique. Dans cette thèse, en effet, si j'ai utilisé l'énactivisme autopoïétique pour montrer que les dysfonctions sous-jacentes à un trouble mental peuvent se trouver ailleurs que dans le cerveau de la personne présentant ce trouble, je n'ai pas exploré en profondeur les travaux en sciences cognitives énactives portant sur la naturalisation des différentes expériences subjectives anormales.

En quoi consiste cette naturalisation de la phénoménologie? Elle ne vise pas à une réduction de l'expérience mentale d'une personne aux activités neurales qui ont lieu en elle. Elle vise plutôt à appliquer le principe méthodologique selon lequel il faut « articuler par des contraintes mutuelles le champ des phénomènes révélé par l'expérience au champ corrélatif des phénomènes établi par les sciences cognitives » (Varela, 1996, traduction de Bitbol dans Varela, 2017, p. 330). Ce principe méthodologique suppose donc le principe épistémique selon lequel les données que fournissent les études phénoménologiques et les données que donnent les études en sciences cognitives s'éclairent et s'autodéterminent mutuellement dans une dialectique perpétuelle où jamais elles ne sont réduites aux unes ou aux autres. Carel (2016, p. 35), s'inspirant de Wheeler (2013), semble soutenir quelque chose de semblable lorsqu'elle affirme que « la phénoménologie et les sciences cognitives s'informent et se contraignent l'une l'autre ». À ces considérations méthodologiques et épistémiques, on peut ajouter que, quand sont formalisées les relations qu'entretiennent les données issues de la phénoménologie et des sciences cognitives, les outils mathématiques privilégiés sont ceux de la théorie des systèmes dynamiques. En somme, idéalement, la naturalisation de la phénoménologie vise donc à intégrer trois éléments, à savoir « (1) les analyses phénoménologiques de l'expérience (2) la théorie des systèmes dynamiques et (3) les expérimentations empiriques sur les systèmes biologiques » (Gallagher et Zahavi, 2012, p. 38).

Il serait pertinent de s'intéresser aux résultats des études empiriques en psychopathologie où est naturalisée la phénoménologie des troubles mentaux pour au moins deux raisons. La première est que cela donnerait une compréhension plus fine de la vie mentale des gens souffrant de troubles mentaux. La deuxième raison est que cela permettrait de découvrir de nouvelles techniques d'intervention thérapeutiques. Par exemple, les travaux de Claire Petitmengin (2005, p. 85) laissent penser qu'il est possible d'apprendre aux personnes souffrant d'épilepsie à reconnaitre la phénoménologie très subtile qui précède de quelques minutes une crise et à agir d'une manière ou d'une autre sur cette dernière à l'aide de ce que Petitmengin appelle des « contremesures » pour la faire avorter.

Troisièmement, l'adoption de la CÉOST et de la notion d'institution cognitive qui lui est inhérente pourrait favoriser le développement d'une psychiatrie intrinsèquement autocritique. Une telle psychiatrie s'interroge perpétuellement sur la validité de ses fondements (ontologiques, épistémologiques, normatifs et axiologiques) de même que sur ses pratiques et les savoirs qu'elle génère. Pourquoi l'adoption de la CÉOST et de la notion d'institution cognitive favorise-t-elle une psychiatrie autocritique? Parce que, selon la CÉOST, la psychiatrie est elle-même une institution cognitive et que l'on doit, en conséquence, s'y intéresser et l'interroger si l'on veut être en mesure de comprendre les agents cognitifs dont l'esprit est constitué par elle, c'est-à-dire les patientes et les patients, les dirigeantes et dirigeants d'instituts en santé mentale, les chercheuses et chercheurs en santé mentale, les cliniciennes et cliniciens, et même les gens de la société en général. Plus précisément, il faut s'intéresser à l'impact qu'ont sur ces agents cognitifs les mécanismes, les pratiques institutionnelles, les normes, les règles et les

technologies qu'on trouve dans l'institution psychiatrique (selon Gallagher, 2013, p. 11-12 c'est à ce type d'éléments que l'on doit s'intéresser quand l'on cherche à comprendre une institution cognitive). Comme on peut le constater, « la notion d'institution cognitive est en soi un outil utile pour développer une position critique qui nous permet d'examiner les pratiques institutionnelles du moment » (Slaby et Gallagher, 2015, p. 38).

L'étude de l'institution psychiatrique renseignera, par exemple, à propos des critères sur lesquels se basent les évaluations des cliniciennes et des cliniciens qui cherchent à départager la santé mentale de la pathologie mentale. Ces critères sont-ils adéquats ? Il est important de se poser cette question parce que s'ils ne le sont pas des gens pourraient être jugés malades alors qu'ils ne le sont pas (et inversement). Ce genre d'analyse pourrait également renseigner sur les critères qui guident la prise de décision (sur ce qui motive le choix d'un mode d'intervention thérapeutique ou le choix d'un cadre de recherche comme celui des RDoC, par exemple). Il est donc particulièrement important d'évaluer si ces critères sont bons parce qu'ils ont des incidences réelles (bonnes et/ou mauvaises) sur la santé mentale des gens. Par ailleurs, de telles études pourraient également renseigner sur la façon dont une société se représente la maladie mentale en général (par une analyse des médias, par exemple). En outre, elles pourraient informer sur la manière dont les patientes et les patients conçoivent leur condition mentale particulière et sur la manière dont les membres du personnel médical conçoivent ces patientes et patients qui présentent cette condition. Enfin, elles pourraient éventuellement montrer que, dans certains cas, la dysfonction mentale de la patiente ou du patient se trouve dans l'institution psychiatrique elle-même. Par exemple, on peut penser que, lorsque l'homosexualité était conçue comme un trouble de santé mentale (dans les DSM-I, II et III), la dysfonction se trouvait dans l'institution psychiatrique et non dans la personne.

ANNEXE A

ORIGINE DU CONCEPT D'INSTITUTION COGNITIVE (ET DE LA THÉORIE DE L'ESPRIT SOCIALEMENT ÉTENDU)

Dans cette annexe, on apprendra que, pour avoir une compréhension plus complète de ce qu'est le construit théorique « institution cognitive », il faut saisir la conception de l'esprit socialement étendu qui est son assisse conceptuelle (Petracca et Gallagher, 2020, p. 4). On montrera également que la conception de l'esprit socialement étendu est mieux appréhendée si l'on comprend qu'elle provient d'une critique que fait Gallagher à l'endroit de la conception de l'esprit étendu formulée par Clark et Chalmers (1998) (Petracca et Gallagher, 2020, p. 4). Enfin, on verra qu'il vaut mieux, pour concevoir l'extension de l'esprit, adopter la théorie de l'esprit socialement étendu plutôt que la théorie de l'esprit étendu pour deux raisons. La première est qu'elle suppose la conception énactive de la cognition (vue à la section 3.1) qui est plus adéquate, selon Gallagher (2013, p. 11) que la conception computationnelle et fonctionnaliste de la cognition que suppose la théorie de l'esprit étendu. La deuxième raison est qu'elle n'est pas affectée, selon Gallagher (2013, p. 8-11), par les trois critiques principales qui sont faites à l'endroit de la théorie de l'esprit étendu, à savoir celles que l'on nomme respectivement le « coupling-constitution fallacy » (Adams et Aizawa, 2001, 2008, chapitre 6), «l'enflure cognitive» [cognitive bloat] (Rupert, 2004, p. 401-405) et le « mark of the mental » (Adams et Aizawa, 2008, p. 10; voir aussi Weiskopf et Adams, 2015, p. 149-53).

Dans ce qui suit, d'abord, je présente l'appareillage conceptuel de la théorie de l'esprit étendu. Après, j'explique pourquoi Gallagher rejette cette dernière et lui préfère la théorie de l'esprit socialement étendu. Je termine cette annexe en montrant pourquoi Gallagher pense que la théorie de l'esprit socialement étendu ne rencontre pas les trois problèmes mentionnés précédemment auxquels se bute la théorie de l'esprit étendu.

L'esprit étendu de Clark et Chalmers

La conception de l'esprit étendu a été développée par Clark et Chalmers (1998). Elle suppose la thèse ontologique selon laquelle l'esprit est constitué de *représentations*, de *propositions* et, ultimement, *d'attitudes propositionnelles* (voir Gallagher, 2013, p. 5; Menary, 2013, p. 31; Varga, 2013, p. 54-55 pour un avis semblable). Ces trois notions sont présentées dans ce qui suit.

On définit les *représentations* comme des structures porteuses d'information (Pitt, 2020). Ces représentations sont toujours implémentées dans une structure physique quelconque (le cerveau est souvent cette structure, mais cette dernière pourrait aussi être autre chose comme l'appareil responsable de la cognition dans le crâne d'un robot humanoïde). Elles ont également la propriété d'être « à propos » de quelque chose le « quelque chose » peut être abstrait (ex. : une licorne, le gouvernement, le déterminant « le », le verbe « manger ») ou concret (le soleil à l'extérieur, l'écran ou le clavier devant moi).

¹⁴⁶ J'aimerais remarquer au passage que, parce qu'elles sont « à propos » de quelque chose, on lie souvent la notion de représentation à la notion phénoménologique d'intentionnalité (Rescorla, 2020, section 5). En effet, cette dernière réfère elle aussi à cette propriété fondamentale de l'esprit «d'être à propos de, de représenter ou de tenir place à des choses, des propriétés ou des états de choses» (Jacob, 2019). On a vu cette notion à la section 2.2.1.

Quand ces représentations sont combinées avec d'autres représentations, selon (par exemple) les règles de ce que Fodor (1975) a nommé le langage de la pensée, elles forment une *proposition*, laquelle réfère à la signification d'une phrase déclarative qui peut être vraie ou fausse. Pour rendre cela plus concret, on peut prendre l'exemple d'un agent cognitif qui possède, dans son langage, les représentations « LES », « CHATS », « CHIENS », « JAPPENT » et « MIAULENT » (je suis ici Mandik, 2014, p. 103). Selon ce que permettent et ne permettent pas les règles du langage de la pensée, cet agent peut avoir dans son esprit certaines propositions et ne peut pas en avoir d'autres. Par exemple, il pourrait avoir la proposition vraie « LES CHIENS JAPPENT », mais il ne pourrait pas avoir la proposition « CHATS CHIENS LES », qui ne respecte pas les règles du langage de la pensée et qui, en conséquence, n'est ni vraie ni fausse. Cet agent cognitif pourrait avoir, cependant, la proposition « LES CHATS JAPPENT ». Il pourrait avoir cette proposition parce que, bien que cette dernière soit fausse en regard de ce que contient le monde, elle respecte néanmoins le type de règles susmentionnées.

Ultimement, ces propositions deviennent, chez un agent cognitif, des *attitudes propositionnelles*. Celles-ci réfèrent à la posture que prend un agent cognitif relativement à une proposition. On peut utiliser la proposition « il fait soleil » pour illustrer cela. Plusieurs attitudes propositionnelles sont possibles en regard de cette proposition. On peut *croire* qu'il fait soleil. On aura alors l'attitude propositionnelle « je crois qu'il fait soleil ». On peut *percevoir* qu'il fait soleil. On aura alors l'attitude propositionnelle « je perçois qu'il fait soleil ». On peut également *désirer* qu'il fasse soleil. On aura alors l'attitude propositionnelle « je désire qu'il fasse soleil ». Ces attitudes propositionnelles ont des propriétés sémantiques : les croyances sont vraies ou fausses, les perceptions sont exactes ou inexactes et les désirs sont assouvis ou inassouvis (Rescorla, 2020, section 5).

La thèse selon laquelle l'esprit est constitué de représentations, de propositions et d'attitudes propositionnelles est sous-jacente à la conception selon laquelle l'esprit fonctionne de manière computationnelle. Selon ce modèle de l'esprit, la cognition est comprise comme une computation, c'est-à-dire comme un traitement des représentations selon un ensemble de règles fixes (par exemple, les règles du langage de la pensée de Fodor). Cette conception computationnelle de l'esprit suppose que l'esprit présente une dynamique d'opérations linéaires qui suit la séquence suivante : perception → cognition → action (Hurley, 2001, p. 7). Plus précisément, selon ce modèle de l'esprit, (1) les intrants sensoriels provenant de l'extérieur ou de l'intérieur de l'agent sont traités par les modules perceptuels (2) la cognition tire de ces intrants des représentations (voir la définition présentée précédemment), traite ces dernières selon certaines règles et produit une attitude propositionnelle et (3) une action (motrice, par exemple) est opérée par l'agent cognitif en fonction de cette attitude propositionnelle. Enfin, il est à noter que ce modèle suppose que les processus cognitifs (perceptuels, par exemple) d'un agent ne sont pas influencés par ses actions. Le schéma suivant représente ce modèle de l'esprit où, comme le disait Susan Hurley, la cognition est prise en sandwich entre la perception et l'action :

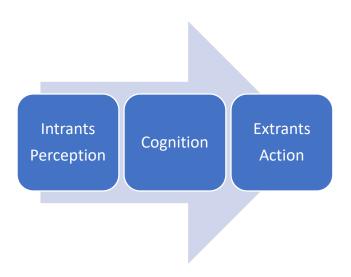


Figure A.1 Modèle en sandwich de l'esprit (inspiré de Hurley, 2001, p. 9)

La conception computationnelle de l'esprit étendu de Clark et Chalmers doit aussi être comprise dans un *cadre fonctionnaliste*. Essentiellement, selon ce cadre,

un état mental [lequel est une attitude propositionnelle ou une représentation] peut être décrit *totalement* en décrivant trois choses : (a) ses causes habituelles, (b) les effets qu'il a sur d'autres états mentaux et (c) quels comportements il peut produire à lui seul ou combiné avec d'autres états mentaux. (Stich et Donaldson, 2018, p. 357, mes *italiques*)

Pour illustrer cela, on peut utiliser la croyance (laquelle est un état mental) « je crois qu'il fait soleil » qui survient sur le cerveau d'un individu. Celle-ci peut être *causée* par le fait que l'individu perçoit qu'il n'y a pas de nuages et qu'il fait jour. Un des *effets* que cette croyance peut avoir est de générer chez lui la croyance « je crois qu'il continuera à faire soleil jusqu'à tard ce soir ». Enfin, les croyances « je crois qu'il fait soleil » et « je crois qu'il continuera à faire soleil jusqu'à tard ce soir » peuvent *causer plusieurs comportements* chez l'individu, par exemple celui de prendre son vélo pour faire une promenade, et ce, dans le but d'aller se baigner dans un lac en soirée.

Il est à noter que Stich et Donaldson (2018) affirment que l'on « décrit totalement » un état mental lorsqu'on décrit (a), (b) et (c) (mentionnés précédemment dans la citation) parce que ce faisant, selon l'approche fonctionnaliste, on donne l'ontologie complète le cet état mental, laquelle est identique *au rôle* ou à *la fonction* que l'état mental joue dans le système (formé des éléments de [a], [b] et [c]) où il se trouve (Stich et Donaldson, 2018, p. 199).

Cette ontologie des états mentaux est très différente de l'ontologie d'un diamant qui est plutôt décrite en termes de caractéristiques physiques comme la dureté et les propriétés optiques (Polger, 2020, section 2). Comme on peut le constater, dans le cadre fonctionnaliste, l'ontologie d'un état mental n'a rien à voir avec ses propriétés physiques ou avec ce dont il est constitué, mais avec ce qu'il *fait* dans le système (Polger, 2020). Cela signifie qu'un état mental ou une représentation peut, en principe, survenir sur des systèmes constitués de n'importe quelle substance comme de la substance organique dont est fait le cerveau des organismes vivants ou comme de la substance non organique dont serait constitué un hypothétique cyborg.

Maintenant que ces notions ont été définies, on peut revenir à la conception de l'esprit étendu qui est l'objet de cette section. En plus de supposer un cadre computationnel et fonctionnaliste, celle-ci a également pour assise ce que Clark et Chalmers ont appelé le *principe de parité* qu'ils formulent de cette manière :

Si, quand nous sommes confrontés à une tâche, une partie du monde fonctionne comme un processus qui, *s'il avait lieu dans notre tête*, serait reconnu par nous sans hésitation comme faisant partie du processus

-

¹⁴⁷ Pour l'état mental « je crois qu'il fait soleil », je pense avoir donné une ontologie *partielle* parce que je ne pense pas avoir épuisé tout ce que cause cette croyance, les états mentaux qu'elle cause elle-même et les comportements qu'elle peut potentiellement engendrer.

cognitif, alors cette partie du monde *fait partie* [...] du processus cognitif. (Clark et Chalmers, 1998, p. 8)

Pour illustrer ce principe de parité, on peut prendre l'image de la figure A.2 suivante :

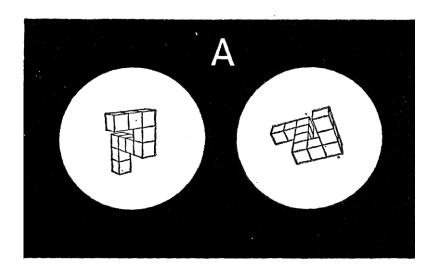


Figure A.2 Exemple de rotation mentale faite sur une forme en trois dimensions (Shepard et Metzler, 1971, p. 702)

On peut dire qu'on a ici affaire à une seule figure géométrique positionnée de deux manières différentes dans l'espace parce qu'on est capable, grâce à nos fonctions cognitives, de faire une rotation mentale. Mais on aurait également pu faire cette rotation à l'aide d'un clavier servant à faire tourner cette figure sur un écran d'ordinateur. Dans ce cas, selon les tenants de la théorie de l'esprit étendu, le clavier et l'écran d'ordinateur sont des ressources du monde *constitutives* de cette tâche cognitive (et sous-jacentes à celle-ci) parce qu'en vertu du principe de parité, ces dernières participent à « un processus qui, s'il avait lieu dans notre tête, serait reconnu par nous

sans hésitation comme une partie du processus cognitif [ici, celle d'effectuer une rotation mentale] » (Clark et Chalmers, 1998, p. 8)¹⁴⁸.

On peut aussi illustrer le principe de parité avec une expérience de pensée développée par Clark et Chalmers (laquelle s'inscrit clairement dans la conception computationnelle/fonctionnaliste de l'esprit vue précédemment). Cette expérience de pensée met en scène les personnages Inga et Otto (Clark et Chalmers, 1998). Celle-ci et celui-ci habitent New York et veulent se rendre au musée d'art contemporain qui se trouve sur la 53° rue. Considérant cet objectif (que l'on pourrait traduire par l'attitude propositionnelle « je souhaite me rendre au musée d'art contemporain »), Inga et Otto ont ainsi à effectuer la tâche cognitive de récupérer dans leur esprit la croyance vraie selon laquelle le musée d'art contemporain se trouve sur la 53° rue (j'écris croyance « vraie » parce que je fais l'hypothèse que, dans cette expérience de pensée, cette croyance est conforme au fait). Inga qui a une santé normale opère cette tâche cognitive grâce à sa mémoire biologique puis se rend au musée. Otto qui présente, pour sa part, la maladie d'Alzheimer opère cette tâche cognitive avec son cerveau et son cahier de notes puis se dirige lui aussi vers le musée.

Il est clair que le cerveau d'Inga est sous-jacent à cette tâche cognitive qui a fait poindre une croyance chez elle, laquelle a ensuite causé qu'Inga se rende au musée. Il est également parfaitement plausible, pour Clark et Chalmers, que le cerveau d'Otto *et* son cahier de notes prennent part à cette tâche cognitive qui consiste à récupérer la croyance « je crois que le musée d'art moderne se trouve sur la 53^e rue », laquelle croyance engendre le comportement d'Otto d'aller au musée. Clark et Chalmers soutiennent cela en s'appuyant sur le principe de parité. En effet, le cahier de notes d'Otto est une

¹⁴⁸ À noter également que le cadre fonctionnaliste permet que la base physique sur laquelle advient la rotation mentale (laquelle est un processus mental) soit le cerveau, le clavier, l'ordinateur et l'écran.

ressource dans l'environnement qui participe à « un processus qui, s'il avait lieu dans sa tête, serait reconnu par nous sans hésitation comme une partie du processus cognitif » (Clark et Chalmers, 1998, p. 8).

La théorie de l'esprit étendu prête le flanc à au moins trois critiques. La première met de l'avant l'argument du « coupling-constitution fallacy » (Adams et Aizawa, 2001, 2008, chapitre 6). Celui-ci dit que l'on ne peut pas conclure que X et Y sont des *constituants* du substrat physique sur lequel survient l'esprit (ou, plus particulièrement, une tâche cognitive quelconque) du seul fait que X et Y soient causalement couplés. Par exemple, on ne peut pas dire que le substrat physique de l'esprit d'Otto est constitué de son cahier de notes *et* de son cerveau (et des informations comme l'emplacement du musée qui surviennent sur eux) du seul fait que le cerveau d'Otto est couplé causalement avec ce cahier. En d'autres termes, cet argument montre que l'on ne doit pas confondre les entrées causales qui influent l'esprit avec les parties qui constituent ce dernier (Clark, 2013, p. 203).

La deuxième critique met de l'avant l'argument de l'enflure cognitive [cognitive bloat] (Rupert, 2004, p. 401-405). Ce dernier suppose que, si l'on permet à l'esprit d'être constitué par des états mentaux qui surviennent sur des choses ou des processus qui ont lieu à l'extérieur du cerveau d'un individu, on peut alors penser que cette extension n'aura pas de limite et que l'esprit pourrait être conçu comme étant constitué d'informations issues de ressources qui ne devraient pas, à première vue, le constituer (Weiskopf et Adams, 2015, p. 145). Par exemple, il serait absurde de penser que toutes les personnes qui ont un téléphone cellulaire avec lequel elles peuvent appeler, quand elles le désirent, un opérateur qui a accès à une liste immense de numéros de téléphone ont des croyances à propos de tous ces numéros de téléphone (Rupert, 2004, p. 402-403).

La troisième critique met de l'avant l'argument du « mark of the mental » formulé par Adams et Aizawa (2008, p. 10; voir aussi Weiskopf et Adams, 2015, p. 149-53). Ceuxci affirment que, pour être en mesure de déterminer si un état ou un processus est mental, on a besoin d'une définition de ce qui est mental. C'est pour cela qu'ils proposent la définition selon laquelle est mental tout état ou processus qui a un contenu non dérivé. Deux exemples concrets aideront à comprendre cette définition. Le premier est une carte géographique représentant le Québec. Le contenu de cette carte — en d'autres mots ce que représente cette carte — est dérivé parce qu'il provient ultimement des intentions et des buts des humains qui l'ont inventée. Le deuxième exemple est l'écran d'ordinateur que je perçois devant moi. Cette perception que j'expérimente subjectivement a un contenu non dérivé parce qu'il ne provient pas de buts ou d'intentions de qui que ce soit. Pour cette raison, selon Adams et Aizawa, cet état perceptif est un état mental. Cela précisé, il convient d'examiner à nouveau la croyance « je crois que le musée d'art contemporain se trouve sur la 53^e rue » qui se trouve dans la mémoire biologique d'Inga (et qui survient sur son cerveau). Est-il possible de la qualifier d'état mental? Oui, selon cette définition, parce que son contenu est non dérivé. Par ailleurs, cette même croyance qui se trouve plutôt sous forme écrite dans le cahier de notes d'Otto est-elle un état mental? On peut dire qu'elle n'en est pas un parce que le contenu de la phrase est dérivé. Adams et Aizawa pensent donc que la théorie de l'esprit étendu ne tient pas la route et que la cognition est nécessairement intracrânienne (Adams et Aizawa, 2008, p. 10).

Clark a tenté de répondre à ces trois critiques en affirmant qu'on est en droit de soutenir qu'X (lequel est une ressource quelconque comme un cahier de notes, un téléphone intelligent, un clavier, un écran, etc.) est un substrat physique d'un processus cognitif ou d'un état mental d'un individu Y si :

(1) X est [...] disponible et utilisé par Y chaque fois que c'est nécessaire

- (2) l'information issue de X est approuvée par Y (« elle ne devrait pas, en règle générale, faire l'objet d'un examen critique », affirme Clark)
- (3) X est facilement accessible à Y quand Y en a besoin
- (4) l'information issue de X a été approuvée à un moment donné dans le passé par Y et c'est la raison pour laquelle elle est approuvée maintenant par Y. (adaptation de Clark, 2008, p. 79)

Par exemple, en vertu de ces quatre critères, on peut dire que le cahier de notes d'Otto fait partie de son esprit parce que (1) il est disponible chaque fois qu'Otto en a besoin (2) l'information qui en est issue est approuvée par Otto (3) il est facilement accessible (en effet, Otto l'a toujours avec lui) et (4) les informations qui s'y trouvent ont été approuvées par Otto dans le passé.

En théorie de l'esprit étendu suppose conception somme, la la computationnelle/fonctionnaliste de l'esprit. On retrouve également à sa base le principe de parité. S'ajoutent à ce dernier quatre critères formulés par Clark qui permettent de déterminer lesquelles des ressources de l'environnement constituent ou sont un substrat physique sur lequel (entre autres) surviennent les processus cognitifs d'un individu. Ces critères servent aussi à protéger la théorie de l'esprit étendu des critiques qui lui sont formulées.

L'esprit socialement étendu

Gallagher défend, comme Clark et Chalmers, l'hypothèse selon laquelle la cognition peut s'étendre dans l'environnement. Cependant, il est en désaccord avec eux sur deux éléments théoriques fondamentaux de leur conception de l'esprit étendu.

Premièrement, il conteste le *principe de parité* et les *quatre critères* susmentionnés dont se servent ces derniers pour déterminer quelles sont les ressources dans

l'environnement qui sont des substrats physiques sur lesquels surviennent certains processus cognitifs.

Gallagher (2013, p. 5) critique le principe de parité en soulevant le fait qu'il suppose paradoxalement la conception intracrânienne de la cognition que combattent les défenseurs de l'esprit étendu. Le principe de parité stipule, en effet, que ce sont seulement les processus qui ont lieu *dans la tête* ou qui pourraient avoir lieu dans la tête qui peuvent être qualifiés de cognitifs. En d'autres termes, l'étalon que contient ce principe pour déterminer si un processus est cognitif est intracrânien. Cela ne devrait pas être le cas selon Gallagher. Celui-ci croit, en effet, que l'on ne devrait pas exclure systématiquement de la cognition les processus ¹⁴⁹ qui n'ont pas lieu ou qui ne pourraient pas avoir lieu dans la tête.

Gallagher affirme, par ailleurs, que les quatre critères de Clark et Chalmers qui servent à déterminer si une composante fait partie ou ne fait pas partie de l'esprit sont inadéquats pour plusieurs raisons (Gallagher, 2013, p. 5). D'abord, il soutient qu'on ne devrait pas exclure de l'activité cognitive d'un agent une ressource rarement disponible et difficilement accessible. Par exemple, le système légal, pour la plupart des gens, n'est ni facilement disponible (comme l'exige le critère 1) ni aisément accessible (comme le commande le critère 3), mais il peut néanmoins faire partie des ressources cognitives d'un agent qui doit résoudre un problème légal (Gallagher, 2013, p. 5). Gallagher affirme aussi que le critère 2, selon lequel l'information contenue dans une ressource doit être acceptée plus ou moins automatiquement par le sujet, est aussi problématique. Pourquoi, en effet, si on fait preuve d'esprit critique, devrait-on rejeter l'information issue d'une ressource quelconque X? Par exemple, il serait inadéquat de

¹⁴⁹ Gallagher ne parle pas de choses, de parties ou de ressources parce que la conception énactive de l'esprit qu'il adopte pose que la cognition est, en son essence, un processus ou une activité.

dire que si Otto doute de l'information qui se trouve dans son carnet, alors son carnet ne participe plus à sa cognition (Gallagher, 2013, p. 6). Gallagher ne discute pas le quatrième critère. Néanmoins, on peut penser qu'il ne l'accepterait pas non plus puisqu'il est très semblable au critère 2.

En somme, selon Gallagher, qu'une ressource X ne respecte pas le principe de parité et les quatre critères formulés par Clark et Chalmers n'est pas une raison d'exclure que ce X puisse être un substrat sur lequel survient un processus de l'esprit d'un individu Y.

Le deuxième élément théorique que Gallagher rejette est la conception computationnelle/fonctionnaliste de l'esprit qu'adoptent Clark et Chalmers. Il lui préfère la conception énactive autopoïétique de la cognition (voir section 3.1). Dans ce cadre, la cognition n'est pas conçue comme une computation sur des représentations survenant sur des structures matérielles, mais en termes d'activités de l'agent (lesquelles impliquent toutes sortes de processus à l'intérieur et à l'extérieur de cet agent). Plus précisément, la cognition y est appréhendée comme étant constituée et émergeant d'interactions qu'entretient continuellement un agent incarné avec l'environnement avec qui il forme un système dynamique (Petracca et Gallagher, 2020, p. 19).

Adopter cette manière énactive de concevoir l'esprit a pour conséquence que la notion d'extension de l'esprit n'est plus comprise de la même manière. Dans ce cadre, ce ne sont pas (comme le soutiennent Clark et Chalmers) les ressources (ou les véhicules) sur lesquelles surviennent des représentations et des croyances qui sont conçues comme des extensions du cerveau sur lequel survient la cognition. En effet, Gallagher et Petracca soutiennent que, dans le cadre énactif, quand l'on conceptualise la notion d'extension, l'accent doit plutôt être mis sur les *processus* cognitifs qui sont rendus

possibles par ces véhicules ou ressources (Petracca et Gallagher, 2020, p. 8, note 5). Autrement dit, alors que la conception computationnelle/fonctionnaliste qu'adoptent Clark et Chalmers fait que ceux-ci insistent sur les ressources, véhicules ou structures extérieures à l'individu sur lesquels surviennent des croyances et des processus cognitifs, la conception énactive conduit Gallagher à plutôt insister sur les ressources, véhicules et structures qui permettent la production de certains nouveaux processus cognitifs et sans lesquels ils ne pourraient tout simplement pas avoir lieu. Gallagher et Petracca pensent donc que Clark et Chalmers accordent trop d'importance aux véhicules de la cognition (Petracca et Gallagher, 2020, p. 8, note 5), et ce, au détriment des processus cognitifs eux-mêmes. Toujours à ce propos, ils ajoutent que Clark et Chalmers ne devraient pas trop s'attarder à ces véhicules parce qu'il est faux qu'un crayon et du papier ou un cahier de notes soient des extensions de notre cerveau (Petracca et Gallagher, 2020, p. 8, note 5). En effet, ces parties du monde ne sont que des ressources qui permettent l'existence de certains processus cognitifs et sans lesquelles ces derniers ne pourraient pas avoir lieu.

C'est donc dans le cadre de l'énactivisme autopoïétique que Gallagher conçoit sa théorie de l'esprit socialement étendu. Gallagher dira d'un processus cognitif qu'il est socialement étendu quand la ressource qui le permet et sans laquelle il ne pourrait exister est une institution cognitive. En effet, ces dernières permettent aux individus de s'engager dans des activités cognitives qui ne peuvent pas être produites par une seule tête et même par plusieurs têtes (Gallagher, 2013, p. 7). De manière générale, on peut concevoir une institution cognitive comme une « affordance sociale » (Gallagher, 2013, p. 4), c'est-à-dire comme une entité sociale qui offre des possibilités d'actions ou d'activités cognitives. De manière plus précise, on peut définir l'institution cognitive comme un ensemble de pratiques cognitives produites à des moments et des lieux spécifiques (Gallagher, 2013, p. 6) « [...] qui fonctionne comme une entité de

résolution de problèmes lorsqu'il [l'ensemble de pratiques] est correctement *couplé* avec des agents cognitifs » (Petracca et Gallagher, 2020, p. 19, mes *italiques*).

Gallagher comprend de deux manières (à mon avis, complémentaires) cette notion de *couplage*. L'une est simple et l'autre plus technique. Selon la notion plus simple, le couplage en question sera jugé adéquat¹⁵⁰ s'il y a modification neurale dans le cerveau du sujet [(c.-à-d., l'agent cognitif)] qui s'engage avec les institutions (Gallagher, 2013, p. 10). Selon la notion plus technique (qui s'inscrit explicitement dans de la conception énactive autopoïétique de la cognition vue en détail à la section 3.1), Gallagher et Slaby affirment que

les aspects clés du couplage énactif sont : (1) qu'il s'agit d'un processus dynamique (c'est-à-dire d'un processus par l'intermédiaire duquel une codépendance est établie entre les systèmes couplés [(ici l'agent et l'institution cognitive)], de sorte que ce qui se passe dans ou sur un système dépend en partie de la situation de l'autre [système]); (2) que l'engagement récurrent [de l'agent cognitif] avec l'externalité [(ici l'institution cognitive)] conduit à une congruence structurelle (Thompson, 2007, p. 45); (3) que l'organisme (ou l'agent) engagé [...] [dans ce processus] conserve son autonomie (sa propre auto-organisation interne). (Slaby et Gallagher, 2015, p. 35-36)¹⁵¹

C'est lorsqu'un tel couplage existe entre l'agent cognitif et l'institution cognitive (ou, autrement dit, c'est lorsque celui-ci et celle-ci forment un système dynamique) que l'institution cognitive peut être considérée comme une des *composantes* des processus

¹⁵⁰ À noter que cette façon de concevoir un couplage approprié est différente de celle qu'ont développée Clark et Chalmers (1998) qui commande que soient respectés le principe de parité et les quatre critères vus précédemment.

¹⁵¹ Å noter que, comme le disent Slaby et Gallagher (2015, p. 35), cette notion de couplage énactif ne s'applique pas seulement au couplage «agent cognitif-institution cognitive». Elle s'applique aussi, par exemple, aux couplages «agent cognitif-artéfact»,« agent cognitif-outil» ou «agent cognitif-technologie». La nature du couplage variera alors un peu dépendamment de ce avec quoi l'agent cognitif est couplé (mais les trois critères susmentionnés seront toujours respectés).

cognitifs de l'agent. C'est aussi à ce moment qu'il y a extension de certains processus cognitifs de l'agent dans (ou vers) les institutions cognitives. L'extension de la cognition, dans cette perspective, se fait de *l'intérieur* de l'individu vers l'extérieur de celui-ci (dans l'institution cognitive). Mais on peut aussi concevoir l'extension de la cognition selon un mouvement inverse. En effet, on peut affirmer que, depuis *l'extérieur*, les processus cognitifs et les pratiques cognitives inhérentes aux institutions cognitives entrent à *l'intérieur* du cerveau des agents cognitifs. C'est ce que soutient Gallagher:

[L]a cognition ne s'étend pas simplement du cerveau vers l'extérieur pour incorporer des outils, des technologies et des institutions, [...] elle fonctionne [aussi] parfois de l'extérieur vers l'intérieur; les outils, les technologies et les institutions façonnent souvent nos processus cognitifs, font fonctionner notre cerveau de certaines manières et peuvent même provoquer des changements plastiques dans la structure neuronale. (Gallagher, 2013, p. 7)

Les institutions cognitives sont très nombreuses dans ce que l'on pourrait appeler le monde des significations humaines. Parmi elles, on retrouve, entre autres, l'institution de la science (Slaby et Gallagher, 2015), l'institution des genres féminin et masculin (Merritt, 2013), l'institution économique (Petracca et Gallagher, 2020) et, à mon avis, l'institution psychiatrique¹⁵².

Pour illustrer concrètement ce qu'est une institution cognitive, Gallagher a souvent utilisé le cas du système légal (Gallagher, 2011, 2013; Gallagher et Crisafi, 2009; Slaby et Gallagher, 2015, p. 36-37). Comme toutes les institutions cognitives, ce dernier ne

¹⁵² Ici, l'expression « institution (cognitive) psychiatrique » réfère aux normes, règles, concepts, pratiques institutionnelles et savoirs que l'on trouve en psychiatrie. Elle ne réfère pas spécifiquement à l'institutionnalisation en psychiatrie, c'est-à-dire au fait de placer des gens sans leur consentement dans des instituts psychiatriques.

doit pas être conçu comme une entité abstraite, mais plutôt comme un ensemble de *pratiques* cognitives qui le constitue (Slaby et Gallagher, 2015, p. 36). Ces pratiques cognitives sont à l'origine de ce que l'on pourrait appeler des *produits cognitifs* parmi lesquels on retrouve, par exemple, le contrat (mais aussi le jugement de culpabilité, la définition des statuts des parties, la révision des lois, etc.). Plus précisément, un contrat (comme un bail) « est l'expression (dans ce cas un accord juridique) de plusieurs esprits [qui établissent], dans une mémoire extérieure, une décision convenue [...] » (Slaby et Gallagher, 2015, p. 36). Quand on fait un contrat, celui-ci s'ajoute au système légal « qui transcende les particularités de l'esprit de tout individu » (Gallagher, 2013, p. 6). Les contrats, comme tous les produits cognitifs, « contribuent à [nos processus cognitifs] et façonnent nos processus cognitifs à l'occasion d'une réflexion plus poussée ou lors de la résolution d'un problème » (Slaby et Gallagher, 2015, p. 36). Il en est de même pour les jugements rendus au tribunal. Ceux-ci, en effet, doivent respecter un « ensemble de règles qui sont établies par le système [légal] » (Slaby et Gallagher, 2015, p. 36). Aussi, Gallagher ajoute à ce sujet que

[l]e processus par l'intermédiaire duquel les jugements sont rendus dépendra de l'engagement cognitif d'un certain nombre de personnes à l'égard d'un ensemble de lois, dont les composantes pertinentes sont mises en évidence en fonction des particularités précises de l'affaire, au fur et à mesure que la procédure se déroule. (Gallagher, 2013, p. 6)

Bref, les juges qui formulent des jugements juridiques ne formulent pas ces derniers seulement à l'aide de leur esprit. En effet, elles sont capables de formuler ces jugements parce qu'elles sont activement couplées (au sens énactif mentionné précédemment) avec le système légal sans lequel, je le répète, elles ne pourraient pas formuler de tels jugements. C'est ce que Gallagher dit (en citant Clark et Chalmers) : « [r]etirer la partie externe de ce processus cognitif — enlever l'institution légale — et "la compétence comportementale du système va s'effondrer, tout comme cela adviendrait si on lui

enlevait une partie de son cerveau". » (Clark et Chalmers, 1998, p. 9, cité dans Gallagher, 2013, p. 7)

Il est important d'ajouter que le système légal se constitue lui-même grâce à l'activité cognitive des agents qui sont couplés avec lui et en fonction des problèmes qu'il aide à résoudre. En effet, le système légal (comme toutes les institutions cognitives) n'est pas un système statique, mais un système dynamique. De fait, il change avec le temps. Par exemple, la structure légale sous-jacente au système juridique reposant essentiellement sur le serment et l'ordalie opérante jusqu'au milieu du Moyen-Âge est bien différente de la structure légale sous-jacente au système accusatoire anglais que l'on voit apparaître après et où y figure un jury (dont la composition est toutefois bien différente de celle d'aujourd'hui) (Poirier et al., 2019, section 3.1). De manière plus générale, on pourrait dire que le couplage entre l'agent cognitif et le système légal (ou toute autre institution cognitive) crée de la signification (Slaby et Gallagher, 2015, p. 37) et, en cela, contribue à la construction du monde de sens des êtres humains (c'est ce que dit d'ailleurs la CÉAC [section 3.1] que suppose la notion d'institution cognitive).

Avant de terminer cette annexe, il est intéressant de noter qu'il vaut mieux concevoir l'extension de l'esprit selon la théorie de l'esprit socialement étendu plutôt que selon la théorie de l'esprit étendu. En effet, selon Gallagher (2013, p. 8-11), la conception énactive de la cognition que suppose la théorie de l'esprit socialement étendu permet d'éviter les problèmes qui accablaient la théorie de l'esprit étendu de Clark et Chalmers que l'on a vus précédemment, à savoir le « causal coupling-constitution fallacy », l'enflure cognitive [cognitive bloat] et le « mark of the mental ».

Pour montrer que l'adoption de l'approche énactive permet d'éviter l'argument du « coupling-constitution fallacy », Gallagher utilise, entre autres, l'exemple d'une étude de Danziger, Levav et Avnaim-Pesso (2011) qui montre que

le pourcentage de décisions favorables [d'une ou d'un juge à la cour] chute progressivement de 65 % à presque zéro au cours de chaque session de décision [par exemple, entre le petit déjeuner et le dîner] et revient brusquement à 65 % après une pause [alimentaire]. (Danziger *et al.*, 2011, p. 6889)

Cette observation suggère que « les décisions judiciaires peuvent être influencées par des variables externes qui ne devraient pas avoir d'incidence sur les décisions de justice » (Danziger et al., 2011, p. 6889). Étant donné ce type de résultat, on peut se demander si l'état de l'estomac de la juge est un constituant de ses processus cognitifs qui s'occupent des problèmes légaux. La réponse de Gallagher (2013, p. 8) est qu'il n'est pas un constituant de l'esprit parce que si l'état de l'estomac de la juge a un impact sur le verdict qu'elle énonce, il n'a cependant pas d'incidence sur l'institution juridique avec laquelle la juge est étroitement couplée. Par exemple, l'état de l'estomac de la juge n'a pas d'incidence sur les règles utilisées pour faire une preuve légale.

Gallagher soutient également que la théorie de l'esprit socialement étendu n'est pas touchée par l'argument de l'enflure cognitive. Gallagher (2013, p. 9-10) pense que cet argument n'affecte pas cette théorie parce que, dans l'approche énactive, les institutions cognitives qui font partie de l'activité cognitive d'un agent sont les institutions cognitives avec lesquelles l'agent est *correctement couplé* (Gallagher utilise aussi le terme « engagé »). En d'autres mots, ce sont les couplages énactifs qui impliquent des facteurs neuraux et non neuraux qui circonscrivent le domaine de la cognition d'un individu (Gallagher, 2013, p. 10). Selon ce critère, les numéros de téléphone que connait l'opérateur ne font pas partie de la cognition d'un individu qui peut appeler cet opérateur quand il le veut avec son cellulaire. Cet individu, en effet,

n'est pas couplé de manière adéquate avec ces numéros de téléphone en ce sens que ces derniers ne causent pas de modifications neurales dans le cerveau de cet individu.

Enfin, Gallagher soutient que la théorie de l'esprit socialement étendu n'est pas affectée par l'argument du « Mark of the mental ». Il démontre cela en remettant en question la thèse d'Adams et Aizawa (vue précédemment) selon laquelle est mental tout état ou processus qui a un contenu non dérivé. Gallagher formule deux critiques à propos de cette définition. La première est qu'elle est problématique en elle-même. Il en est ainsi parce que, selon Gallagher et Miyahara (2012, p. 139), ce qu'Adams et Aizawa appellent un contenu non dérivé n'est en fait pas « non dérivé » parce qu'il est dérivé des interactions que les personnes entretiennent entre elles dans leur environnement. En effet, pour Gallagher et Miyahara, l'intentionnalité première (c'est-à-dire celle qui est non dérivée) ou, plus précisément, ce qu'ils appellent l'intentionnalité opérative [operative intentionality] est intrinsèque aux actions et aux mouvements des systèmes vivants. Pour illustrer cela, on peut prendre l'exemple de Miyahara repris dans un article écrit par ce dernier et Gallagher (2012, p. 136) : si je suis en auto et que je vois quelqu'un qui tourne la tête et regarde rapidement du côté droit et du côté gauche à la croisée d'une rue que je m'apprête à traverser, mon réflexe de ralentir et de me disposer à presser sur les freins sera alors directement causée par les mouvements de la tête de la personne, et non par une quelconque inférence que j'opèrerais suivant ces mouvements de tête.

La deuxième critique est que cette définition suppose la conception computationnelle/fonctionnaliste de l'esprit que rejette Gallagher. En effet, selon ce dernier :

[l]a cognition ne concerne pas le contenu (qu'il soit dérivé ou non) porté par des véhicules (neuraux ou extraneuraux); la cognition est un engagement actif et émotionnel avec le monde grâce auquel nous sommes capables de résoudre des problèmes, de contrôler des comportements, de comprendre, de juger, d'expliquer et, de manière générale, de faire certains types de choses — dont une grande partie est façonnée constitutivement par des outils, des facteurs environnementaux, des pratiques sociales, etc. (Gallagher, 2013, p. 11)

Bref, pour Gallagher, la cognition est *avant tout* constituée par des activités qui engagent tout l'organisme vivant avec le monde. Je mets « avant tout » en *italique* parce que Gallagher spécifie que « les concepts d'attitude propositionnelle, d'état mental, de représentation, de véhicule et même de contenu non dérivé sont dérivés [de telles activités de l'agent] et *sont inexplicables à moins de faire référence à de telles activités* » (Gallagher, 2013, p. 11, mes *italiques*).

ANNEXE B

SUR LA THÉORIE DES SYSTÈMES DYNAMIQUES

L'expression « théorie des systèmes dynamiques » (TSD) est quelque peu trompeuse parce que la *théorie* des systèmes dynamique n'est pas en elle-même une théorie *scientifique* du monde naturelle (Beer, 2014, p. 135) comme le sont, par exemple, la théorie de l'évolution, la mécanique classique ou la théorie quantique. En effet, *la TSD est plutôt une branche des mathématiques* contenant les outils formels permettant de décrire le comportement des systèmes dynamiques.

À quoi réfère la notion de système dynamique que la TSD permet d'appréhender? Dans l'expression « système dynamique », le concept « système » réfère soit à un système naturel, soit à un modèle mathématique de ce système naturel « comme le voit et le conceptualise un observateur » (Thompson, 2007, p. 38-39). Le qualificatif « dynamique », quant à lui, indique que l'état de ce système change dans le temps (Shapiro, 2019, p. 146; Thompson, 2007, p. 38-39). Les systèmes dynamiques sont

donc des systèmes naturels (que l'on peut représenter par des modèles mathématiques) qui changent dans le temps¹⁵³.

Parce qu'à peu près tous les systèmes naturels changent avec le temps, la plupart d'entre eux sont des systèmes dynamiques (Spivey, 2020, p. 42). Par exemple, le marché boursier américain qui met en scène des centaines de millions d'investisseurs en interaction les uns avec les autres est un système dynamique (Spivey, 2020, p. 42). Le cortex préfrontal où des centaines de millions de neurones sont en interaction les uns avec les autres est lui aussi un système dynamique (Spivey, 2020, p. 42). Une personne humaine est un système dynamique. Il y a, en outre, des systèmes dynamiques composés de deux personnes. Par exemple, un parent et son nourrisson forment un tel système de même que les personnes formant un vieux couple qui va bien (section 2.1.3). Par ailleurs, il existe des systèmes dynamiques composés d'une personne et d'une (ou plusieurs) institution cognitive (section 2.1.4). Par exemple, les avocates et les avocats forment un système de ce type avec le système légal suivant lequel elles ou ils opèrent des raisonnements juridiques (section 2.1.4). Enfin, un humain et le monde de significations où il vit forment un système dynamique.

Il arrive souvent que l'on dise de ces systèmes dynamiques qu'ils sont *complexes*. Cela signifie que la manière dont ils se comportent « n'est ni aléatoire, ni ordonné et prévisible, mais un peu entre les deux, [parce que] présentant des schémas [comportementaux] instables et changeants » (Thompson, 2007, p. 40).

¹⁵³ Pour être plus précis, on pourrait dire, comme me l'a fait remarquer un membre du jury de ma thèse, qu'un système dynamique est un système dont le prochain état est fonction de son état actuel. Le modèle mathématique du pendule simple, que je présente à la fin de cette section, illustre bien cela.

Lorsqu'on étudie le comportement d'un système dynamique, on tente de faire essentiellement trois choses (Shapiro, 2019, p. 146-147). Premièrement, on essaie d'identifier les parties qui changent dans le système. Deuxièmement, on tente de déterminer toutes les manières possibles suivant lesquelles ces parties pourraient changer. Troisièmement, on tente de formuler une règle d'évolution « qui décrit la manière dont le système évolue, ou change, d'un instant à l'autre » (Shapiro, 2019, p. 147). On essaie généralement de construire une telle règle, entre autres, parce qu'un système dynamique est un système « qui est mieux caractérisé par la façon dont il évolue dans le temps, et non par un état stable particulier qu'il pourrait brièvement présenter » (Spivey, 2020, p. 140-141).

Pour illustrer cela, Shapiro (2019, p. 146-147) utilise le modèle mathématique du pendule simple. Ce dernier décrit une masse ponctuelle qui est fixée à l'extrémité d'un fil inextensible sans masse et qui oscille sous l'effet de la pesanteur (Pendule simple, 2020) (comme on peut le constater, il s'agit ici d'un modèle mathématique et pas d'un système réel). Lorsqu'on s'imagine que la masse dans le modèle du pendule oscille, une première observation que l'on peut faire est que la valeur de deux variables de ce système dynamique change d'état, à savoir la vitesse angulaire (υ) et la position (θ) (Shapiro, 2019, p. 147). De cela on peut tirer que « [1]'espace d'état d'un pendule comprend toutes les vitesses et positions possibles que pourrait présenter le pendule » (Shapiro, 2019, p. 147). Troisièmement, la règle d'évolution du pendule est celle-ci : υ = $d\theta/dt$. Suivant cette dernière, la vitesse angulaire et la position verticale du pendule au temps t_1 peuvent être déterminées à partir de la vitesse angulaire et de la position du pendule au temps t_0 (Shapiro, 2019, p. 147).

ANNEXE C

ÉMERGENCE ET ÉNACTIVISME AUTOPOÏÉTIQUE

Ils existent plusieurs conceptions différentes de l'émergence (Bunge, 2003a; Hempel et Oppenheim, 1948, p. 146-152; Humphreys, 2016; Kim, 1999, 2006). Dans cette thèse, par souci de cohérence, j'adopte la conception qu'ont développée deux défenseurs importants de l'énactivisme autopoïétique, à savoir celle de Thompson et Varela (2001, p. 420; voir aussi Thompson, 2007, Appendice B). Ceux-ci comprennent l'émergence de la manière suivante :

Définition : Un réseau R de composants interreliés présente un processus émergent E avec des propriétés émergentes P si et seulement si :

- (1) E est un processus global qui instancie P et qui découle de la dynamique non linéaire D des interactions locales entre les composants de R.
- (2) E et P ont une influence (« descendante ») du global sur le local qui détermine la dynamique D des composants de R.

Et (possiblement):

(3) E et P ne sont pas exhaustivement déterminés par les propriétés intrinsèques des composants de R, c'est-à-dire qu'ils présentent un « holisme relationnel ». (Thompson et Varela, 2001, p. 420)

Suivant Thompson (2007, p. 418-431), je vais examiner cette définition. D'abord, il est important de bien distinguer conceptuellement E et P. En effet, E est un phénomène

temporel alors que P est atemporelle (Thompson, 2007, p. 418). Suivant cela, il faut comprendre que P est instanciée ou réalisée par E qui se déploie dans le temps. Par exemple, un processus (temporel) autopoïétique instancie la propriété (statique) d'être en vie (Thompson, 2007, p. 418-419). Il faut ajouter à cela que E qui instancie P présente des capacités causales (Thompson, 2007, p. 419)¹⁵⁴. D'une part, il provoque des changements dans l'environnement où il évolue. Par exemple, une bactérie altère la quantité de sucre se trouvant dans l'environnement où elle évolue. D'autre part, E qui instancie P crée un contexte où se déroulent des événements qui ne pourraient pas avoir lieu à l'extérieur de ce contexte. Par exemple, la synthèse des protéines a lieu dans un système vivant (une cellule) et elle ne pourrait pas avoir lieu ailleurs (Thompson, 2007, p. 419).

Je vais maintenant examiner la proposition 1. Cette dernière dit que E qui instancie P émerge de la dynamique non linéaire D qu'entretiennent les composants de R (lequel est un système dynamique). À première vue, on pourrait penser être en présence d'un système dynamique qui présente (a) un niveau *inférieur* où se trouve les composants de R qui entretiennent une relation D et (b) un niveau *supérieur* où l'on observe E qui instancie P. Mais cette manière de concevoir ce système est incorrecte. Il faut plutôt l'appréhender comme un système où il n'y a pas vraiment de niveaux. En effet, comme l'explique Thompson,

[d]ans un tel système, la distinction entre les parties préexistantes [à savoir les composants de R] et le tout [à savoir E qui instancie P] qui survient [en raison de la dynamique D qu'entretiennent les composants de R] devient

¹⁵⁴ Étant donné ce qui est dit dans le commentaire de la proposition 2 que l'on verra plus loin, il serait peut-être plus prudent de dire que E qui instancie P présente quelque chose qui *ressemble* à des propriétés causales.

problématique. Non seulement le tout émerge des composants, mais aussi les composants émergent du tout. (Thompson, 2007, p. 420, mes italiques)

C'est pour cette raison que Thompson conçoit l'émergence en termes de co-émergence dynamique. Selon cette manière de concevoir l'émergence, « les parties et le tout co-émergent et se spécifient mutuellement l'un l'autre » (Thompson, 2007, p. 431). Bref, il est difficile de concevoir qu'un système qui affiche une telle dynamique puisse présenter des niveaux.

Je vais désormais m'intéresser à la proposition 2. Que signifie « E et P ont une influence déterminante du global au local ("vers le bas") sur la dynamique D des composants de R »? Il faut comprendre cette influence en termes de relations de contraintes.

Que signifie exactement la notion « contrainte » dans ce contexte? Selon la théorie des systèmes complexes, les contraintes peuvent être comprises comme des propriétés relationnelles que les parties possèdent en vertu de leur intégration ou de leur unification (et non de leur agrégation) dans un réseau systémique. La « contrainte » est donc une notion formelle ou topologique (Deacon, 2003). (Thompson, 2007, p. 424)

Dans cette perspective, cette influence est conçue comme «l'influence de l'organisation topologique du système sur les processus qui le constituent » ou comme «l'influence formelle ou topologique du tout en regard de ses parties » (Thompson, 2007, p. 426). Par exemple, le fer acquiert la propriété d'être capable d'incorporer de l'oxygène lorsqu'il se trouve dans l'hémoglobine laquelle propriété, à son tour, a un effet sur la macrostructure et l'organisme vivant (Fuchs, 2018, p. 96). Elle permet à ce dernier, entre autres, de rester en vie. Est-ce que cette influence peut être conçue en termes causaux? Répondre à cette question nécessiterait de se positionner dans le débat sur la nature de la causalité et, plus spécifiquement, dans le débat sur l'émergence et la causalité descendante. Thompson préfère ne pas se prononcer à ce sujet. En effet, il se contente « d'expliciter ce que les théoriciens des systèmes dynamiques ont à l'esprit

lorsqu'ils parlent de l'influence de l'ensemble des parties dans les systèmes complexes » (Thompson, 2007, p. 427).

Enfin, je vais me pencher sur la proposition 3. Il est difficile de saisir le sens exact que Thompson donne à cette dernière. On ne sait pas, en effet, si Thompson défend une interprétation ontologique ou épistémique de l'émergence. En effet, s'il semble accepter la thèse de l'holisme relationnel selon laquelle un système étudié en mécanique quantique est ontologiquement émergent au sens où (a) « l'état de [ce] système n'est pas constitué [ou déterminé] par les états de ses parties » et où (b) « seul l'ensemble du système peut être considéré comme étant dans un état défini» (Thompson, 2007, p. 428, mes italiques), il semble, par ailleurs, indécis sur le type d'émergence qui caractérise les systèmes dynamiques non décomposables étudiés selon les sciences autres (comme les neurosciences) que la physique quantique. À première vue, on peut penser qu'il défend une conception épistémique de l'émergence. De fait, comme le remarque Van Gulick (2011, p. 145), Thompson affirme que « [l]a « non-décomposabilité » et la « décomposabilité » sont des catégories heuristiques épistémologiques et non ontologiques [et qu'il n'a pas] l'intention d'argumenter pour une thèse métaphysique d'holisme ontologique sur la base de la non-décomposabilité » (Thompson, 2007, p. 423).

Mais, comme le souligne ce même Van Gulick (2011, p. 145), Thompson semble soutenir ailleurs une conception ontologique de l'émergence :

Les phénomènes à toutes les échelles ne sont pas des entités ou des substances, mais des processus relativement stables, et puisque ces processus atteignent une stabilité à différents niveaux de complexité, tout en interagissant avec d'autres niveaux, tous sont également réels et aucun n'a de primauté ontologique absolue. (Thompson, 2007, p. 441)

Pour montrer l'ambiguïté de la position de Thompson à ce sujet, j'ajouterais au propos de Van Gulick que Thompson, dans cette citation, emploie la notion de niveau qu'il semble rejeter ailleurs (voir le commentaire de la proposition 2).

Pour rendre la conception de l'émergence de Thompson et Varela plus concrète, je termine cette annexe en donnant deux exemples de systèmes dynamiques qui instancient des propriétés émergentes. Le premier exemple est celui des systèmes vivants. On peut dire que ces systèmes qui sont en vie émergent de l'interaction dynamique D des composants interreliés du réseau R dont ils sont faits. Autrement dit, ces systèmes vivants *sont* des processus émergents E. Et ces processus, comme on l'a vu (section 3.1), instancient au moins dix propriétés émergentes P, les six premières étant qu'ils sont (1) opérationnellement clos, (2) thermodynamiquement ouverts, (3) précaires, (4) adaptatifs, (5) autonomes et (6) nécessairement cognitifs.

Le deuxième exemple est le cas du couple de personnes âgées vu à la section 2.1.3. Ces dernières sont les composantes interreliées d'un réseau R. Le système dynamique qu'elles forment présente un processus émergent E (le mécanisme mental de la récupération des souvenirs) qui instancie des propriétés émergentes P, lesquelles sont au nombre de trois : (1) les souvenirs sont plus détaillés, (2) les souvenirs sont plus riches émotionnellement et plus vifs et (3) les souvenirs présentent une plus-value de compréhension.

ANNEXE D

QUELQUES BRÈVES RÉFLEXIONS SUR LES MODÈLES ÉNACTIFS DE DE HAAN, DE NIELSEN ET WARD ET DE FUCHS

Alors que la rédaction de ma thèse était bien entamée, sont parus les travaux en philosophie de la psychiatrie de Sanneke de Haan (2020b, 2020a) et de Kristopher Nielsen et Tony Ward (2018, 2020a). C'est aussi à ce moment que je me suis intéressé plus sérieusement aux recherches du philosophe et psychiatre Thomas Fuchs (2018, voir particulièrement le chapitre 7). Les projets que développent ceux-ci s'inscrivent en grande partie dans le cadre de l'énactivisme autopoïétique. Pour cette raison, leurs idées et les miennes sont en grande partie compatibles. Elles sont aussi à certains égards similaires. Cependant, elles ne sont pas identiques. Dans ce qui suit, j'explique brièvement en quoi nos idées sont à la fois semblables et dissemblables.

Les idées que de Haan, Nielsen, Ward, Fuchs et moi développons partagent au moins deux points en communs. Premièrement, nous nous entendons sur la thèse selon laquelle les RDoC sont réductionnistes et neurocentrés (de Haan, 2020a, p. 4; Fuchs, 2018, p. 254-255; Nielsen et Ward, 2018, p. 805). En ce qui la concerne, de Haan affirme que les RDoC « supposent une conception neuroréductionniste de ce que sont les troubles mentaux » (de Haan, 2020b, p. 26). Selon elle, « la recherche [au sein des RDoC] est orientée vers la découverte des "mécanismes neuronaux sous-jacents" aux troubles psychiatriques » (de Haan, 2020a, p. 4). Pour leur part, Nielsen et Ward

soutiennent que les RDoC « sont trop neurocentrés » (Nielsen et Ward, 2018, p. 805). Pour sa part, Fuchs classe nettement les RDoC dans le paradigme neuroréductionniste en psychiatrie auquel il n'adhère pas (Fuchs, 2018, p. 253-254).

Deuxièmement, de Haan, Nielsen, Ward, Fuchs et moi nous accordons sur le fait que l'on doit adopter une posture épistémique pluraliste et non réductionniste lorsqu'on explique les troubles mentaux. En effet, la philosophe de Haan pense que les troubles mentaux ont des dimensions biologique, psychologique, socioculturelle et existentielle que nos modèles doivent être capables d'intégrer les unes aux autres (de Haan, 2020a, p. 4)¹⁵⁵. Nielsen et Ward adoptent aussi une approche pluraliste. En effet, ils proposent un modèle 3E (incrusté [embedded], incarné [embodied] et énactif [enactive]) qui saisit (a) quels facteurs biologiques constituent un phénomène clinique 156 d'un trouble mental (Nielsen et Ward, 2018, p. 815) et (b) de quelle manière la composition de ce phénomène clinique est en partie déterminée par l'environnement physique et socioculturel dans lequel est incrusté l'organisme (Nielsen et Ward, 2018, p. 815). Fuchs adopte également une posture non réductionniste. En effet, il affirme que les processus mentaux dysfonctionnels « ne peuvent pas être réduits au cerveau ou à des activités neuronales localisées [...] » (Fuchs, 2018, p. 253). Aussi, il ajoute que ces processus « sont incarnés, intrinsèquement intentionnels et liés au contexte [...]» (Fuchs, 2018, p. 253). Enfin, Fuchs dit qu'«ils sont inséparables du monde

¹⁵⁵ La dimension existentielle est moins souvent mentionnée dans la littérature. Elle réfère « aux postures réflexives que nous prenons relativement à nos expériences et nos situations» (de Haan, 2020a, p. 4). Pour exemplifier cette dimension, de Haan donne l'exemple « de la peur de la peur d'avoir une attaque panique » (de Haan, 2020a, p. 4).

¹⁵⁶ Nielsen et Ward parlent aussi d'un riche échantillon de quelques symptômes qui forment un système. Ce serait d'abord ce système qu'il faudrait expliquer. Pour appréhender ces systèmes de symptômes, Nielsen et Ward ont développé un cadre explicatif métaméthodologique qu'ils ont nommé le « Relational Analysis of Phenomena » (RAP) (voir Nielsen et Ward, 2020b). Je n'entre pas dans les détails de ce dernier parce que je n'en ai pas besoin pour montrer quelles sont les ressemblances et les différences entre nos modèles respectifs.

intersubjectif des significations partagées et des interactions partagées » (Fuchs, 2018, p. 253). L'approche écologique de Fuchs (2018, p. 255), basée sur la conception incarnée et énactive de la cognition, saisit ce monde intersubjectif de significations dans lequel se trouvent les cerveaux des différentes personnes où ont lieu ces processus mentaux dysfonctionnels.

Je vais maintenant expliquer pourquoi ce que font de Haan, Nielsen, Ward et Fuchs dans leurs travaux n'est pas identique à ce que je fais dans les miens.

Premièrement, les problématiques qui nous intéressent ne sont pas exactement les mêmes. En effet, bien que de Haan, Nielsen, Ward et Fuchs s'intéressent aux RDoC et proposent des approches énactives en psychiatrie qui seraient supérieures à leur avis au cadre de recherche du NIMH, elle et ils ne s'intéressent pas *spécifiquement* aux RDoC comme je l'ai fait dans ma thèse. Plus précisément, elle et ils ne font pas une présentation détaillée des RDoC ni n'exposent et n'explicitent leurs bases philosophiques¹⁵⁷. En effet, de Haan veut développer un modèle énactif supérieur aux modèles unidimensionnels [*one-sided*] (dont les RDoC font selon elle partie), *mais aussi* aux modèles bidimensionnels [*two-sided*], aux modèles intégratifs comme le très connu modèle biopsychosocial (Engel, 1977, 1980) et au récent modèle de Borsboom (2017) selon lequel les troubles mentaux sont des réseaux de symptômes. Quant à eux, Nielsen et Ward veulent développer une approche énactive pour classifier (Nielsen et

¹⁵⁷ On pourrait contre-argumenter que Nielsen et Ward (2018, p. 803-806) font une telle analyse philosophique lorsqu'ils caractérisent les RDoC selon les critères que donnent Kendler et Zachar (2007) pour analyser les classifications psychiatriques. Cette objection est partiellement juste parce qu'il me semble que mon analyse est plus détaillée et exhaustive que la leur. En effet, Nielsen et Ward font connaître peu de citations des défenseurs et des critiques des RDoC. Aussi, ils ne mettent pas ces citations directement en lien avec une théorie philosophique précise comme je le fais (entre autres) quand je tente de démontrer qu'une interprétation adaptationniste de la TBS de Boorse est au fondement de la conception du trouble mental des RDoC (voir section 1.5.3).

Ward, 2018), concevoir (Nielsen et Ward, 2018, 2020a) et expliquer (Nielsen et Ward, 2020b) les troubles mentaux qui serait supérieure à celle des RDoC, *mais aussi* à celle du DSM et aux modèles semblables à celui de Borsboom. Enfin, Fuchs, dans son livre *Ecology of the brain* (2018), s'attaque *de manière générale* à l'approche neuroréductionniste en sciences cognitives (donc pas spécifiquement aux RDoC comme je le fais) et défend une approche écologique (incarnée et énactive). Ce n'est qu'au chapitre 7 de son livre que Fuchs traite rapidement des RDoC. C'est aussi dans cette portion de son livre qu'il montre en quoi la posture écologique qu'il adopte en sciences cognitives a des conséquences en psychiatrie et sur la manière de concevoir les troubles mentaux.

Deuxièmement, bien qu'elles ne soient pas incompatibles avec la conception énactive et objective de la santé et du trouble (CÉOST) et partagent avec elle des caractéristiques, les conceptions du trouble mental que proposent de Haan, Nielsen, Ward et Fuchs sont différentes de la CÉOST. Je présente maintenant ces dernières et explique après en quoi elles sont semblables et dissemblables à la CÉOST.

La conception du trouble mental de de Haan suppose que les troubles mentaux affectent les personnes (et pas seulement leur cerveau) qu'elle définit comme « des êtres corporels, sociaux et réflexifs » (de Haan, 2020a, p. 11). La conception de de Haan dit que, chez les personnes, les troubles psychiatriques sont « des troubles de la fabrication de la signification [sens-making] » (de Haan, 2020a, p. 10). Cette activité de création de sens peut être définie « comme l'interaction évaluative d'un organisme [ici une personne] avec son environnement » (de Haan, 2020a, p. 10). Cette interaction est dite « évaluative » parce que c'est d'elle qu'émerge le monde symbolique rempli des valeurs de la personne. Au sein de ce monde, exister, pour cette dernière, ne consiste pas seulement à être en vie. Cela consiste plutôt à habiter un monde de significations.

Parce que lorsqu'il y a présence d'un trouble mental « les interactions évaluatives d'une personne et de son monde s'égarent » (de Haan, 2020a, p. 10), il y a quelque chose dans le sens que la personne donne à son monde qui ne va plus. Comme le dit de Haan,

à un niveau très général, nous pouvons dire que la façon dont la personne donne un sens à son monde est biaisée dans une direction spécifique : le monde semble trop menaçant, ou [trop] dénué de sens, ou [trop] significatif, ou [trop] chaotique. (de Haan, 2020a, p. 10)

De Haan attribue *quatre caractéristiques* à la fabrication de sens pathologique (voir de Haan, 2020b, p. 206-215) :

- (1) Souvent, quand il y a présence d'un trouble mental, *la fabrication* de sens par la personne est inadéquate relativement au contexte où elle se trouve. Par exemple, une personne mince qui se perçoit grosse et qui, pour cela, modifie sa manière de s'alimenter (ou cesse de s'alimenter) adopte une posture anormale.
- (2) Fréquemment, lorsqu'il y a un trouble mental, *la fabrication de sens* qui n'est pas en adéquation avec le contexte (point 1) *est* également *inflexible*. Par exemple, une personne qui présente un TOC et qui, dans sa demeure qui est propre, se lave systématiquement et compulsivement les mains une centaine de fois par jour par peur d'être infectée par des germes présente une telle rigidité.
- (3) Souvent, quand il y a présence d'un trouble mental, *la personne n'est pas capable de prendre du recul ou d'avoir une perspective relativement à sa condition psychiatrique*. Par exemple, quand une personne a une dépression majeure, elle est incapable de prendre du

recul et d'imaginer qu'un jour elle sera à nouveau heureuse (de Haan, 2020b, p. 212).

(4) Souvent, lorsqu'une personne présente une fabrication de sens dysfonctionnelle, elle souffre.

En général, affirme de Haan, lorsqu'une personne présente un trouble psychiatrique, ces quatre caractéristiques sont présentes (de Haan, 2020b, p. 212). Toutefois, elles n'ont pas toujours à l'être (de Haan, 2020b, p. 212). Par exemple, une personne psychopathe ne souffre pas de sa condition.

Enfin, de Haan précise que, pour qu'il y ait présence d'un trouble mental chez la personne, cette fabrication malsaine de sens qui entraine une perte de sens, ou un surplus de sens, ou une frayeur, ou un chaos, doit être stable (de Haan, 2020a, p. 10, mes italiques). Par exemple, la perte de sens qu'entraine une dépression majeure ne dure pas seulement une journée. Elle s'étend plutôt sur plusieurs mois, voire sur plusieurs années.

Par ailleurs, il faut mentionner que de Haan distingue les troubles physiques des troubles mentaux. Elle dit qu'alors que, « dans les troubles somatiques, la fabrication de sens chez les patients peut être secondairement affectée, les troubles psychiatriques sont *principalement* des problèmes de la fabrication de sens » (de Haan, 2020a, p. 10).

Ce que veut dire de Haan par « secondairement affecté » est que les problèmes de fabrication de sens sont des effets secondaires du trouble somatique. Cela implique que si on guérissait le trouble somatique, ces derniers disparaitraient. Par exemple, si on traitait la tumeur cérébrale qui perturbe l'interaction évaluative qu'une personne entretient avec son monde, cette perturbation disparaitrait et la fabrication de sens qu'opère la personne redeviendrait normale. Donc, les troubles somatiques ont des

causes physiques et c'est en traitant ces dernières qu'on les guérit et qu'on fait disparaitre le trouble de la fabrication de sens qui les accompagnait (de Haan, 2020a, p. 11). Quant aux troubles mentaux, de Haan affirment que leur existence est plutôt due à des raisons (de Haan, 2020a, p. 11). En effet, l'interaction qu'entretiennent les personnes avec le monde auquel elles donnent un sens est modulée par des raisons. Que le monde où évoluent les personnes puisse changer explique que cette interaction soit précaire. En effet, certaines raisons qui guidaient une personne et la confortaient à un certain moment dans un monde donné peuvent cesser d'être viables à un autre moment et/ou dans un autre monde. Pour illustrer cela, de Haan donne l'exemple suivant :

[...] il peut être judicieux d'être très prudent et attentif au danger à un moment et dans un lieu donné, mais dans des circonstances plus sûres, l'incapacité à faire confiance peut vous empêcher de donner un sens approprié à votre situation et aux personnes qui vous entourent. (de Haan, 2020b, p. 201)

De Haan tire la conclusion suivante de ce qui précède : « Les troubles psychiatriques sont [...] énactés : ils se dissolvent si l'on parvient à changer [en mieux] notre façon d'interagir avec le monde. » (de Haan, 2020a, p. 11) Par exemple, une personne réfugiée de guerre qui, dans le pays paisible où elle immigre, cesse graduellement d'être ultra attentive aux dangers potentiels qu'elle perçoit partout rétablit une interaction normale avec son monde et le trouble panique dont elle souffrait à son arrivée s'atténue¹⁵⁸.

Maintenant que l'on a vu la conception du trouble mental de de Haan, on peut examiner celle de Nielsen et Ward. D'un point de vue ontologique, ceux-ci conçoivent une

¹⁵⁸ De Haan remarque à juste titre que ce n'est pas en changeant la manière d'interagir avec le monde que l'on peut guérir d'un trouble physique (de Haan, 2020a, p. 11). Par exemple, une personne ne guérit pas du cancer du pancréas en se mettant en paix avec le fait que sûrement celui-ci la tuera.

psychopathologie comme un phénomène (cognitif, perceptuel et comportemental) qui présente une complexité constitutionnelle et causale (Nielsen et Ward, 2020a, p. 120). Ce phénomène, selon eux, émerge de réseaux de mécanismes (Nielsen, 2020) et, plus précisément, d'un individu qui est un système (mécaniste) « cerveau-corpsenvironnement » (Nielsen et Ward, 2018, p. 813).

Cela posé, on peut maintenant examiner suivant quelles conditions le phénomène (cognitif, perceptuel et comportemental) susmentionné peut être qualifié de dysfonctionnel selon ces auteurs. Ces derniers affirment qu'

[e]st considéré comme mentalement dysfonctionnel tout ensemble de comportements (y compris la cognition, la perception ou tout ce que fait l'organisme) exécutés par un organisme qui viole de manière significative ses propres normes fonctionnelles, en ce sens qu'il [cet ensemble de comportements] agit à l'encontre de ses propres besoins d'*autoentretien* et d'*adaptation*. (Nielsen et Ward, 2020a, p. 116)

Les notions « autoentretien » (ou « automaintien ») et « adaptation » sont à comprendre ici de la manière dont elles sont définies dans le cadre de l'énactivisme autopoïétique (voir la section 3.1 et 3.2.1 de ma thèse). Quant aux normes fonctionnelles, elles sont conçues par Nielsen et Ward (2020a, p. 115) comme des tendances comportementales qui permettent aux organismes de s'automaintenir, de s'adapter et, éventuellement, de s'épanouir en fonction des contraintes de l'environnement socioculturel. Il faut aussi noter que ces normes fonctionnelles opèrent en même temps sur les niveaux biologique et socioculturel (Nielsen et Ward, 2020a, p. 111 et 115). Pour illustrer cela, Nielsen et Ward donnent un exemple de Maiese (2016) qu'ils formulent de la manière suivante : « Nous souhaitons être de bons conducteurs non seulement pour éviter les accidents [comportement (biologique) de base], mais aussi pour démontrer notre maîtrise [comportement socioculturel complexe], ce qui a des implications sociales positives pour nous. » (Nielsen et Ward, 2020a, p. 115)

Nielsen remarque que, même s'ils nuisent aux normes fonctionnelles de la personne, les comportements dysfonctionnels associés aux troubles mentaux persistent parce qu'ils ont pour but d'automaintenir (de manière inadéquate) la personne (Nielsen, 2020). Par exemple, s'automutiler peut être une manière (inadéquate) de réguler ses émotions (Nielsen et Ward, 2020a, p. 115). On peut aussi penser que ces comportements dysfonctionnels persistent parce qu'ils forment un réseau causal circulaire dans lequel ils s'automaintiennent les uns les autres (Nielsen et Ward, 2020b). Par exemple, dans la boulimie, le cycle « ingestion massive d'aliments → sentiment de culpabilité → purge → ingestion massive d'aliments » est un réseau de ce type (Nielsen et Ward, 2020b, p. 172). Le cycle de l'évitement, dans le trouble obsessif compulsif, en est un autre (Nielsen et Ward, 2020b, p. 172).

Enfin, on peut exposer la conception du trouble mental de Fuchs (2012, 2018, section 7.2). Pour la comprendre, il faut connaître une hypothèse de Fuchs sur le cerveau et son concept de causalité circulaire. D'abord, il convient de présenter son hypothèse :

Dès la naissance, le cerveau est intégré aux interrelations [qui ont lieu] entre la personne et l'environnement et il [le cerveau] doit être conçu comme un organe de médiation et de transformation des processus biologiques, mentaux et sociaux qui sont liés dans une interaction circulaire. (Fuchs, 2012, p. 332)

Fuch ajoute à cela que « [d]ans cette interaction [où le cerveau est l'organe de médiation], la subjectivité [...] joue un rôle central, tout comme les interactions sociales de la personne avec les autres » (Fuchs, 2012, p. 332).

Pour comprendre la manière dont le cerveau est intégré dans cette relation entre la personne et l'environnement, il faut avoir en tête que les organismes vivants (comme les personnes) s'inscrivent et évoluent dans un réseau où un type particulier de causalité est opérant. Fuchs nomme ce genre de causalité la *causalité circulaire* (Fuchs, 2012, p.

332, 2018, p. 94). Celle-ci prend, entre autres, deux formes. Il y a, d'abord, la causalité circulaire *verticale* qui réfère aux relations causales existant, dans un organisme vivant, entre les parties et les différentes macrostructures formées par ces parties (où l'organisme lui-même est une macrostructure) (Fuchs, 2018, p. 94-98). La relation causale opérante ici en est une, entre autres, de contrainte où (1) la macrostructure sélectionne les propriétés et les comportements de ces composants et en bloque d'autres et où (2) ces composants, à leur tour, ont un impact sur la macrostructure. Pour illustrer cela sur le plan biologique, Fuchs (2018, p. 96) donne l'exemple du fer qui acquière la propriété d'être capable d'incorporer de l'oxygène lorsqu'il se trouve dans l'hémoglobine. Cette propriété, à son tour, a un effet sur la macrostructure et l'organisme vivant. Elle permet à ce dernier, entre autres, de rester en vie. On peut illustrer ce phénomène de contrainte rétroactive à d'autres niveaux d'organisations. Par exemple, l'expérience subjective d'une personne a un impact sur ses processus physiologiques (Fuchs, 2012, p. 334) et ces derniers ont à leur tour une influence sur l'expérience phénoménologique de cette dernière.

Maintenant que l'on connait la causalité circulaire verticale, il convient de s'intéresser à la causalité circulaire *horizontale* (Fuchs, 2018, p. 98-99). Il existe deux sortes de causalité horizontale. La première réfère aux relations causales ayant lieu entre les parties internes opérant sur un même niveau dans un organisme vivant (Fuchs, 2018, p. 98). Les relations causales qui existent entre les cellules qui composent l'organisme vivant (Fuchs, 2018, p. 98) ou celles qui existent entre les différents circuits du cerveau sont des relations causales de ce type. Le deuxième type de causalité circulaire horizontale réfère aux relations causales qui existent entre l'organisme vivant et l'environnement où évolue ce dernier (Fuchs, 2018, p. 98). C'est, par exemple, en percevant l'environnement, en se déplaçant en lui et en agissant dans celui-ci que l'organisme vivant entre dans ce type de relation causale circulaire horizontale. Quand

l'organisme vivant est une personne, ce type de causalité est à l'œuvre quand celle-ci entre en relation sociale avec d'autres personnes ou avec la culture d'un endroit.

Maintenant que l'on sait de quelle façon Fuchs conçoit le cerveau et les causalités circulaires verticale et horizontale, on peut présenter sa conception des troubles mentaux. Selon elle, ces derniers ont au moins trois particularités. La première est que les troubles mentaux

[...] sont caractérisés [...] par une perturbation de la causalité circulaire verticale, c'est-à-dire [par une perturbation] de l'interaction entre les processus de niveau inférieur et les capacités supérieures de l'organisme. [...] [...] [Fuch spécifie que] [cette perturbation] affecte principalement la relation du patient [ou de la patiente] à lui-même ou à elle-même, ce qui [à son avis] influence continuellement le cours de la maladie, y compris les processus neuronaux sur le plan micro. (Fuchs, 2018, p. 256)

Fuchs illustre, entre autres, cette perturbation de boucle causale circulaire verticale avec le cas des troubles anxieux. Il dit que ces derniers présentent ce type de patron causal :

[A]pparition des caractéristiques physiologiques du stress (activation du système [nerveux] sympathique, éveil, pouls plus rapide, essoufflement) \rightarrow perception des symptômes physiques comme « menaçants » \rightarrow [...] évaluations catastrophistes \rightarrow augmentation du stress physiologique, etc. (Fuchs, 2018, p. 258)

Conformément avec ce qu'il dit précédemment sur la causalité verticale, Fuchs observe que l'expérience subjective qu'a la personne anxieuse influe sur le cours (et je dirais sur le maintien dans le temps) de sa maladie (Fuchs, 2018, p. 258). En effet, plus une personne est anxieuse, plus elle a des symptômes somatiques désagréables et plus elle a de tels symptômes, plus elle est anxieuse.

Le deuxième attribut des troubles mentaux est qu'ils

[...] sont caractérisés par une perturbation de la *causalité circulaire* horizontale, c'est-à-dire [par une perturbation] des relations sociales et de la capacité à répondre adéquatement aux demandes et aux attentes des autres. Cela conduit [selon Fuchs] à des boucles de rétroaction négative dans les cycles sociofonctionnels, lesquels ont également une influence cruciale sur l'évolution de la maladie. (Fuchs, 2018, p. 256)

On peut utiliser le cas de la personne anxieuse, qui figure précédemment, pour illustrer ce qu'est concrètement une perturbation de la causalité circulaire horizontale. Il est possible que cette personne évite de sortir de chez elle parce qu'il lui arrive de faire des crises de panique. Ainsi, plus elle ne sort pas de chez elle, moins elle entretient de relations sociales. Cet isolement est évidemment mauvais pour elle et contribue à la détérioration de sa condition. Et si sa condition se détériore, elle voudra alors encore moins sortir de chez elle. En conséquence, une boucle causale horizontale dysfonctionnelle sera maintenue.

La troisième caractéristique est que

[c]es deux types de processus causaux circulaires sont [médiés par l'intermédiaire du] cerveau, mais ne peuvent pas être localisés en lui. C'est pourquoi [, selon Fuchs,] la réduction des troubles mentaux à des troubles cérébraux n'est en principe pas possible. (Fuchs, 2018, p. 256)

Fuchs, ici, adopte une posture antiréductionniste de principe. Il soutient essentiellement que les troubles mentaux ne peuvent pas être réduits à des troubles du cerveau. En effet, ils sont plutôt, comme on vient de le voir, des troubles de la causalité circulaire dans laquelle la personne évolue et qui la traverse. Certes, le cerveau a un rôle (celui d'être un médiateur) dans ce réseau causal, mais il n'est pas le seul à en avoir un.

Maintenant que j'ai présenté les conceptions du trouble mental de de Haan, de Nielsen et Ward et de Fuchs, je peux brièvement expliquer pourquoi elles sont en même temps similaires et différentes de la mienne.

Trois raisons expliquent qu'elles sont semblables à la CÉOST. Premièrement, elles s'inscrivent toutes dans le cadre de l'énactivisme autopoïétique. Par exemple, la conception de Nielsen et Ward contient les notions d'auto-organisation (qui est identique à celle d'automaintien) et d'adaptation. Quant à la conception de de Haan, elle fait un usage massif de la notion de fabrication de significations. Enfin, toutes ces conceptions mettent de l'avant que les organismes et, plus spécifiquement, les personnes entretiennent des liens étroits et rétroactifs avec leur environnement.

Deuxièmement, comme je l'ai déjà dit, elles sont toutes non réductionnistes et pluralistes. En effet, elles supposent que les troubles mentaux ne sont pas réductibles à des troubles du cerveau et que, conséquemment, nos modèles doivent saisir non seulement leurs composantes psychobiologiques, mais aussi leurs composantes corporelles, sociales et culturelles (et même existentielles, selon de Haan).

Troisièmement, nos conceptions tiennent toutes compte de manière plus ou moins explicite de la dimension temporelle des troubles mentaux. En effet, la fabrication de sens (normale ou anormale) dont parle de Haan ne peut s'effectuer que dans la durée. Dans le même ordre d'idées, la causalité circulaire que met de l'avant Fuchs ne peut se déployer que dans un espace temporel. Quant à eux, Nielsen et Ward soutiennent que les troubles mentaux persistent et se maintiennent dans le temps. Similairement, de Haan dit des troubles mentaux qu'ils sont stables dans le temps¹⁵⁹.

¹⁵⁹ J'aimerais mentionner au passage que cette caractéristique temporelle que contient chacune de ces conceptions et la CÉOST les distingue des conceptions de Boorse et de Wakefield qui m'apparaissent plus statiques.

Maintenant qu'on a vu quelques points qu'avaient en commun ces trois conceptions du trouble mental et la CÉOST, je vais examiner brièvement en quoi elles diffèrent de celle-ci. Trois différences me viennent à l'esprit.

Premièrement, la CÉOST met à l'œuvre une théorie des attributions fonctionnelles, à savoir l'approche organisationnelle des fonctions modifiée (AOFm) qui est basée sur l'AOF de Mossio, Moreno et Saborido (Mossio et al., 2009, 2010; Mossio et Moreno, 2015, chapitre 3) (voir section 3.2.1). L'AOFm permet à la CÉOST de déterminer quelles sont les composantes fonctionnelles intérieures ou extérieures à l'enveloppe corporelle d'une personne. Les conceptions de de Haan, de Nielsen et Ward et de Fuchs ne semblent pas conçues pour effectuer une telle tâche¹⁶⁰. Cela s'explique par le fait que de Haan et Fuchs n'utilisent pas de conception de la fonction dans leurs théories du trouble mental. Quant à Nielsen et Ward, ils ne mettent également pas à profit une théorie des attributions fonctionnelles. Certes, leur conception du trouble mental contient la notion de normes fonctionnelles. Cependant, cette dernière n'est pas fondée dans une conception de la fonction. Elle est plutôt issue du concept de normativité naturelle que l'on trouve dans le cadre de l'énactivisme autopoïétique (sections 3.1 et 3.2.1) et aussi dans les travaux de Christensen et Bickard (Christensen, 2012; Christensen et al., 2002) et dans ceux d'Okrent (2017) (voir Nielsen et Ward, 2020a, p. 111)¹⁶¹. La normativité naturelle est un référentiel qui sert à déterminer si un comportement ou une composante d'un système vivant opère normalement ou anormalement. Cependant, cette notion de normativité ne permet pas de dire si une

¹⁶⁰ J'aimerais aussi ajouter au passage que Nielsen et Ward (2020b, p. 113-114) semblent s'opposer à la thèse (que je défends) selon laquelle la cognition peut être constituée par des composantes extracorporelles. En fait, ils semblent plutôt adopter la thèse de l'incrustation de l'esprit dans l'environnement (2020b, p. 114).

¹⁶¹ À noter que l'on trouve aussi la notion de normativité naturelle dans les travaux de Mossio, Moreno et Saborido (Mossio *et al.*, 2009, 2010 ; Mossio et Moreno, 2015). C'est avec les idées de ces derniers et celles contenues dans le cadre de l'énactivisme autopoïétique que je construis la CÉOST.

partie quelconque est ou n'est pas une composante *fonctionnelle* d'un système dynamique BioCognitifSocioCulturel (BCSC) (voir section 3.2.1). Pour faire cette tâche, il faut utiliser l'AOFm.

Deuxièmement, contrairement aux conceptions de de Haan, de Nielsen et Ward et de Fuchs, la CÉOST distingue le trouble et la dysfonction (section 3.2.1). Selon la CÉOST, un trouble est une dysfonction que le système BCSC ne parvient pas à réguler, ce qui cause la dissolution du système (autrement dit, sa mort). Quant à elle, la dysfonction est un bris dans un système que ce dernier parvient à réguler grâce à des mécanismes compensatoires qui lui permettent de rester viable. Cette distinction est pertinente puisque les êtres humains présentent tous des dysfonctions, qu'ils aient un trouble ou non. Cette distinction me permet aussi de proposer que présenter une dysfonction soit l'occasion d'une adaptation (section 3.2.1). Lorsque cette dysfonction se trouve dans une personne, elle est l'occasion d'un changement qui, bien qu'il ne soit pas optimal, peut être viable et bénéfique. Dans le même ordre d'idées, quand cette dysfonction se trouve dans une société, elle peut potentiellement être l'occasion d'un progrès social. Bref, je trouve intéressant et rafraichissant que l'on puisse concevoir la présence d'une dysfonction dans un système de manière positive.

Troisièmement, la CÉOST contient la notion de différenciation organisationnelle développée par Mossio, Saborido et Moreno (2010, p. 166)¹⁶². Cette dernière réfère au fait que les systèmes dynamiques BCSC contiennent nécessairement (et plus précisément produisent et automaintiennent) un ensemble d'entités dynamiques différentes et localisables et un ensemble de mécanismes différents et localisables qui contribuent de manière spécifique et concertée, par l'intermédiaire de l'intégration de

¹⁶² J'aimerais mentionner au passage que cette notion de différenciation organisationnelle, bien qu'elle soit, à mon avis, compatible avec lui, ne se trouve pas dans le cadre de l'énactivisme autopoïétique.

leurs différentes fonctions, à l'automaintien de leur organisation et, conséquemment, à leurs capacités fonctionnelles générales (voir section 3.2.1). Cette notion permet, entre autres, à la CÉOST de saisir qu'un système BCSC (comme une personne ou une société) présentant un trouble (parce que la norme téléologique naturelle [NTN] qui leur est inhérente est violée) tend vers une *in*différenciation organisationnelle. À ma connaissance, les conceptions du trouble mental de de Haan, de Nielsen et Ward et de Fuchs ne saisissent ni dans une personne ni dans une société un tel phénomène.

Cela termine ma présentation des conceptions du trouble mental de de Haan, de Nielsen et Ward et de Fuchs et mon exposition des raisons qui montrent qu'elles sont à la fois similaires à la CÉOST et différentes d'elle. Je ne prétends pas que mon exposition soit exhaustive. Je suis certain que j'aurais pu relever d'autres points qu'elles partagent et d'autres propriétés qui les distinguent. Cela dit, j'aimerais souligner que je trouve les conceptions de de Haan, de Nielsen et Ward et de Fuchs très intéressantes et pertinentes. Que leurs conceptions respectives et la CÉOST soient différentes montre que le cadre de l'énactivisme autopoïétique est un terrain conceptuel fertile qu'il vaut la peine d'explorer.

RÉFÉRENCES

- Adams, F. et Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, *14*(1), 43-64. doi: 10.1080/09515080120033571
- Adams, F. et Aizawa, K. (2008). *The bounds of cognition*. Oxford, Royaume-Uni: Blackwell Publishing.
- American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders* (1^{re} éd.). Washington, DC: American Psychiatric Association Publishing.
- American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2^e éd.). Washington, DC: American Psychiatric Association Publishing
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3^e éd.). Washington, DC: American Psychiatric Association Publishing
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4e éd.). Washington, DC: American Psychiatric Association Publishing
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*. *Text revision* (4^e éd. révisée). Washington, DC: American Psychiatric Association Publishing
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5^e éd.). Washington, DC: American Psychiatric Association Publishing
- Andersen, L. M. (2017). Mechanisms and reduction in psychiatry. Dans M. Massimi, J.-W. Romeijn et G. Schurz (dir.), *EPSA15 selected papers: The 5th conference of the European philosophy of science association in Düsseldorf* (p. 111-124). New York, NY: Springer. doi: 10.1007/978-3-319-53730-6 10

- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. Dans H. Guetzkow (dir.), *Groups, leadership and men: Research in human relations* (177-190). Pittsburgh, PA: Carnegie Press.
- Baron, S. et Miller, K. (2019). *An introduction to the philosophy of time*. Cambridge, Royaume-Uni: Polity Press.
- Barrett, R. J. (1996). *The psychiatric team and the social definition of schizophrenia: An anthropological study of person and illness* (1^{re} éd.). Cambridge, Royaume-Uni: Cambridge University Press.
- Bechtel, W. (2007). Reducing psychology while maintaining its autonomy via mechanistic explanation. Dans M. Schouton et H. L. de Jong (dir.), *The matter of the mind: Philosophical essays on psychology, neuroscience and reduction* (p. 172-198). Oxford, Royaume-Uni: Blackwell Publishing.
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. New York, NY: Routledge.
- Bechtel, W. (2009). Constructing a philosophy of science of cognitive science. *Topics in Cognitive Science*, 1(3), 548-569. doi: 10.1111/j.1756-8765.2009.01039.x
- Bechtel, W. et Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *36*(2), 421-441. doi: 10.1016/j.shpsc.2005.03.010
- Bechtel, W. et McCauley, R. N. (1999). Heuristic identity theory (or back to the future): The mind-body problem against the background of research strategies in cognitive neuroscience. Dans M. Hahn et S. C. S. Stoness (dir.), *Proceedings of the twenty first annual conference of the cognitive science society* (p. 67-72). Mahwah, NJ: Lawrence Erlbaum Associates.
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York, NY: Harper & Row.
- Beck, A. T. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy*, 1(1), 5-37.
- Bedwell, J. S., Gooding, D. C., Chan, C. C. et Trachik, B. J. (2014). Anhedonia in the age of RDoC. *Schizophrenia Research*, *160*(1-3), 226-227. doi: 10.1016/j.schres.2014.10.028.

- Beer, R. D. (2014). Dynamical systems and embedded cognition. Dans K. Frankish et W. M. Ramsey (dir.), *The Cambridge handbook of artificial intelligence* (p. 128-148). Cambridge, Royaume-Uni: Cambridge University Press. doi: 10.1017/CBO9781139046855.009
- Benson, O., Gibson, S. et Brand, S. L. (2013). The experience of agency in the feeling of being suicidal. *Journal of Consciousness Studies*, 20(7-8), 56-79.
- Bercik, P., Denou, E., Collins, J., Jackson, W., Lu, J., Jury, J., ... Collins, S. M. (2011). The intestinal microbiota affects central levels of brain-derived neurotropic factor and behavior in mice. *Gastroenterology*, *141*(2), 599-609. doi: 10.1053/j.gastro.2011.04.052
- Berenbaum, H. (2013). Classification and psychopathology research. *Journal of Abnormal Psychology*, 122(3), 894-901. doi: 10.1037/a0033096
- Bickle, J. et Kostko, A. (2017). Personalized psychiatry and scientific causal explanations: Two accounts. Dans J. Poland et S. Tekin (dir.), *Extraordinary Science and psychiatry: Responses to the crisis in mental health research* (p. 137-162). Cambridge, MA: MIT Press.
- Bitbol, M. et Cohen-Varela, A. (2017). Introduction. Dans *Le cercle créateur : Écrits* (1976-2001) (p. 9-31). Paris : Éditions du Seuil.
- Bitbol, M. et Luisi, P. L. (2004). Autopoiesis with or without cognition: Defining life at its edge. *Journal of The Royal Society Interface*, *I*(1), 99-107. doi: 10.1098/rsif.2004.0012
- Bogen, J. et Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, 97(3), 303-352. doi: 10.2307/2185445
- Bolton, D. (2008). What is Mental Disorder? An essay in philosophy, science, and values. Oxford, Royaume-Uni: Oxford University Press.
- Bolton, D. et Gillett, G. (2019). *The biopsychosocial model of health and disease: New philosophical and scientific developments*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-11899-0
- Boorse, C. (1975). On the distinction between disease and illness. *Philosophy & Public Affairs*, 5(1), 49-68. Récupéré de https://www.jstor.org/stable/2265020
- Boorse, C. (1976a). What a theory of mental health should be. *Journal for the Theory of Social Behaviour*, 6(1), 61-84. doi: 10.1111/j.1468-5914.1976.tb00359.x

- Boorse, C. (1976b). Wright on functions. *The Philosophical Review*, 85(1), 70-86. doi: 10.2307/2184255
- Boorse, C. (1977). Health as a theoretical concept. *Philosophy of Science*, 44(4), 542-573. doi: https://doi.org/10.1086/288768
- Boorse, C. (1997). A rebuttal on health. Dans J. M. Humber et R. F. Almeder (dir.), *What is disease?* (p. 1-134). Totowa, NJ: Humana Press.
- Boorse, C. (2002). A rebuttal on functions. Dans A. Ariew, R. Cummins et M. Perlman (dir.), *Functions: New essays in the philosophy of psychology and biology* (p. 63-112). Oxford, Royaume-Uni: Oxford University Press.
- Boorse, C. (2014). A second rebuttal on health. *Journal of Medicine and Philosophy*, 39(6), 683-724. doi: 10.1093/jmp/jhu035
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5-13. doi: 10.1002/wps.20375
- Bray, A. (2008). The extended mind and borderline personality disorder. *Australasian Psychiatry*, 16(1), 8-12.
- Bunge, M. (1967). *Philosophy of science: From explanation to justification* (réédition 1998, vol. 2). New Brunswick, NJ: Transaction Publishers.
- Bunge, M. (1997). Mechanism and explanation. *Philosophy of the Social Sciences*, 27(4), 410-465. doi: 10.1177/004839319702700402
- Bunge, M. (2003a). *Emergence and convergence: Qualitative novelty and the unity of knowledge*. Toronto, ON: University of Toronto Press.
- Bunge, M. (2003b). *Philosophical dictionary* (2^e éd.). Amherst, NY: Prometheus Books.
- Bunge, M. (2006). *Chasing reality: Strife over realism*. Toronto, ON: University of Toronto Press.
- Bunge, M. (2010). *Matter and mind: A philosophical inquiry*. New York, NY: Springer.
- Busner, J. et Targum, S. D. (2007). The clinical global impressions scale: Applying a research tool in clinical practice. *Psychiatry (Edgmont)*, 4(7), 28-37.

- Campbell, N. A. et Reece, J. B. (2012). *Biologie* (4^e éd.). Montréal, QC : Pearson ERPI.
- Carcone, D. et Ruocco, A. C. (2017). Six years of research on the National Institute of Mental Health's Research Domain Criteria (RDoC) initiative: A systematic review. *Frontiers in Cellular Neuroscience*, 11(46). doi: 10.3389/fncel.2017.00046
- Carel, H. (2013). *Illness: The cry of the flesh*. Durham, ON: Acumen Publishing Ltd.
- Carel, H. (2016). *The phenomenology of illness*. Oxford, Royaume-Uni : Oxford University Press.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford, Royaume-Uni: Clarendon Press.
- Cellules de Bénard. (2021, 14 janvier, 10 h 34). Dans *Wikipédia, l'encyclopédie libre*. Récupéré le 27 juin 2021 de https://fr.wikipedia.org/w/index.php?title=Cellules_de_B%C3%A9nard&oldid=178778775
- Centre de toxicomanie et de santé mentale. (2019). La crise est réelle. Dans *CAMH Agent du changement* | Récupéré de https://www.camh.ca/fr/agent-du-changement/la-crise-est-reelle
- Christensen, W. (2012). Natural sources of normativity. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(1), 104-112. doi: 10.1016/j.shpsc.2011.05.009
- Christensen, W. D., Bickhard, M. H. et The Hegeler Institute. (2002). The process dynamics of normative function. *The Monist*, 85(1), 3-28. doi: 10.5840/monist20028516
- Clark, A. (2008). Supersizing the mind: Embodiment, action, and cognitive extension. Oxford, Royaume-Uni: Oxford University Press.
- Clark, A. (2013). *Mindware: An introduction to the philosophy of cognitive science* (2^e éd.). Oxford, Royaume-Uni : Oxford University Press.
- Clark, A. et Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19. doi: 10.1111/1467-8284.00096

- Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E. et Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest*, 18(2), 72-145. doi: 10.1177/1529100617727266
- Cooper, R. (2002). Disease. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 33(2), 263-282. doi: 10.1016/S0039-3681(02)00018-3
- Cooper, R. (2007). *Psychiatry and philosophy of science*. Montréal, QC : McGill-Queen's University Press.
- Coughtrey, A. E., Shafran, R., Lee, M. et Rachman, S. (2013). The treatment of mental contamination: A case series. *Cognitive and Behavioral Practice*, 20(2), 221-231. doi: 10.1016/j.cbpra.2012.07.002
- Cournoyea, M. (2013). Ancestral assumptions and the clinical uncertainty of evolutionary medicine. *Perspectives in Biology and Medicine*, 56(1), 36-52. doi: 10.1353/pbm.2013.0006
- Craver, C. et Tabery, J. (2015). Mechanisms in science. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (hiver 2015). Récupéré de http://plato.stanford.edu/archives/win2015/entries/science-mechanisms/
- Craver, C. F. (2007). Explaining the brain: Mechanisms and the mosaic unity of neuroscience. Oxford, Royaume-Uni: Oxford University Press.
- Craver, C. F. et Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. Chicago, IL: University of Chicago Press.
- Cronbach, L. J. et Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302. doi: 10.1037/h0040957
- Cuthbert, B. (2014a). *RDoC kickoff webinar* [Vidéo en ligne]. Récupéré de https://www.youtube.com/watch?v=w3BGQwHVdd8
- Cuthbert, B. et Insel, T. R. (2010). The data of diagnosis: New approaches to psychiatric classification. *Psychiatry: Interpersonal and Biological Processes*, 73(4), 311-314. doi: 10.1521/psyc.2010.73.4.311

- Cuthbert, B. et Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC medicine*, 11(1), 1-8. doi: 10.1186/1741-7015-11-126
- Cuthbert, B. N. (2014b). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*, 13(1), 28-35. doi: 10.1002/wps.20087
- Cuthbert, B. N. (2015). Research Domain Criteria: Toward future psychiatric nosologies. *Dialogues in Clinical Neuroscience*, *17*(1), 89-97. doi: 10.31887/DCNS.2015.17.1/bcuthbert
- Cuthbert, B. N. et Kozak, M. J. (2013). Constructing constructs for psychopathology: The NIMH Research Domain Criteria. *Journal of Abnormal Psychology*, 122(3), 928-937. doi: 10.1037/a0034572
- Cuthbert, B. N. et Morris, S. E. (2012). Research Domain Criteria: Cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience*, 14(1), 29-37.
- Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York, NY: Putnam.
- Danziger, S., Levav, J. et Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17), 6889-6892. doi: 10.1073/pnas.1018033108
- Davis, D. (2003, 26 octobre). Losing the mind. Dans *Los Angeles Times*. Récupéré de http://articles.latimes.com/2003/oct/26/magazine/tm-survivors43
- de Haan, S. (2020a). An enactive approach to psychiatry. *Philosophy, Psychiatry, & Psychology, 27*(1), 3-25. doi: 10.1353/ppp.2020.0001
- de Haan, S. (2020b). *Enactive psychiatry*. Cambridge, Royaume-Uni : Cambridge University Press.
- De Rosnay, J. (1975). *Le macroscope: vers une version globale*. Paris : Éditions du Seuil.
- Deacon, T. et Koutroufinis, S. (2014). Complexity and dynamical depth. *Information*, 5(3), 404-423. doi: 10.3390/info5030404

- Deacon, T. W. (2003). The hierarchic logic of emergence: Untangling the interdependence of evolution and self-organization. Dans B. H. Weber et D. J. Depew (dir.), *Evolution and learning: The Baldwin effect reconsidered* (p. 273-308). Cambridge, Royaume-Uni: Cambridge University Press.
- Demazeux, S. et Pidoux, V. (2015). Le projet RDoC : la classification psychiatrique de demain? *Médecine/sciences*, *31*(8-9), 792-796. doi: 10.1051/medsci/20153108019
- Dennett, D. C. (1991). Consciousness explained. Boston, MA: Back Bay Books.
- Dennett, D. C. (2001). *The fantasy of first-person science* [Document non publié]. Tufts University. Récupéré de https://ase.tufts.edu/cogstud/dennett/papers/chalmersdeb3dft.htm
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology* and the Cognitive Sciences, 4(4), 429-452. doi: 10.1007/s11097-005-9002-y
- Di Paolo, E. A. (2009). Extended life. *Topoi*, 28(1), 9-21. doi: 10.1007/s11245-008-9042-3
- Di Paolo, E. A. et Thompson, E. (2014). The enactive approach. Dans L. Shapiro (dir.), *The Routledge handbook of embodied cognition* (1^{re} éd., p. 69-78). New York, NY: Routledge.
- Diabète Québec. (2021). Les symptômes. Dans *Diabète Québec*. Récupéré de https://www.diabete.qc.ca/fr/comprendre-le-diabete/tout-sur-le-diabete/symptomes/les-symptomes/
- Dinan, T. G., Stilling, R. M., Stanton, C. et Cryan, J. F. (2015). Collective unconscious: How gut microbes shape human behavior. *Journal of Psychiatric Research*, 63, 1-9. doi: 10.1016/j.jpsychires.2015.02.021
- Disner, S. G., Beevers, C. G., Haigh, E. A. P. et Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467-477. doi: 10.1038/nrn3027
- Doerr-Zegers, O. (2018). Space and time in the obsessive-compulsive phenomenon. *Psychopathology*, *51*(1), 31-37. doi: 10.1159/000485630
- Dretske, F. I. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

- Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition*, 30(3), 478-487. doi: 10.3758/BF03194948
- Enders, G. (2015). Le charme discret de l'intestin : tout sur un organe mal aimé... Arles : Actes Sud.
- Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 196(4286), 129-136. doi: 10.1126/science.847460
- Engel, G. L. (1980). The clinical application of the biopsychosocial model. *American Journal of Psychiatry*, 137(5), 535-544. doi: 10.1176/ajp.137.5.535
- Eronen, M. I. (2015). Robustness and reality. *Synthese*, *192*(12), 3961-3977. doi: 10.1007/s11229-015-0801-6
- Eronen, M. I. (2019a). Psychopathology and truth: A defense of realism. *The Journal of Medicine & Philosophy*, 44(4), 507-520. doi: 10.1093/jmp/jhz009
- Eronen, M. I. (2019b). Robust realism for the life sciences. *Synthese*, *196*(6), 2341-2354. doi: 10.1007/s11229-017-1542-5
- Escobar, J. M. (2012). Autopoiesis and darwinism. *Synthese*, *185*(1), 53-72. doi: 10.1007/s11229-011-9875-y
- Etkin, A. (2014a). Brain mapping study. Dans *Stanford medicine/Etkin Lab*. Récupéré de http://www.etkinlab.stanford.edu/brain-mapping-study
- Etkin, A. (2014b). Mapping and manipulating circuits for emotion and cognition in anxiety and depression. Dans *Research Portfolio Online Reporting Tools* (*RePORTER*). Récupéré de https://projectreporter.nih.gov/project_info_description.cfm?aid=9109051&icde=30709495
- Faraone, S. V., Seidman, L. J., Buka, S., Goldstein, J. M., Lyons, M., Kremen, W. S. et Glatt, S. J. (2013). Festschrift celebrating the career of Ming T. Tsuang. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 162(7), 551-558. doi: 10.1002/ajmg.b.32194
- Faucher, L. (2012). Evolutionary psychiatry and nosology: Prospects and limitations. Baltic International Yearbook of Cognition, Logic and Communication, 7(1), 1-64. doi: 10.4148/biyclc.v7i0.1776

- Faucher, L. (2021). Facts, facts, facts: HD analysis goes factual. Dans L. Faucher et D. Forest (dir.), *Defining mental disorder: Jerome Wakefield and his critics* (p. 47-70). Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9949.001.0001
- Faucher, L. et Goyer, S. (2015). RDoC: Thinking outside the DSM box without falling into a reductionist trap. Dans S. Demazeux et P. Singy (dir.), *The DSM-5 in perspective* (p. 199-224). New York, NY: Springer. doi: 10.1007/978-94-017-9765-8 12
- Faucher, L. et Goyer, S. (2016). Le Research Domain Criteria (RDoC), le réductionnisme et la psychiatrie clinique. *Revue de Synthèse*, *137*(1), 117-149. doi: 10.1007/s11873-016-0292-8
- Faucher, L. et Goyer, S. (2017). RDoC's special kind of reductionism and its possible impact on clinical psychiatry. Dans *The Routledge handbook of neuroethics* (p. 412-428). New York, NY: Routledge.
- Fodor, J. A. (1975). *The language of thought*. New York, NY: Thomas Y. Crowell.
- Frigg, R. et Hartmann, S. (2020). Models in science. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (printemps 2020). Récupéré de https://plato.stanford.edu/archives/spr2020/entries/models-science/
- Froese, T. (2017). Life is precious because it is precarious: individuality, mortality, and the problem of meaning. Dans G. Dodig-Crnkovic et R. Giovagnoli (dir.), *Representation and reality in humans, animal and machines* (p. 33-50). New York, NY: Springer.
- Fuchs, T. (2012). Are mental illnesses diseases of the brain? Dans S. Choudhury et J. Slaby (dir.), *Critical neuroscience: A handbook of the social and cultural contexts of neuroscience* (1^{re} éd., p. 331-344). Oxford, Royaume-Uni: Blackwell Publishing. doi: 10.1002/9781444343359.ch16
- Fuchs, T. (2018). *Ecology of the brain: The phenomenology and biology of the embodied mind*. Oxford, Royaume-Uni: Oxford University Press.
- Galatzer-Levy, I. R. et Bryant, R. A. (2013). 636 120 ways to have posttraumatic stress disorder. *Perspectives on Psychological Science*, 8(6), 651-662. doi: 10.1177/1745691613504115
- Gallagher, S. (2011). The overextended mind. *Versus: Quaderni di studi semiotici*, 113, 57-68. Récupéré de https://scholars.uow.edu.au/display/publication94104

- Gallagher, S. (2013). The socially extended mind. *Cognitive Systems research*, 25-26, 4-12. doi: 10.1016/j.cogsys.2013.03.008
- Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford, Royaume-Uni: Oxford University Press.
- Gallagher, S. (2018). Decentering the brain: Embodied cognition and the critique of neurocentrism and narrow-minded philosophy of mind. *Constructivist Foundations*, 14(1), 8-21. Récupéré de http://constructivist.info/14/1/008
- Gallagher, S. et Crisafi, A. (2009). Mental institutions. *Topoi*, 28(1), 45-51. doi: 10.1007/s11245-008-9045-0
- Gallagher, S. et Miyahara, K. (2012). Neo-pragmatism and enactive intentionality. Dans J. Schulkin (dir.), *Action, perception and the brain* (p. 117-146). Londres: Palgrave Macmillan. doi: 10.1057/9780230360792 6
- Gallagher, S. et Zahavi, D. (2012). *The phenomenological mind* (2° éd.). New York, NY: Routledge.
- Gallagher, S., Mastrogiorgio, A. et Petracca, E. (2019). Economic reasoning and interaction in socially extended market institutions. *Frontiers in Psychology*, 10, 1856. doi: 10.3389/fpsyg.2019.01856
- Gao, J., Wang, H., Liu, Y., Li, Y., Chen, C., Liu, L., Wu, Y., Li, S. et Yang, C. (2014). Glutamate and GABA imbalance promotes neuronal apoptosis in hippocampus after stress. *Medical Science Monitor*, 20, 499-512. https://doi.org/10.12659/MSM.890589
- Garson, J. (2013). The functional sense of mechanism. *Philosophy of Science*, 80(3), 317-333. doi: 10.1086/671173
- Garson, J. (2015). *The biological mind: a philosophical introduction*. New York, NY: Routledge.
- Garson, J. (2016). A critical overview of biological functions. New York, NY: Springer.
- Garson, J. (2017a). Mechanisms, phenomena, and functions. Dans *The Routledge handbook of mechanisms and mechanical philosophy* (1^{re} éd., p. 104-115). New York, NY: Routledge.

- Garson, J. (2017b). Against organizational functions. *Philosophy of Science*, 84(5), 1093-1103. https://doi.org/10.1086/694009
- Garson, J. (2019a). There are no ahistorical theories of function. *Philosophy of Science*, 86(5), 1146-1156. doi: 10.1086/705472
- Garson, J. (2019b). What biological functions are and why they matter. Cambridge, Royaume-Uni: Cambridge University Press. doi: 10.1017/9781108560764
- Geoffrion, P. (2009). Le groupe de discussion. Dans B. Gauthier (dir.), *Recherche sociale : de la problématique à la collecte des données* (5^e éd., p. 391-414). Québec, QC : Les Presses de l'Université du Québec.
- Giroux, E. (2012). Présentation du texte de Christopher Boorse : le concept théorique de santé. Dans É. Giroux et M. Lemoine (dir.), *Philosophie de la médecine : santé, maladie, pathologie* (vol. 2, p. 53-59). Paris : Vrin.
- Giroux, E. et Lemoine, M. (dir.). (2012). *Philosophie de la médecine : santé, maladie, pathologie*. Paris : Vrin.
- Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, 44(1), 49-71. doi: 10.1007/BF00172853
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69(S3), 342-353. doi: 10.1086/341857
- Glennan, S. (2017). *The new mechanical philosophy*. Oxford, Royaume-Uni : Oxford University Press.
- Godfrey-Smith, P. (2001). Three kinds of adaptationism. Dans H. Orzack et E. Sober (dir.), *Adaptationism and optimality* (p. 335-357). Cambridge, Royaume-Uni: Cambridge University Press.
- Gottesman, I. I. et Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry*, 160(4), 636-645. doi: 10.1176/appi.ajp.160.4.636
- Gouvernement du Québec. Centre d'étude sur la pauvreté et l'exclusion. (2019). Les troubles de la santé mentale et la pauvreté au Québec. Rédigé par A. Lechaume et F. Savard. Québec: Ministère du Travail, de l'Emploi et de la Solidarité sociale du Québec. Récupéré de https://www.mtess.gouv.qc.ca/publications/pdf/CEPE_sante-mentale-pauvrete_2019.pdf

- Goyer, S. (2013). *Pour un modèle de l'explication pluraliste et mécaniste en psychiatrie* (Mémoire de maîtrise). Université du Québec à Montréal. Récupéré de https://archipel.uqam.ca/5449/
- Graham, G. (2013). The disordered mind: An introduction to philosophy of mind and mental illness (2° éd.). New York, NY: Routledge.
- Gupta, M., Potter, N. N. et Goyer, S. (2019). Diagnostic reasoning in psychiatry: Acknowledging an explicit role for intersubjective knowing. *Philosophy, Psychiatry, & Psychology*, 26(1), 49-64. doi: 10.1353/ppp.2019.0003
- Hacking, I. (1983). Representing and intervening: Introductory topics in the philosophy of natural science. Cambridge, Royaume-Uni: Cambridge University Press.
- Harkness, A. R., Reynolds, S. M. et Lilienfeld, S. O. (2014). A review of systems for psychology and psychiatry: Adaptive systems, personality psychopathology five (PSY-5), and the DSM-5. *Journal of Personality Assessment*, *96*(2), 121-139. doi: 10.1080/00223891.2013.823438
- Harris, C. B., Barnier, A. J., Sutton, J. et Keil, P. G. (2014). Couples as socially distributed cognitive systems: Remembering in everyday social and material contexts. *Memory Studies*, 7(3), 285-297. doi: 10.1177/1750698014530619
- Hausman, D. M. (2011). Is an overdose of paracetamol bad for one's health? *The British Journal for the Philosophy of Science*, 62(3), 657-668. doi: 10.1093/bjps/axr008
- Hempel, C. G. et Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2), 135-175. doi: 10.1086/286983
- Higgins, E. S. (2009). Do ADHD drugs take a toll on the brain? *Scientific American Mind*, 20(4), 38-43. doi: 10.1038/scientificamericanmind0709-38
- Hoffman, G. A. et Zachar, P. (2017). RDoC's metaphysical assumptions: Problems and promises. Dans J. Poland et T. Şerife (dir.), *Extraordinary science and psychiatry: Responses to the crisis in mental health* (p. 59-86). Cambridge, MA: MIT Press.
- Horwitz, A. V. et Wakefield, J. C. (2010). *Tristesse ou dépression? Comment la psychiatrie a médicalisé nos tristesses*. Bruxelles, Belgique : Mardaga.

- Hsiang, S. M., Burke, M. et Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, *341*(6151), 1235367. doi: 10.1126/science.1235367
- Humphreys, P. (2016). *Emergence: A philosophical account*. Oxford, Royaume-Uni: Oxford University Press.
- Hurley, S. L. (2001). Perception and action: Alternative views. *Synthese*, *129*(1), 3-40, doi: 10.1023/A:1012643006930
- Husserl, E. (1990 [1928]). On the phenomenology of the consciousness of internal time. Dordrecht: Kluwer.
- Hutto, D. D. et Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge, MA: MIT Press.
- Hutto, D. D. et Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. Cambridge, MA: MIT Press.
- Hyman, S. E. (2007). Can neuroscience be integrated into the DSM-V? *Nature Reviews Neuroscience*, 8(9), 725-732. doi: 10.1038/nrn2218
- Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology*, *6*(1), 155-179. doi: 10.1146/annurev.clinpsy.3.022806.091532
- Ingrand, P. (2017). Loi normale de Laplace-Gauss. *Journal d'imagerie diagnostique et interventionnelle*, *1*, S4-S8. doi: 10.1016/S2543-3431(18)30066-6
- Insel, T. R. (2010). Faulty circuits. Scientific American, 302(4), 44-51.
- Insel, T. R. (2011, 22 décembre). Treatment development: Where do we go from here? *National Institute of Mental Health Director's blog*. Récupéré de https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2011/treatment-development-where-do-we-go-from-here.shtml
- Insel, T. R. (2013a). *Thomas Insel: Toward a new understanding of mental illness* [Vidéo en ligne]. Récupéré de https://www.youtube.com/watch?v=PeZ-U0pj9LI
- Insel, T. R. (2013b, 29 avril). Transforming diagnosis. *National Institute of Mental Health Director's Blog*. Récupéré de

- http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml et http://psychrights.org/2013/130429NIMHTransformingDiagnosis.htm
- Insel, T. R. (2014). The NIMH Research Domain Criteria (RDoC) project: Precision medicine for psychiatry. *American Journal of Psychiatry*, 171(4), 395-397. doi: 10.1176/appi.ajp.2014.14020138
- Insel, T. R. (2018). *Tom Insel, MD on better therapeutics for mental illness* [Vidéo en ligne]. Récupéré de https://www.youtube.com/watch?v=6isuhf2rOJI
- Insel, T. R. et Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science*, *348*(6234), 499-500. doi: 10.1126/science.aab2358
- Insel, T. R. et Sahakian, B. J. (2012). Drug research: A plan for mental illness. *Nature*, *483*(7389), 269-269. doi: 10.1038/483269a
- Insel, T. R., Collins, P. Y. et Hyman, S. E. (2015). Darkness invisible: The hidden global costs of mental illness. *Foreign Affairs*, 94(1), 127-135. Récupéré de https://www.researchgate.net/publication/279036526 Darkness Invisible The Hidden Global Costs of Mental Illness
- International OCD Foundation. (2021). Medications for OCD. Dans *International OCD Foundation*. Récupéré de https://iocdf.org/about-ocd/ocd-treatment/meds/
- Jacob, F. (1981). Le jeu des possibles : essai sur la diversité du vivant. Paris : Fayard.
- Jacob, P. (2019). Intentionality. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (hiver 2019). Récupéré de https://plato.stanford.edu/archives/win2019/entries/intentionality/
- Jasanoff, A. (2018). The biological mind: How brain, body, and environment collaborate to make us who we are. New York, NY: Basic Books.
- Jodoin, L. (2010). L'héritage intellectuel de Mario Bunge : entre science et philosophie. *Philosophiques*, *37*(2), 439-455. doi: 10.7202/045191ar
- Jonas, H. (1973). Organismus und freiheit: Ansätze zu einer philosophischen biologie. Göttingen: Vandenhoeck & Ruprecht.
- Joshi, S., Mooney, S. J., Rundle, A. G., Quinn, J. W., Beard, J. R. et Cerdá, M. (2017). Pathways from neighborhood poverty to depression among older adults. *Health & Place*, 43, 138-143. doi: 10.1016/j.healthplace.2016.12.003

- Junker, B. H. (1960). Field work: An introduction to the social sciences. Chicago, IL: University of Chicago Press.
- Kelly, A. B., Beach, S. R. H. et Fincham, F. D. (2003). Communication skills in couples: A review and discussion of emerging perspectives. Dans J. O. Greene et B. R. Burleson (dir.), *Handbook of communication and social interaction skills* (p. 723-751). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kendell, R. E. (1975). The concept of disease and its implications for psychiatry. *The British Journal of Psychiatry*, 127(4), 305-315. doi: 10.1192/bjp.127.4.305
- Kendler, K. S. (2005). Toward a philosophical structure for psychiatry. *American Journal of Psychiatry*, 162(3), 433-440. doi: 10.1176/appi.ajp.162.3.433
- Keyes, C., Nolte, L. et Williams, T. I. (2018). The battle of living with obsessive compulsive disorder: A qualitative study of young people's experiences. *Child and Adolescent Mental Health*, 23(3), 177-184. doi: 10.1111/camh.12216
- Kim, J. (1993). The non-reductivist's troubles with mental causation. Dans J. Heil et A. Mele (dir.), *Mental Causation* (p. 189-210). Oxford, Royaume-Uni: Clarendon Press.
- Kim, J. (1999). Making sense of emergence. *Philosophical Studies*, 95(1-2), 3-36. doi: 10.1023/a:1004563122154
- Kim, J. (2006). Emergence: Core ideas and issues. *Synthese*, *151*(3), 547-559. doi: 10.1007/s11229-006-9025-0
- Kingma, E. (2007). What is it to be healthy? *Analysis*, *67*(274), 128-133. doi: 10.1111/j.1467-8284.2007.00662.x
- Kingma, E. (2017). Disease as scientific and as value-laden concept. Dans T. Schramme et S. Edwards (dir.), *Handbook of the philosophy of medicine* (p. 45-63). New York, NY: Springer.
- Kingma. (2010). Paracetamol, poison, and polio: Why Boorse's account of function fails to distinguish health and disease. *The British Journal for the Philosophy of Science*, 61(2), 241-264. doi: 10.1093/bjps/axp034
- Kirk, S. A., Gomory, T. et Cohen, D. (2015). *Mad Science: Psychiatric coercision, diagnosis, and drugs*. Londres: Transaction Publishers.

- Kitzinger, J. (1995). Qualitative research: Introducing focus groups. *British Medical Journal*, 311(7000), 299-302. doi: 10.1136/bmj.311.7000.299
- Klee, P. (1997). *Introduction to the philosophy of science: Cutting nature at its seams*. Oxford, Royaume-Uni: Oxford University Press.
- Koch, H. (2020). Social dysfunction. Dans R. Hufendie, D. James et R. van Riel (dir.), *Social functions in philosophy: Metaphysical, normative, and methodological perspectives* (1^{re} éd., p. 45-69). New York, NY: Routledge.
- Kohn, L. et Christiaens, W. (2014). Les méthodes de recherches qualitatives dans la recherche en soins de santé : apports et croyances. *Reflets et perspectives de la vie économique*, *LIII*(4), 67-82. doi: 10.3917/rpve.534.0067
- Kozak, M. J. et Cuthbert, B. N. (2016). The NIMH Research Domain Criteria initiative: background, issues, and pragmatics. *Psychophysiology*, *53*(3), 286-297. doi: 10.1111/psyp.12518
- Kozak, M. J. et Miller, G. A. (1982). Hypothetical constructs versus intervening variables: A re-appraisal of the three-systems model of anxiety assessment. *Behavioral Assessment*, 4(3), 347-358. Récupéré de https://psycnet.apa.org/record/1983-00072-001
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, 84(3), 394-421. doi: 10.1016/j.biopsycho.2010.03.010
- Krystal, J. (2016). Interview with CINP president elect: Prof. John Krystal. *Progress In Mind*. Récupéré de http://progress.im/en/content/interview-cinp-president-elect-prof-john-krystal
- Lane, N. (2015). The vital question: Energy, evolution, and the origins of complex life. New York, London: W. W. Norton.
- Lebowitz, M. S. et Ahn, W. (2014). Effects of biological explanations for mental disorders on clinicians' empathy. *Proceedings of the National Academy of Sciences*, 111(50), 17786-17790. doi: 10.1073/pnas.1414058111
- Lende, D. H. (2012). Poverty poisons the brain. *Annals of Anthropological Practice*, 36(1), 183-201. doi: 10.1111/j.2153-9588.2012.01099.x
- Lewens, T. (2009). Seven types of adaptationism. *Biology & Philosophy*, 24(2), 161-182. doi: 10.1007/s10539-008-9145-7

- Lilienfeld, S. O. et Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology*, *12*(1), 435-463. doi: 10.1146/annurev-clinpsy-021815-093122
- MacCorquodale, K. et Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*(2), 95-107. doi: 10.1037/h0056029
- Machamer, P., Darden, L. et Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1-25. doi: 10.1086/392759
- Maiese, M. (2012). Rethinking attention deficit hyperactivity disorder. *Philosophical Psychology*, 25(6), 893-916. doi: 10.1080/09515089.2011.631998
- Maiese, M. (2016). *Embodied selves and divided minds*. Oxford, Royaume-Uni: Oxford University Press.
- Mandik, P. (2014). *This is philosophy of mind: An introduction*. Oxford, Royaume-Uni: Blackwell Publishing.
- Maturana, H. R. (1975). The organization of the living: A theory of the living organization. *International Journal of Man-Machine Studies*, 7(3), 313-332. doi: 10.1016/S0020-7373(75)80015-0
- Maturana, H. R. et Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht: D. Reidel Publishing Company.
- Maturana, H. R. et Varela, F. J. (1992). *The tree of knowledge: The biological roots of human understanding*. Boulder, CO: Shambhala.
- Maxwell, J. A. (2004). Using qualitative methods for causal explanation. *Field Methods*, 16(3), 243-264. doi: 10.1177/1525822X04266831
- Maxwell, J. A. (2012). The importance of qualitative research for causal explanation in education. *Qualitative Inquiry*, 18(8), 655-661. doi: 10.1177/1077800412452856
- McCauley, R. N. et Bechtel, W. (2001). Explanatory pluralism and the heuristic identity theory. *Theory and Psychology*, 11(6), 736-760. doi: 10.1177/0959354301116002
- McGuire, M. et Troisi, A. (1998). *Darwinian psychiatry*. Oxford, Royaume-Uni: Oxford University Press.

- McNally, R. J. (2011). *What is mental illness?* Cambridge, MA: Harvard University Press.
- McTeague, L. M., Goodkind, M. S. et Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of Psychiatric Research*, 83, 37-46. doi: 10.1016/j.jpsychires.2016.08.001
- Meehl, P. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality*, 60, 117-174. Récupéré de https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.555.6702&rep=rep1&type=pdf
- Meincke, A. S. (2019). The disappearance of change: Towards a process account of persistence. *International Journal of Philosophical Studies*, 27(1), 12-30. doi: 10.1080/09672559.2018.1548634
- Menary, R. (2013). Cognitive integration, enculturated cognition and the socially extended mind. *Cognitive Systems Research*, 25-26, 26-34. doi: 10.1016/j.cogsys.2013.05.002
- Merritt, M. (2013). Instituting impairment: Extended cognition and the construction of female sexual dysfunction. *Cognitive Systems Research*, 25-26, 47-53. doi: 10.1016/j.cogsys.2013.03.005
- Merton, R. K. (1968). *Social theory and social structure*. New York, NY: Simon and Schuster.
- Mialet, H. (2014). The pugilist and the cosmologist: Response to Loïc Wacquant's 'Homines in extremis'. Body & Society, 20(2), 91-99. doi: 10.1177/1357034X14524799
- Miller, G. A. et Kozac, M. J. (1993). Three-systems assessment and the construct of emotion. Dans N. Birbaumer et A. Öhman (dir.), *The structure of emotion: Physiological, cognitive and clinical aspects* (p. 31-47). Seattle, WA: Hogrefe & Huber.
- Millikan, R. G. (1989). Biosemantics. *The Journal of Philosophy*, 86(6), 281-297. doi: 10.2307/2027123
- Morgan, D. L. (2008). Focus Groups. Dans L. M. Given (dir.), *The SAGE* encyclopedia of qualitative research methods (1^{re} éd., vol. 1 et 2, p. 352-354). Thousand Oaks, CA: SAGE Publications.

- Morris, S. E., Vaidyanathan, U. et Cuthbert, B. N. (2016). Changing the diagnostic concept of schizophrenia: The NIMH Research Domain Criteria initiative. Dans M. Li et W. D. Spaulding (dir.), *The Neuropsychopathology of Schizophrenia* (vol. 63, p. 225-252). New York, NY: Springer. doi: 10.1007/978-3-319-30596-7
- Mossio, M. et Moreno, A. (2015). *Biological autonomy: A philosophical and theoretical enquiry*. New York, NY: Springer.
- Mossio, M., Saborido, C. et Moreno, A. (2009). An organizational account of biological functions. *The British Journal for the Philosophy of Science*, 60(4), 813-841. doi: 10.1093/bjps/axp036
- Mossio, M., Saborido, C. et Moreno, A. (2010). Fonctions : normativité, téléologie et organisation. Dans J. Gayon et A. Ricqlès (dir.), *Les fonctions : des organismes aux artéfacts* (p. 159-166). Paris : Presses Universitaires de France.
- Moyes, C. D. et Schulte, P. M. (2006). *Principles of animal physiology*. San Francisco, CA: Benjamin Cummings Publishers.
- Murphy, D. (2015). Validity, realism, and normativity. Dans P. Zachar, D. S. Stoyanov, M. Aragona et A. Jablensky (dir.), *Alternative perspectives on psychiatric validation: DSM, IDC, RDoC, and Beyond* (p. 60-75). Oxford, Royaume-Uni: Oxford University Press.
- Murphy, D. et Roe, K. (2011). Function, dysfunction, and adaptation? Dans A. D. Block et P. R. Adriaens (dir.), *Maladapting Minds: Philosophy, Psychiatry, and Evolutionary Theory* (p. 216-237). Oxford, Royaume-Uni: Oxford University Press.
- Nadelhoffer, T. et Sinnott-Armstrong, W. (2013). Is psychopathy a mental disease? Dans N. A. Vincent (dir.), *Neuroscience and Legal Responsibility* (p. 229-255). Oxford, Royaume-Uni: Oxford University Press.
- Nagel, E. (1961). The structure of science: Problems in the logic of scientific explanation. New York, NY: Harcourt, Brace & World.
- National Institute of Mental Health. (2008). *National Institute of Mental Health Strategic plan*. Récupéré de https://www.hsdl.org/?view&did=755067
- National Institute of Mental Health. (2015a). National Institute of Mental Health Strategic plan for research.

- National Institute of Mental Health. (2015b). *RDOC Negative Valence Webinar 5-11-15* [Vidéo en ligne]. Récupéré de https://www.youtube.com/watch?v=CQiQ0yiFQo0
- National Institute of Mental Health. (2017). *Home*. Récupéré de https://www.nimh.nih.gov/index.shtml
- National Institute of Mental Health. (2019a). *Introduction*. Récupéré de https://www.nimh.nih.gov/about/strategic-planning-reports/introduction.shtml
- National Institute of Mental Health. (2019b). *Research Domain Criteria (RDoC)*. Récupéré de https://www.nimh.nih.gov/research-priorities/rdoc/index.shtml
- National Institute of Mental Health. (2020a). *National Institute of Mental Health Strategic Plan for Research*. Récupéré de https://www.nimh.nih.gov/sites/default/files/documents/about/strategic-planning-reports/2020 nimh strategic plan 508.pdf
- National Institute of Mental Health. (2020b). *Update on Genes in the RDoC Matrix*. Récupéré de https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/update-on-genes-in-the-rdoc-matrix.shtml
- National Institute of Mental Health. (2021a). *About RDoC*. Récupéré de https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/about-rdoc
- National Institute of Mental Health. (2021b). *RDoC Matrix*. Récupéré de https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/constructs/rdoc-matrix
- Neveu, R. (2005). Faire face aux TOC. Paris: Retz.
- Newman, E., Jernigan, T. L., Lisdahl, K. M., Tamm, L., Tapert, S. F., Potkin, S. G., ... MTA Neuroimaging Group. (2016). Go/no go task performance predicts cortical thickness in the caudal inferior frontal gyrus in young adults with and without ADHD. *Brain Imaging and Behavior*, 10(3), 880-892. doi: 10.1007/s11682-015-9453-x
- Nicholson, D. J. et Dupré, J. (dir.). (2018). *Everything flows: Towards a processual philosophy of biology*. Oxford, Royaume-Uni: Oxford University Press.
- Nielsen, K. (2020). Think of mental disorders as the mind's 'sticky tendencies'. Dans *Aeon*. Récupéré de https://aeon.co/ideas/think-of-mental-disorders-as-the-minds-sticky-tendencies

- Nielsen, K. et Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism, and embedment. *Theory & Psychology*, 28(6), 800-822. doi: 10.1177/0959354318808394
- Nielsen, K. et Ward, T. (2020a). Mental disorder as both natural and normative: Developing the normative dimension of the 3° conceptual framework for psychopathology. *Journal of Theoretical and Philosophical Psychology*, 40(2), 107-123. doi: 10.1037/teo0000118
- Nielsen, K. et Ward, T. (2020b). Phenomena complexes as targets of explanation in psychopathology: The relational analysis of phenomena approach. *Theory & Psychology*, 30(2), 164-185. doi: 10.1177/0959354320906462
- Noë, A. (2004). Action in perception. Cambridge, MA: MIT Press.
- Noë, A. (2012). Varieties of presence. Cambridge, MA: Harvard University Press.
- Nummenmaa, L., Glerean, E., Hari, R. et Hietanen, J. K. (2014). Bodily maps of emotions. *Proceedings of the National Academy of Sciences*, 111(2), 646-651. doi: 10.1073/pnas.1321664111
- O'Regan, J. K. et Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), 939-973. doi: 10.1017/S0140525X01000115
- Okrent, M. (2017). *Nature and normativity: Biology, teleology, and meaning*. New York, NY: Routledge.
- Paris, J. et Kirmayer, L. J. (2016). The National Institute of Mental Health Research Domain Criteria: A bridge too far. *The Journal of Nervous and Mental Disease*, 204(1), 26-32. doi: 10.1097/NMD.0000000000000435
- Parnas, J. (2014). The RDoC program: Psychiatry without psyche? *World Psychiatry*, 13(1), 46-47. doi: 10.1002/wps.20101
- Patrick, C. J., Venables, N. C., Yancey, J. R., Hicks, B. M., Nelson, L. D. et Kramer, M. D. (2013). A construct-network approach to bridging diagnostic and physiological domains: Application to assessment of externalizing psychopathology. *Journal of Abnormal Psychology*, 122(3), 902-916. doi: 10.1037/a0032807

- Pendule simple. (2020, 13 décembre, 14 h 3). Dans *Wikipédia, l'encyclopédie libre*. Récupéré le 6 avril 2021 de https://fr.wikipedia.org/w/index.php?title=Pendule_simple&oldid=177586743
- Peretz, H. (2004). Les méthodes en sociologie : l'observation (2^e éd.). Paris : La Découverte.
- Petitmengin, C. (2005). Un exemple de recherche neuro-phénoménologique : l'anticipation des crises d'épilepsie. *Intellectica*, 40(1), 63-89. doi: 10.3406/intel.2005.1359
- Petitot, J., Varela, F. J., Pachoud, B. et Roy, J.-M. (dir.). (1999). *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science*. Palo Alto, CA: Stanford University Press.
- Petracca, E. et Gallagher, S. (2020). Economic cognitive institutions. *Journal of Institutional Economics*, 16(6), 1-19. doi: 10.1017/S1744137420000144
- Pettit, P. (1993). A Definition of physicalism. *Analysis*, *53*(4), 213-223. doi: 10.1093/analys/53.4.213
- Pigliucci, M. et Kaplan, J. (2006). *Making sense of Evolution: The conceptual foundations of evolutionary biology*. Chicago, IL: University of Chicago Press.
- Pitt, D. (2020, printemps). Mental representation. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy*. Récupéré de https://plato.stanford.edu/archives/spr2020/entries/mental-representation/
- Poirier, P., Faucher, L. et Bourdon, J.-N. (2019). Cultural blankets: Epistemological pluralism in the evolutionary epistemology of mechanisms. *Journal for General Philosophy of Science*, 52(49). doi: 10.1007/s10838-019-09472-8
- Polger, T. W. (2020). Functionalism. Dans *Internet Encyclopedia of Philosophy*. Récupéré de https://iep.utm.edu/functism/
- Pominville-Racette, M. (2018). Les impacts d'une démarche d'empowerment sur la pauvreté : étude de cas de l'Épicerie solidaire de Rosemont (Mémoire de maîtrise). Université du Québec à Montréal. Récupéré de https://archipel.uqam.ca/11559/1/M15500.pdf

- Porter, D. (2019). RDoC, psychopathology, and naturalism: What's new is what's old? *Journal of Humanistic Psychology*, *59*(1), 6-25. doi: 10.1177/0022167818778663
- Potter, N. N. (2013). Empathic foundations of clinical knowledge. Dans K. W. Fulford, M. Davies, R. G. T. Gipps, J. Z. Sadler, G. Stanghellini, T. Thornton et G. Graham (dir.), *The Oxford handbook of philosophy and psychiatry* (p. 293-306). Oxford, Royaume-Uni: Oxford University Press.
- Powell, R. et Scarffe, E. (2019a). Rehabilitating "disease": Function, value, and objectivity in medicine. *Philosophy of Science*, 86(5), 1168-1178. doi: 10.1086/705520
- Powell, R. et Scarffe, E. (2019b). Rethinking "Disease": a fresh diagnosis and a new philosophical treatment. *Journal of Medical Ethics*, *45*(9), 579-588. doi: 10.1136/medethics-2019-105465
- Price, J., Sloman, L., Gardner, R., Gilbert, P. et Rohde, P. (1994). The social competition hypothesis of depression. *British Journal of Psychiatry*, *164*(3), 309-315. doi: 10.1192/bjp.164.3.309
- Ratcliffe, M. (2015). *Experiences of depression: A study in phenomenology*. Oxford, Royaume-Uni: Oxford University Press.
- Rescorla, M. (2020). The computational Theory of Mind. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (printemps 2020). Récupéré de https://plato.stanford.edu/archives/spr2020/entries/computational-mind/
- Richardson, R. C. (2007). *Evolutionary psychology as maladapted psychology* (1^{re} éd.). Cambridge, MA: MIT Press.
- Rioux, C. (2014). L'approche organisationnelle des fonctions est-elle trop restrictive? *Phares*, *14*, 31-52. Récupéré de https://revuephares.com/wp-content/uploads/2014/05/Phares-XIV-03-Catherine-Rioux.pdf
- Rottenberg, J. (2014). *The depths: The evolutionary origins of the depression epidemic*. New York, NY: Basic Books.
- Rupert, R. D. (2004). Challenges to the hypothesis of extended cognition. *The Journal of Philosophy*, 101(8), 389-428. doi: 10.5840/jphil2004101826

- Saborido, C. et Moreno, A. (2015). Biological pathology from an organizational perspective. *Theoretical Medicine and Bioethics*, *36*(1), 83-95. doi: 10.1007/s11017-015-9318-8
- Saborido, C., Moreno, A., González-Moreno, M. et Clemente, J. (2016).

 Organizational malfunctions and the notions of health and disease. Dans É.

 Giroux (dir.), *Naturalism in the philosophy of health: Issues and implications* (p. 101-120). New York, NY: Springer.
- Salamé, P. et Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of verbal learning and verbal behavior*, 21(2), 150-164. doi: 10.1016/S0022-5371(82)90521-7
- Saldana, J. (2011). *Fundamentals of qualitative research*. Oxford, Royaume-Uni: Oxford University Press.
- Sanislow, C. A. (2016). Connecting psychopathology meta-structure and mechanisms. *Journal of Abnormal Psychology*, *125*(8), 1158-1165. doi: 10.1037/abn0000207
- Sanislow, C. A., Ferrante, M., Pacheco, J., Rudorfer, M. V. et Morris, S. E. (2019). Advancing translational research using NIMH Research Domain Criteria and computational methods. *Neuron*, *101*(5), 779-782. doi: 10.1016/j.neuron.2019.02.024
- Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinssen, R. K., ... Cuthbert, B. N. (2010). Developing constructs for psychopathology research: Research Domain Criteria. *Journal of Abnormal Psychology*, *119*(4), 631-639. doi: 10.1037/a0020909
- Savoie-Zajc, L. (2009). L'entrevue semi-dirigée. Dans B. Gauthier (dir.), *Recherche sociale : de la problématique à la collecte des données* (p. 337-360). Québec, QC : Les Presses de l'Université du Québec.
- Schaffer, J. (2003). Is there a fundamental level? *Nous*, *37*(3), 498-517. doi: 10.1111/1468-0068.00448
- Schaffner, K. F. (2006). Reduction: the Cheshire cat problem and a return to roots. *Synthese*, *151*(3), 377-402. doi: 10.1007/s11229-006-9031-2
- Schaffner, K. F. (2008). Etiological models in psychiatry: Reductive and nonreductive approaches. Dans K. S. Kendler et J. Parnas (dir.), *Philosophical*

- issues in psychiatry: Explanation, phenomenology, and nosology (p. 48-90). Baltimore, MD: Johns Hopkins University Press.
- Schaffner, K. F. (2011). Reduction in biology and medicine. Dans F. Gifford (dir.), *Philosophy of medicine* (vol. 16, p. 137-157). Amsterdam: Elsevier.
- Schaffner, K. F. (2013). Reduction and reductionism in psychiatry. Dans B. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini et T. Thornton (dir.), *The Oxford handbook of philosophy and psychiatry* (p. 1003-1022). Oxford, Royaume-Uni: Oxford University Press.
- Schaffner, K. F. (2016). *Behaving: What's genetic, what's not, and why should we care?* Oxford, Royaume-Uni: Oxford University Press.
- Scheff, T. J. (1966). Being mentally ill: A sociological theory. Chicago, IL: Aldine.
- Schrödinger, E. (1944). What is life? Cambridge: Cambridge University Press.
- Seibt, J. (2009). Forms of emergent interaction in General Process Theory. *Synthese*, *166*(3), 479-512. doi: 10.1007/s11229-008-9373-z
- Seibt, J. (2018). Ontological tools for the process turn in biology. Dans J. Dupré et D. J. Nicholson (dir.), *Everything flows: Towards a processual philosophy of biology* (vol. 1, p. 113-136). Oxford, Royaume-Uni: Oxford University Press. doi: 10.1093/oso/9780198779636.003.0006
- Shanks, N. (2004). God, the devil, and Darwin: A critique of intelligent design theory. Oxford, Royaume-Uni: Oxford University Press.
- Shapiro, L. (2019). *Embodied cognition* (2^e éd.). New York, NY: Routledge.
- Sharp, C., Kim, S., Herman, L., Pane, H., Reuter, T. et Strathearn, L. (2014). Major depression in mothers predicts reduced ventral striatum activation in adolescent female offspring with and without depression. *Journal of Abnormal Psychology*, 123(2), 298-309. doi: 10.1037/a0036191
- Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J. P., Greenstein, D., ... Rapoport, J. L. (2007). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences*, 104(49), 19649-19654. doi: 10.1073/pnas.0707741104
- Shepard, R. N. et Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701-703. doi: 10.1126/science.171.3972.701

- Simmons, J. M. et Quinn, K. J. (2014). The NIMH Research Domain Criteria (RDoC) project: Implications for genetics research. *Mammalian Genome*, 25(1-2), 23-31. doi: 10.1007/s00335-013-9476-9
- Sisti, D., Young, M. et Caplan, A. (2013). Defining mental illnesses: Can values and objectivity get along? *BMC psychiatry*, *13*(346), 1-4. doi: 10.1186/1471-244X-13-346
- Slaby, J. et Gallagher, S. (2015). Critical neuroscience and socially extended minds. *Theory, Culture & Society*, 32(1), 33-59. doi: 10.1177/0263276414551996
- Smart, J. J. C. (2017). The mind/brain identity theory. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (printemps 2017). Récupéré de https://plato.stanford.edu/archives/spr2017/entries/mind-identity/
- Smith, P. A. (2015, 23 juin). Can the bacteria in your gut explain your mood? *The New York Times Magazine*. Récupéré de https://www.nytimes.com/2015/06/28/magazine/can-the-bacteria-in-your-gut-explain-your-mood.html
- Spivey, M. J. (2020). *Who you are: The science of connectedness*. Cambridge, MA: The MIT Press.
- Stanghellini, G., Ballerini, M., Presenza, S., Mancini, M., Northoff, G. et Cutting, J. (2017). Abnormal time experiences in major depression: An empirical qualitative study. *Psychopathology*, 50(2), 125-140. doi: 10.1159/000452892
- Stegenga, J. (2018a). *Care and cure: an introduction to philosophy of medicine*. Chicago, IL: University of Chicago Press.
- Stegenga, J. (2018b). *Medical nihilism*. Oxford, Royaume-Uni : Oxford University Press.
- Stich, S. et Donaldson, T. (2018). *Philosophy: Asking questions seeking answers*. Oxford, Royaume-Uni: Oxford University Press.
- Stoljar, N. (2018). Feminist perspectives on autonomy. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (hiver 2018). Récupéré de https://plato.stanford.edu/archives/win2018/entries/feminism-autonomy/
- Stoyanov, D., Correia, D. T. et Cuthbert, B. N. (2019). The Research Domain Criteria (RDoC) and the historical roots of psychopathology: A viewpoint. *European Psychiatry*, *57*, 58-60. doi: 10.1016/j.eurpsy.2018.11.007

- Sukhodolsky, D. (2014). Using CBT to examine circuitry of frustrative non-reward in aggressive children. Dans *Research Portfolio Online Reporting Tools* (*RePORTER*). Récupéré de https://projectreporter.nih.gov/project_info_description.cfm?aid=9116671&icde=30725483&ddparam=&ddvalue=&ddsub=&cr=5&csb=default&cs=ASC
- Sukhodolsky, D., Wyk, B. C. V., Eilbott, J. A., McCauley, S. A., Ibrahim, K., Crowley, M. J. et Pelphrey, K. A. (2016). Neural mechanisms of cognitive-behavioral therapy for aggression in children and adolescents: Design of a randomized controlled trial within the National Institute for Mental Health Research Domain Criteria construct of frustrative non-reward. *Journal of Child and Adolescent Psychopharmacology*, 26(1), 38-48. doi: 10.1089/cap.2015.0164
- Sullivan, J. A. (2016). Stabilizing constructs through collaboration across different research fields as a way to foster the integrative approach of the Research Domain Criteria (RDoC) project. *Frontiers in Human Neuroscience*, 10, 1-4. doi: 10.3389/fnhum.2016.00309
- Szasz, T. (1961). The myth of mental illness: Foundations of a theory of personal conduct. New York, NY: Hoeber-Harper.
- Tal, E. (2017). Measurement in science. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (automne 2017). Récupéré de https://plato.stanford.edu/archives/fall2017/entries/measurement-science/
- Thagard, P. (1999). *How scientists explain disease*. Princeton, NJ: Princeton University Press.
- Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind* (1^{re} éd.). Cambridge, MA: Harvard University Press.
- Thompson, E. (2017, 27 janvier). The enactive approach. Dans *The Brains Blog*. Récupéré de http://philosophyofbrains.com/2017/01/27/the-enactive-approach.aspx
- Thompson, E. (2018). Review of evolving enactivism: Basic minds meet content. Notre Dame Philosophical Reviews. Récupéré de https://ndpr.nd.edu/news/evolving-enactivism-basic-minds-meet-content/
- Thompson, E. et Lutz, A. (2003). Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10(9-10), 31-52.

- Thompson, E. et Varela, F. J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in cognitive sciences*, *5*(10), 418-425. doi: doi.org/10.1016/S1364-6613(00)01750-2
- Thompson, P. et Upshur, R. E. G. (2017). *Philosophy of medicine: An introduction*. New York, NY: Routledge.
- Tsuang, M. T., Stone, W. S., Gamma, F. et Faraone, S. V. (2003). Schizotaxia: Current status and future directions. *Current Psychiatry Reports*, *5*(2), 128-134. doi: 10.1007/s11920-003-0029-x
- Vaidyanathan, U. (2016). Webinar: Analyzing and using RDoC Data in Your Research [Vidéo en ligne]. Récupéré de https://www.youtube.com/watch?v=i RV0oMFdGc
- Valles, S. A. (2012). Evolutionary medicine at twenty: Rethinking adaptationism and disease. *Biology & Philosophy*, 27(2), 241-261. doi: 10.1007/s10539-011-9305-z
- Valles, S. A. (2020). Philosophy of Biomedicine. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (été 2020). Récupéré de https://plato.stanford.edu/archives/sum2020/entries/biomedicine/
- Van Gulick, R. (2011). Life, holism and emergence: Converging themes. *Journal of Consciousness Studies*, 18(5-6), 139-147.
- Varela, F. J. (1989). *Autonomie et connaissance essai sur le vivant*. Paris : Éditions du Seuil.
- Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, *3*(4), 330-349. Récupéré de https://unstable.nl/andreas/ai/langcog/part3/varela npmrhp.pdf
- Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34(1), 72-87. doi: 10.1006/brcg.1997.0907
- Varela, F. J. (2004). *Quel savoir pour l'éthique? Action, sagesse et cognition*. Paris : La Découverte.
- Varela, F. J. (2017). Le cercle créateur : écrits (1976-2001). Paris : Éditions du Seuil.
- Varela, F. J., Evan, T. et Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

- Varela, F. J., Thompson, E. et Rosch, E. (1993). L'inscription corporelle de l'esprit : sciences cognitives et expérience humaine. Paris : Éditions du Seuil.
- Varga, S. (2013). The frames of cognition. *Cognitive Systems Research*, 25-26, 54-60. doi: 10.1016/j.cogsys.2013.03.004
- Varga, S. (2018). Demarcating the realm of cognition. *Journal for General Philosophy of Science*, 49(3), 435-450. doi: 10.1007/s10838-017-9375-y
- Volynets, S., Glerean, E., Hietanen, J. K., Hari, R. et Nummenmaa, L. (2019). Bodily maps of emotions are culturally universal. *Emotion*, 20(7), 1127-1136. doi: 10.1037/emo0000624
- Vrij, A., Van der Steen, J. et Koppelaar, L. (1994). Aggression of police officers as a function of temperature: An experiment with the fire arms training system. *Journal of community & applied social psychology*, 4(5), 365-370. doi: 10.1002/casp.2450040505
- Waddington, C. H. (1969). The practical consequences of metaphysical beliefs on a biologist's work: An autobiographical note. Dans C. H. Waddington (dir.), *Towards a theoretical biology: Sketches* (Vol. 2, p. 72-81). Édimbourg: Edinburgh University Press.
- Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist*, 47(3), 373-388. doi: 10.1037/0003-066X.47.3.373
- Wakefield, J. C. (2006a). Fait et valeur dans le concept de trouble mental: le trouble en tant que dysfonction préjudiciable. *Philosophiques*, *33*(1), 37-63. doi: 10.7202/012946ar
- Wakefield, J. C. (2006b). What makes a mental disorder mental? *Philosophy, Psychiatry, & Psychology*, 13(2), 123-131. doi: 10.1353/ppp.2007.0010
- Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, 6(3), 149-156.
- Wakefield, J. C. (2011). Darwin, functional explanation, and the philosophy of psychiatry. Dans A. D. Block et P. R. Adriaens (dir.), *Maladapting minds: Philosophy, psychiatry, and evolutionary theory* (p. 143-172). Oxford, Royaume-Uni: Oxford University Press.

- Wakefield, J. C. et Conrad, J. A. (2019). Does the harm component of the harmful dysfunction analysis need rethinking? Reply to Powell and Scarffe. *Journal of Medical Ethics*, 45(9), 594-596. doi: 10.1136/medethics-2019-105578
- Ward, D., Silverman, D. et Villalobos, M. (2017). Introduction: The varieties of enactivism. *Topoi*, *36*(3), 365-375. doi: 10.1007/s11245-017-9484-6
- Weber, A. et Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the cognitive sciences*, *I*(2), 97-125. doi: 10.1023/A:1020368120174
- Weiskopf, D. et Adams, F. (2015). *An introduction to the philosophy of psychology*. Cambridge, Royaume-Uni: Cambridge University Press.
- Welshon, R. (2011). *Philosophy, neuroscience, and consciousness*. Montréal, QC : Mcgill-Queens University Press.
- Wheeler, M. (2013). Science friction: Phenomenology, naturalism and cognitive science. Dans H. Carel et D. Meacham (dir.), *Human experience and nature: Examining the relationship between phenomenology and naturalism* (p. 135-168). Cambridge, Royaume-Uni: Cambridge University Press.
- Wimsatt, W. (1981). Robustness, reliability, and overdetermination. Dans M. B. Brewer et B. E. Collins (dir.), *Scientific inquiry and the social sciences: A volume in honor of Donald T. Campbell* (p. 124-161). San Francisco, CA: Jossey-Bass.
- Wimsatt, W. (1994). The ontology of complex systems: Levels of organization, perspectives, and causal thickets. *Canadian Journal of Philosophy Supplementary Volume*, 20, 207-274. doi: 10.1080/00455091.1994.10717400
- Woodruff Smith, D. (2018). Phenomenology. Dans E. N. Zalta (dir.), *The Stanford encyclopedia of philosophy* (été 2018). Récupéré de https://plato.stanford.edu/archives/sum2018/entries/phenomenology/
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford, Royaume-Uni: Oxford University Press.
- Wright, C. et Bechtel, W. (2007). Mechanisms and psychological explanation. Dans D. M. Gabbay, J. Woods et P. Thagard (dir.), *Philosophy of psychology and cognitive science: A volume of the handbook of the philosophy of science series* (p. 31-79). Amsterdam: Elsevier. doi: 10.1016/B978-0-444-51540-7.X5015-4

- Zachar, P. (2012). Progress and the calibration of scientific constructs: The role of comparative validity. Dans K. S. Kendler et J. Parnas (dir.), *Philosophical issues in psychiatry II: Nosology* (1^{re} éd., p. 21-34). Oxford, Royaume-Uni: Oxford University Press. doi: 10.1093/med/9780199642205.001.0001
- Zachar, P. (2014a). A metaphysics of psychopathology. Cambridge, MA: MIT Press.
- Zachar, P. (2014b). Beyond natural kinds: Toward a "relevant" "scientific" taxonomy in psychiatry. Dans J. A. Sullivan et H. Kincaid (dir.), *Classifying psychopathology: Mental kinds and natural kinds* (p. 75-104). Cambridge, MA: MIT Press.
- Zachar, P. et Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *The American Journal of Psychiatry*, *164*(4), 557-565. doi: 0.1176/appi.ajp.164.4.557