

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

POUR UN MODÈLE DE L'EXPLICATION PLURALISTE ET MÉCANISTE EN
PSYCHIATRIE

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAITRISE EN PHILOSOPHIE

PAR

SIMON GOYER

MAI 2013

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). 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.»

Don't even visit that place they'll sharpen their teeth on your smile. I'm glad you didn't, all our songs will be lullabies in no time.

Emily Haines, *Knives Don't Have Your Back*

REMERCIEMENTS

Je tiens particulièrement à témoigner ma reconnaissance à Luc Faucher, mon directeur de maîtrise et professeur de philosophie à l'UQAM, pour son soutien continu, pour la bourse d'études qu'il m'a donnée de même que pour la liberté qu'il me laissa lors de la rédaction de ce mémoire. J'ai gagné en autonomie intellectuelle sous sa direction et lui en suis reconnaissant. Je remercie également Pierre Poirier, Christophe Malaterre et Serge Robert, tous professeurs de philosophie à l'UQAM et membre du jury, pour leurs critiques constructives. Je suis aussi redevable envers Vincent Guillin, lui aussi professeur de philosophie à l'UQAM, qui fut présent lors du premier séminaire et dont deux suggestions — l'une concernant une lecture et l'autre relative aux présentations orales — me furent fort utiles. En outre, j'exprime ma gratitude à Paul Daoust qui a bien voulu lire mon manuscrit afin de voir s'il restait des erreurs de français (je demeure le seul responsable de celles qui resteraient, évidemment). Je remercie, de plus, le *Centre interuniversitaire de recherche sur la science et la technologie* (CIRST) pour la bourse d'études dont j'ai pu bénéficier. Enfin, je dis merci du fond du cœur à tous ceux qui étaient là dans ma vie quotidienne lors de la rédaction de ce mémoire. Merci spécialement à mon père Normand et à ma mère Suzanne pour leur soutien émotif et financier indéfectibles. Merci à mon frère Matthieu pour sa présence et son amitié. Merci à Marie-Ève d'avoir été là et pour son doux regard qui reste en moi.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
LISTE DES FIGURES	vii
RÉSUMÉ.....	viii
Introduction générale du mémoire.....	1
CHAPITRE I	
L'explication et la réduction selon les représentants de l'empirisme logique.....	13
1.1 La conception syntaxique des théories scientifiques en guise d'introduction au premier chapitre.....	13
1.2 L'explication scientifique selon Carl Gustav Hempel (et Paul Oppenheim).....	19
1.3 L'explication des lois et théories scientifiques selon le modèle DN et mon interprétation réductionniste des modèles DN et IS appliqués aux phénomènes naturels	27
1.4 La réduction selon Ernest Nagel	30
1.4.1 Les modèles de Hempel/Oppenheim et Nagel sont semblables.....	30
1.4.2 Le modèle interthéorique de la réduction nagélienne et ses concepts	32
1.4.3 Les énoncés de balises physiques	33
1.4.4 Les deux types de réduction et les hypothèses de liaison	34
1.4.5 Les deux conditions formelles de la réduction hétérogène	40
1.4.6 Les trois conditions informelles.....	42

1.4.7	Résumé des idées nagéliennes sur la réduction.....	43
1.5	Conclusion du premier chapitre	45

CHAPITRE II

La conception syntaxique des théories scientifiques est un cadre inadéquat pour formuler des explications réductionnistes ou pluralistes en psychiatrie 46

2.1	Introduction du deuxième chapitre.....	46
2.2	La réduction de la psychiatrie par la neurobiologie moléculaire	48
2.3	Les problèmes que pose la caractéristique DN du modèle nagélien.....	50
2.3.1	L'argument de Popper.....	50
2.3.2	L'argument de Feyerabend	52
2.3.3	La conclusion générale de Popper et Feyerabend	56
2.4	La solution de Schaffner	57
2.5	Le problème épistémique que rencontre le modèle de Schaffner	61
2.5.1	L'argument antiréductionniste de la réalisation multiple	61
2.6	Les problèmes que les modèles DN et IS rencontrent dans l'explication du fonctionnement du cerveau	65
2.6.1	Éléments fondamentaux du système nerveux	66
2.6.2	Le modèle des lois de couverture, le potentiel d'action et le problème de la pertinence	69
2.6.3	La libération des neurotransmetteurs et le problème lié à la thèse de la prévisibilité nomique.....	70
2.6.4	Le problème de la représentation de l'explication des troubles de santé mentale que posent les modèles DN et IS.....	72
2.7	Conclusion du deuxième chapitre	80

CHAPITRE III

L'explication réductionniste (Bickle) et pluraliste (Kendler) dans le cadre sémantique (et mécaniste) des théories scientifiques	82
3.1 Introduction du troisième chapitre	82
3.2 La conception sémantique des théories scientifiques.....	83
3.2.1 Les différentes conceptions du terme « modèle ».....	84
3.2.2 De ce que sont un mécanisme et un modèle mécaniste	93
3.3 Le modèle de l'explication mécaniste et réductionniste de Bickle.....	100
3.3.1 Le rôle des protéines CREB dans la transmission sociale des préférences alimentaires (TSPA) chez les souris	104
3.4 L'explication causale selon Bickle.....	108
3.4.1 Les faits empiriques supportant la thèse neurocausale de Bickle	116
3.5 Critique de l'épistémologie réductionniste et de la thèse empirique neurocausale de Bickle.....	120
3.6 Le modèle mécaniste de l'explication pluraliste de Kendler	140
3.7 La construction d'une explication mécaniste et l'exemple de la dépendance à l'alcool.....	148
3.8 Le modèle interventionniste de la causalité et la psychiatrie	154
3.8.1 Présentation du modèle interventionniste	156
3.9 Conclusion du troisième chapitre.....	174
Conclusion générale du mémoire	176
APPENDICE A Les cinq principes de Kandel.....	179
APPENDICE B Les huit principes de Kendler	185
RÉFÉRENCES	188

LISTE DES FIGURES

FIGURE 1 LE MODÈLE DE LA RÉDUCTION DE SCHAFFNER (SCHÉMA TIRÉ DE DIJADJI-BAHAMANI, FRIGG ET HARTMAN 2010, 399).	59
FIGURE 2 LA DÉPRESSION MAJEURE (SCHÉMA TIRÉ D'INSEL 2010, 46).	75
FIGURE 3 LA DÉPRESSION MAJEURE CHEZ LES FEMMES (SCHÉMA TIRÉ DE KENDLER ET COLL. 1993 CITÉS DANS MURPHY 2006, 136).....	77
FIGURE 4 MODÈLE MATÉRIEL REPRÉSENTANT LE MODÈLE ABSTRAIT $Y = AX+B$ (GRAPHIQUE TIRÉ DE GIERE 1999B, 49).	89
FIGURE 5 LE RÉDUCTIONNISME IMPITOYABLE (SCHÉMA TIRÉ DE BICKLE 2006, 426 CITÉ DANS ERONEN 2010, 22).	102
FIGURE 6 REPRÉSENTATION D'UN ÉNONCÉ EXPLICATIF CAUSAL PSYCHOLOGIQUE.....	110
FIGURE 7 REPRÉSENTATION DES ÉTAPES (D) À (F) DE L'ÉNONCÉ CAUSAL EXPLICATIF CELLULAIRE ET BIOMOLÉCULAIRE PRÉCÉDENT (2) (SCHÉMA TIRÉ DE BICKLE 2003, 69).	113
FIGURE 8 REPRÉSENTION DU FAIT SELON LEQUEL LES PERSONNES VIVANT AU SEIN DE SOCIÉTÉS INÉGALITAIRES SONT BEAUCOUP PLUS SUSCEPTIBLES DE SOUFFRIR DE MALADIES MENTALES QUE LES PERSONNES VIVANT DANS LES SOCIÉTÉS ÉGALITAIRES (GRAPHIQUE TIRÉ DE WILKINSON ET PICKETT 2009, 67).	127
FIGURE 9 LES QUATRE CRITÈRES QUE DOIT RESPECTER L'INTERVENTION IDÉALE (SCHÉMA ADAPTÉ DE CRAVER 2007, 97).	157

RÉSUMÉ

Dans ce mémoire, je présente et j'évalue trois modèles de l'explication en psychiatrie, à savoir (1) l'interprétation réductionniste que je fais des modèles déductif-nomologique (DN) et inductif-statistique (IS) de Carl Gustav Hempel (et Paul Oppenheim); (2) le modèle de l'explication réductionniste et mécaniste formulé par John Bickle et (3) le modèle de l'explication pluraliste et mécaniste développé par le psychiatre et chercheur Kenneth S. Kendler.

Au premier chapitre, je présente la conception syntaxique des théories scientifiques. J'y présente, en outre, les modèles DN et IS qui s'inscrivent dans cette conception. J'y expose aussi mon interprétation réductionniste de ces modèles. Je termine ce chapitre par une présentation du modèle de la réduction interthéorique d'Ernest Nagel.

Au deuxième chapitre, je présente une critique des modèles de la réduction interthéorique de Nagel et de Kenneth F. Schaffner. De plus, j'expose trois problèmes que rencontrent les modèles DN et IS. En montrant les failles de ces derniers, je conclus trois choses : (a) il n'est pas possible d'interpréter de manière réductive les modèles DN et IS; (b) ces modèles, même non interprétés de manière réductionniste, sont problématiques et, pour cela, inapplicables en psychiatrie; (c) la conception syntaxique des théories est un cadre inadéquat pour concevoir l'explication en médecine mentale.

Dans le troisième chapitre, je présente la conception sémantique des théories scientifiques. Suivant cette conception, une théorie est une famille de modèles. On peut définir le terme « modèle » de plusieurs manières. Patrick Suppes, Bas C. Van Fraassen et Ronald N. Giere ont proposé une définition de ce terme. J'adopte celle de Giere. Puis, j'expose la notion de modèle mécaniste qui peut être subsumée sous la conception des modèles de ce dernier. Je montre qu'un modèle mécaniste est pertinent pour représenter l'explication d'un trouble de santé mentale. Je termine en rejetant le modèle mécaniste et réductionniste de Bickle et en défendant le modèle mécaniste et pluraliste de Kendler.

Mots clés : philosophie des sciences, philosophie de la psychiatrie, explication, réduction, conception syntaxique des théories, conception sémantique des théories, modèle, causalité, psychiatrie, troubles de santé mentale, neurosciences.

INTRODUCTION GÉNÉRALE DU MÉMOIRE

La philosophie de la psychiatrie s'inscrit dans le mouvement de naturalisation et de spécialisation qui marque la philosophie depuis environ vingt-cinq ans (Faucher, 2006a, 3). En ce moment, ce domaine de recherche est particulièrement fleurissant. Témoignent de cette effervescence les diverses collections d'essais comme *Philosophy of Psychiatry: A Companion* (Radden, 2004), *Philosophie et psychopathologie* (Faucher, 2006a), *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology* (Kendler & Parnas, 2008) et *Psychiatry as Cognitive Neuroscience: Philosophical Perspectives* (Broome, M. & Bortolotti, 2009), les monographies comme *What is Mental Illness?* (McNally, 2011), *What is Mental Disorder?* (Bolton, 2008), *Psychiatry in the Scientific Image* (Murphy, 2006), *Psychiatry and Philosophy of Science* (Cooper, 2007) et *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder* (Horwitz & Wakefield, 2007), les revues (par exemple, la *Philosophy, Psychiatry, and Psychology*) et, enfin, les très nombreux articles publiés par les chercheurs et philosophes dans diverses revues *n'étant pas* spécialisées en philosophie de la psychiatrie comme, par exemple, le *International Journal of Integrated Care* (par exemple, Schaffner 2010) et le *Philosophy Compass* (par exemple, Murphy 2010b).

Plusieurs types de problématiques sont abordés par ceux qui œuvrent en philosophie de la psychiatrie¹. Certains d'entre eux s'intéressent à des sujets comme l'explication, la réduction et la classification en médecine mentale. Dans ces cas, ces philosophes conçoivent la psychiatrie comme une discipline scientifique particulière et l'examinent selon les méthodes et les concepts de la philosophie des sciences (voir Kendler 2001, 2005, 2008a; Kendler et Campbell 2009; Mitchell 2008a, 2008 b, 2009; Murphy 2006, 2008, 2009, 2010 b; Rachel 2007; Thagard 2008; Schaffner 1994, 2008a). Par exemple, Jeffrey Poland (2006) montre que le concept de schizophrénie n'est pas un concept scientifique et que, pour cette raison, il nuit à la recherche en psychiatrie, à l'intervention thérapeutique et, par voie de conséquence, au patient traité. D'autres chercheurs travaillent, quant à eux, sur le concept de maladie mentale ou de trouble de santé mentale (voir Bolton 2008; Cooper 2005, chapitre 1; Fuchs 2009; Kendler, K. S., Zachar, P., & Craver, C. 2011; Murphy 2006, chapitre 2; Oulis 2010; Wakefield 1992, 2006). Jerome C. Wakefield (2006), par exemple, conçoit le trouble de santé mentale comme une dysfonction préjudiciable, c'est-à-dire comme une dysfonction psychologique biologiquement fondée présente chez un individu et qui est jugée comme préjudiciable par la communauté ou l'environnement social où évolue ce même individu. D'autres s'interrogent, par ailleurs, sur les dimensions éthiques de la psychiatrie. Par exemple, Thomas Szasz (2007) explique en quoi, selon lui, la pratique psychiatrique viole les principes de liberté et de responsabilité individuelle qui font partie de l'humanisme séculier alors que Scott O. Lilienfeld (2007) s'inquiète des problèmes éthiques issus du fait que la pratique en psychiatrie est de moins en moins scientifiquement informée. D'autres réfléchissent à des problèmes philosophiques à la lumière des connaissances en psychopathologie. Par exemple, c'est en se basant sur des recherches scientifiques montrant que les psychopathes comprennent rationnellement les principes moraux, mais ne saisissent

¹ Je présente les différentes branches de la philosophie de la psychiatrie en suivant la classification qu'en fait Murphy (2010a).

pas émotionnellement pourquoi il faudrait les respecter que Shaun Nichols (2002) élabore la thèse selon laquelle l'éthique de type kantienne est intenable. En effet, selon la littérature scientifique, ce serait nos capacités d'empathie — sérieusement perturbées chez les psychopathes — qui nous pousseraient à agir moralement et non des raisonnements au sujet de ce qu'il convient ou non de faire. Enfin, certains philosophes s'attèlent à concevoir les diverses maladies mentales et ce qui les caractérise avec l'aide de l'appareillage conceptuel de la philosophie de l'esprit. Par exemple, George Graham et G. Lynn Stephens (2006) ont écrit un article sur la notion de délire qu'ils définissent comme « un type spécifique de défaillance de la connaissance et de la gestion de soi ». Dans le même ordre d'idées, Thomas Fuchs (2009, 230), en se basant sur ce que nous apprennent les « sciences cognitives incarnées » (« embodied cognitive science ») ou les théories de l'esprit incarné, propose que la schizophrénie résulte non pas seulement d'une dysfonction neurobiologique, mais d'une cascade d'événements subjectifs, neuronaux, sociaux et environnementaux en constante interaction.

En ce qui me concerne, dans ce mémoire, je considère la psychiatrie à la lumière des outils de la philosophie des sciences. En effet, je m'intéresse à la question de l'*explication* en médecine mentale. Plus précisément, je présente et j'évalue la pertinence de trois modèles de l'explication en psychiatrie, à savoir (1) l'interprétation réductionniste que je fais des modèles déductif-nomologique (DN) et inductif-statistique (IS) de Carl Gustav Hempel (et Paul Oppenheim) (Hempel 1965a [1942], 1965 b; Hempel et Oppenheim 1948); (2) le modèle de l'explication réductionniste et mécaniste formulé par John Bickle (2003, 2006, 2007, 2008a, 2008 b)² et (3) le modèle de l'explication pluraliste et mécaniste développé par le

² Je sais que Bickle affirme que son modèle de l'explication n'est pas un modèle mécaniste (Bickle 2006, 429-431). Mais son rejet du mécanisme repose essentiellement sur son refus d'admettre plusieurs niveaux épistémiques dans une explication de l'esprit. Or, en ajoutant le qualificatif

psychiatre et chercheur Kenneth S. Kendler (2001, 2005, 2008a, 2012; Kendler et Campbell 2009; Kendler, Zachar et Craver, C. 2011)³. Ultimement, je rejette les deux modèles réductionnistes et recommande le modèle pluraliste. Plus précisément, d'un point de vue épistémique, cela signifie, d'une part, que je m'oppose, de manière générale, à la thèse étiologique réductionniste selon laquelle les maladies mentales s'expliquent et sont causées fondamentalement par des anomalies biomoléculaires et, d'autre part, cela signifie que je défends l'idée selon laquelle la plupart des troubles mentaux doivent être expliqués par des mécanismes causaux multiples (moléculaires, physiologiques, psychologiques, sociales, etc.).

« réductionniste » au modèle que je qualifie de « mécaniste » de Bickle, je ne pense pas trahir la pensée de ce dernier. Par ailleurs, je sais que William Bechtel et Robert C. Richardson soutiennent eux aussi que le modèle de l'explication de Bickle n'est pas un modèle mécaniste (voir Bechtel et Richardson 2010, xxxvii). Leur thèse repose sur deux idées. La première est l'idée selon laquelle le modèle de l'explication de Bickle n'admet pas, contrairement au modèle mécaniste qu'ils ont développé, d'autres niveaux d'explication que le niveau cellulaire et biomoléculaire (il s'agit du même argument de Bickle que je viens d'énoncer). La deuxième est l'idée selon laquelle «Bickle's account [...] retains laws or generalisations, neither of which play a significant role in our own analyses [of mechanistic models] » (Bechtel et Richardson 2010, xxxvii). Je pense, comme je l'ai dit, m'accommoder de la première critique de Bechtel et Richardson en ajoutant le qualificatif « réductionniste » au modèle de Bickle. En ce qui concerne l'idée de ces derniers selon laquelle le modèle de Bickle suppose les concepts de lois et de généralisations, elle n'est pas conforme à ce que dit Bickle lui-même. De fait, ce dernier affirme qu'« [i]n current cellular and molecular neuroscience, as in cell and molecular biology generally, few explanations are framed in terms of laws or generalizations » (Bickle 2006, 429). Cela dit, cependant, je ne pense pas que Bickle, Bechtel et Richardson seraient contre l'idée qu'il y ait, ultimement, des généralisations régulant les mécanismes causaux. Par exemple, Bickle établit lui-même ce genre de généralisation quand il montre que l'activation des protéines CREB *cause* éventuellement la synthèse des protéines nécessaires à l'engrangement de nos souvenirs (à ce sujet, voir section 3.4). Afin de pouvoir utiliser la notion de généralisation, j'adopte le concept de mécanisme développé par Stuart Glennan (2002; 2005). Celui-ci, en effet, suivant James Woodward (2003), soutient la thèse selon laquelle les parties d'un mécanisme interagissent selon des *généralisations invariantes* (Glennan 2002, S344, mes *italiques*).

³ Je tiens à noter que le modèle de l'explication pluraliste de Kendler n'est pas le seul en son genre. Par exemple, Dominic Murphy (2006, 2008, 2009, 2010b) et Paul Thagard (1998, 2008; Thagard et Findlay, 2012) ont développé chacun un modèle de l'explication en psychiatrie très intéressant qui s'apparente sous bien des aspects à celui de Kendler. Kenneth F. Schaffner (1994, 2008), quant à lui, a aussi construit un modèle pluraliste de l'explication pour la médecine mentale. Je choisis de présenter celui de Kendler, toutefois, parce qu'il me semble intéressant de faire connaître un modèle élaboré par quelqu'un qui pratique la psychiatrie et qui fait de la recherche dans ce domaine selon les principes de son modèle de l'explication.

Selon certains philosophes (Bickle 2003; Cooper 2007, 107; Murphy 2006, 115-120; 2010a, section 2.3; Schaffner 1993, 412; 2010, 44), la tendance épistémique réductionniste en regard des maladies mentales se retrouve dans les idées du psychiatre et chercheur en neuroscience Eric Kandel⁴, un prix Nobel de médecine. De fait, ce dernier, dans son article « Un nouveau cadre conceptuel de travail pour la psychiatrie » (2002 [1998]), a formulé et explicité cinq principes desquels on peut, selon Dominic Murphy, tirer l'idée selon laquelle les processus moléculaires ayant lieu dans le cerveau sont à la base du développement des maladies mentales (Murphy 2010a). Bien sûr, Kandel sait que des facteurs culturels et comportementaux influencent l'expression des gènes, mais, en dernière instance, selon lui, les explications psychiatriques complètes n'auront qu'à mentionner des facteurs causaux génétiques (Murphy 2010a). Ce « réductionnisme génétique » (Murphy 2010a, section 2.3), note Murphy (2006, 116), repose sur deux idées, à mon avis complémentaires, de Kandel, à savoir le principe selon lequel (1) tout ce qui concerne la « "culture" est [...] finalement exprimé dans la "nature" » (Kandel, 2002 [1998], 19) et l'idée selon laquelle (2) « les gènes déterminent le phénotype » (ibidem, 22). Évidemment, le corollaire méthodologique de cette thèse épistémologique concernant l'explication est qu'il est préférable, dans la recherche scientifique, de prioriser (et de financer) les expériences qui permettent de découvrir les mécanismes biomoléculaires et cellulaires qui sont à l'origine des troubles de santé mentale (voir, par exemple, Akil et coll. 2010). Les propos du psychiatre Samuel Barondes en regard de la façon

⁴ J'écris « certains philosophes » parce que d'autres philosophes, comme Max Kistler, par exemple, ne semblent pas considérer Kandel comme un réductionniste. En effet, dans un article où il critique le réductionnisme, Kistler (2009, 76) se sert des propos de Kandel pour illustrer la position antiréductionniste ou, pour être plus précis, l'idée de causalité descendante. En ce qui me concerne, je ne suis pas certain de la façon dont on doit interpréter les travaux de Kandel. Cependant, cela n'a pas d'importance pour mon propos. En effet, que certains soutiennent que Kandel soit réductionniste ou que d'autres ne le soutiennent pas ne change rien au fait que le réductionnisme en psychiatrie est largement répandu dans notre société et que *l'interprétation réductionniste* que certains font des travaux de Kandel illustre bien cette école de pensée (pour un aperçu de cette interprétation réductionniste des principes de Kandel relatifs à la psychiatrie, voir, par exemple, Murphy 2006, 115-120 et Murphy 2010a, section 2.3 de même que ma propre interprétation — à laquelle je ne suis pas certain d'adhérer — que l'on retrouve à l'appendice A de ce mémoire).

dont il faut concevoir et étudier les facultés cognitives et les émotions illustrent très bien cette posture méthodologique réductionniste :

Molecules are the chemical machinery of our brains, so that to study them is to study the actual brain components involved in feeling and remembering instead of contemplating at a higher level of abstraction. (Barondes 1999, 46 cité dans Craver 2007, 12)

Enfin, on peut légitimement faire l'hypothèse que l'adoption de cette thèse étiologique réductionniste s'harmonise très bien avec, d'une part, l'idée selon laquelle l'intervention pharmacologique est la plus pertinente et la plus efficace pour traiter les maladies mentales⁵ et, d'autre part, avec le projet au cœur de la pharmacogénomique qui consiste à développer des médicaments (ici, pour le cerveau/esprit) qui conviennent au profil particulier du génome du patient que l'on veut soigner (voir Stahl 2008).

Quant à elle, toujours chez les chercheurs en neurosciences et les psychiatres, la posture épistémique pluraliste en regard des troubles de santé mentale se retrouve dans les idées de Kendler qui soutient que « les approches explicatives pluralistes sont préférables aux approches monistes explicatives, et en particulier au réductionnisme biologique [(comme celui que mettent de l'avant Bickle et Kandel)] » (Kendler 2005, 433, ma traduction). Cette position épistémique pluraliste relative à l'explication des maladies mentales est aussi celle de Luc Faucher (2006b; voir aussi Faucher, Poirier et Lachapelle, 2006), Ian Gold (2009; Kirmayer et Gold 2012), Dominic Murphy (2006, 2008, 2009, 2010b), Sandra Mitchell (2008a, 2008b, 2009), Alva Noë (2009, xii-xiii), Paul Thagard (1998, 2008; Thagard, et Findlay, 2012), Kenneth F. Schaffner (1994, 2008a) et James Woodward (2008a). Par exemple, dans

⁵ Je développe brièvement cette idée à la fin de la section 3.4.1 et à l'argument 8 de la section 3.5.

une section de son livre *Unsimple Truths* (Mitchell, 2009, p. 5-11), Mitchell met en lumière la complexité étiologique de la dépression majeure. En effet, elle y explique que, suivant les recherches de l'équipe d'Avshalom Caspi (Caspi, A. *et coll.* 2003; Caspi et Moffitt 2006), l'expression ou la non-expression des gènes responsables de cette pathologie mentale est fonction du type d'environnement dans lequel l'individu qui porte ces gènes évolue. Les facteurs causaux environnementaux, dans cette perspective pluraliste, peuvent présenter un pouvoir causal réel et, comme l'affirme Faucher, « il n'y a pas de raison d'accorder une préséance ontologique [(et causale, il me semble)] au niveau biomoléculaire » (Faucher 2006b, 406, ma traduction). J'ajouterais, par ailleurs, que si l'on conçoit l'étiologie des maladies mentales selon cette perspective explicative pluraliste, cela implique que les programmes de recherche en psychiatrie doivent nécessairement être pluridisciplinaires, c'est-à-dire qu'ils doivent favoriser la participation de *toutes les disciplines* scientifiques susceptibles d'être pertinentes dans la compréhension des affections de l'esprit. Par exemple, Gold pense que, pour comprendre ce que fait le cerveau, on doit réfléchir, entre autres, selon les disciplines des neurosciences, mais aussi selon la psychologie, la sociologie et l'anthropologie (Gold 2009, 511). Enfin, en ce qui concerne la pratique en santé mentale, on pourrait penser qu'adopter ce genre d'explication suggère qu'il faille considérer selon un degré égal d'importance toutes les formes d'interventions thérapeutiques (pharmacologique, psychologique, familiale, sociale, etc.). J'expliquerai très brièvement pourquoi je pense que c'est le cas dans mon travail (voir section 3.6).

Afin de bien faire comprendre la thèse explicative pluraliste que je défends, je présente, d'abord, dans le premier chapitre, la conception syntaxique des théories scientifiques formulée par les tenants de l'empirisme logique. Il est important de présenter cette interprétation des théories scientifiques parce que les modèles de l'explication DN (des phénomènes naturels et des théories) et IS (des phénomènes

naturels) ont été formulés dans ce cadre syntaxique (pour connaître ces modèles DN et IS, voir surtout Hempel 1965c; Hempel et Oppenheim 1948). Après avoir présenté les modèles DN et IS, j'expose le très influent modèle de la réduction interthéorique qu'a développé Ernest Nagel (1949, 1961, chap.11, 1970) en s'inspirant du modèle DN de Hempel et Oppenheim⁶. Au cours de ce chapitre, je m'inspire des idées de Hempel, Oppenheim et Nagel pour formuler ce que j'ai nommé « l'interprétation réductionniste des modèles DN et IS appliqués aux *phénomènes naturels* » (section 1.3).

Dans le deuxième chapitre, je montre que cette interprétation réductionniste des modèles de l'explication DN et IS appliqués aux *phénomènes naturels* est intenable. De fait, comme nous le verrons, pour que le modèle des lois de couverture — dans cette interprétation réductionniste — soit utilisable et viable, il faut que la réduction interthéorique telle que conçue par Hempel et, surtout, par Nagel, le soit aussi, et ce, tant pour les sciences en général que pour le cas particulier de la réduction de la psychiatrie par la neurobiologie. Or, ce type de réduction est difficilement opérable, comme l'ont vu Karl Popper et Paul Feyerabend (Feyerabend 1962; 1965; 1979 [1975]; Popper 1991 [1979]), mais surtout Jerry Fodor qui a montré que même la version révisée et améliorée du modèle nagelien interthéorique formulé par Schaffner (Schaffner 1967, 1969 b, 1974) — lequel résout les problématiques formulées par Popper et Feyerabend — rencontre des difficultés (voir Fodor 1974)⁷. Il faut ajouter à

⁶ En effet, il est intéressant de noter que le modèle de Nagel, selon ce dernier (Nagel 1970, 120 et 123) et d'autres philosophes (par exemple, Kim 2006 [1998], 58), est un modèle DN de l'explication des théories.

⁷ Mis à part Schaffner, je sais que d'autres philosophes que l'on associe au mouvement du « New Wave Reductionism » ont développé des modèles de la réduction pour contourner les problèmes que posent les modèles nagelien et schaffnerien (Bickle 1998; Churchland 1979, 80-88 et Hooker 1981). Mais, comme l'a montré Ronald Endicott (1998, 2001) et comme le rappellent les philosophes des sciences Foad Dizadji-Bahmani, Roman Frigg et Stephan Hartmann (2010, 402), tous ces modèles rencontrent, en dernière instance, les mêmes difficultés que le modèle nagelien/schaffnerien. C'est la raison pour laquelle je ne présenterai pas ces autres modèles.

cela que même si le modèle néonagelien de la réduction de Schaffner ne posait pas de problèmes, les modèles de l'explication DN et IS appliqués aux *phénomènes naturels* ne seraient pas davantage appropriés. En effet, ceux-ci présentent des difficultés — à savoir le problème de la pertinence et le problème lié au principe de la prévisibilité nomique — qui les rendent inopérant en neurosciences (Craver 2007, 34-40) et, par voie de conséquence, à mon avis, en psychiatrie⁸. On doit ajouter à ces problèmes une autre difficulté. En effet, selon moi, les modèles DN et IS donnent une mauvaise *représentation* de l'explication d'un trouble de santé mentale. En somme, de manière générale, ce chapitre me permettra de montrer que la conception syntaxique des théories scientifiques dans laquelle ont été développés les modèles de Hempel/Oppenheim, Nagel et Schaffner⁹ ne nous permet pas, ultimement, de formuler des conceptions viables de l'explication (et de la réduction) en médecine mentale.

Puisque la conception syntaxique des théories scientifiques est intenable, je présente, dans le troisième chapitre, la conception sémantique. Plusieurs philosophes ont contribué à l'élaboration de cette conception qui comporte plusieurs formulations (par exemple, Giere 1988; 1999a, 1999b, 2004, 2006, 2010; Suppe 1989; Suppes 1957, 1960; 1962; 1967, 2002; Van Fraassen 1980, 1987). De ces multiples formulations, celle de Ronald N. Giere me semble convenir particulièrement bien pour notre discussion. En effet, celui-ci a développé une notion générale de modèle adéquate, à mon avis, pour représenter une explication — qu'elle soit réductionniste ou pluraliste — en psychiatrie. De fait, pour Giere, un modèle est une entité

⁸ Et en science en général, d'ailleurs (voir Bromberger 1966; Salmon 1971).

⁹ Selon ma compréhension du cheminement intellectuel de Schaffner, le modèle de ce dernier, jusqu'en 1974 environ, s'inscrit dans la conception *syntaxique* des théories scientifiques. Ce n'est qu'à partir des années 80 que Schaffner développe un modèle de l'explication et de la réduction pour la médecine et la biologie s'inscrivant dans la conception *sémantique* des théories scientifiques. La forme achevée de ce modèle est explicitée dans son maître livre *Discovery and Explanation in Biology and Medicine* (1993).

représentationnelle et il se trouve que l'on peut subsumer, selon moi, sous cette façon de concevoir l'idée de modèle, le concept de *modèle mécaniste* élaborée par Stuart Glennan (2002, 2005), lequel *modèle mécaniste* convient pour représenter les troubles de santé mentale qui sont des mécanismes physiques. En effet, ces derniers sont difficilement représentables dans le cadre formel inhérent à l'interprétation syntaxique des théories scientifiques, mais ils le sont aisément selon des modèles mécanistes. Le premier modèle de l'explication mécaniste que je présente — et rejette — est le modèle réductionniste de Bickle. Le deuxième modèle de l'explication mécaniste que je fais connaître — et défends — est le modèle pluraliste de Kendler. Je présente le modèle de Bickle (2003, 2006, 2007, 2008a, 2008 b) parce qu'il décrit adéquatement, me semble-t-il, la structure des raisonnements explicatifs faits par la plupart des psychiatres à l'heure actuelle, lesquels raisonnements s'harmonisent très bien avec les cinq principes réductionnistes qu'a formulés Kandel (Kandel 2002 [1998])¹⁰. Bickle a développé ses idées sur l'explication en étudiant scrupuleusement ce qui se fait en neurosciences cognitives moléculaires (spécialement les travaux concernant les mécanismes moléculaires de la consolidation mnésique). Ce philosophe défend un réductionnisme impitoyable (« ruthless reductionism »). De fait, ce dernier soutient un fondationnalisme neural et éventuellement un fondationnalisme microphysique (Bickle 2003, 156-157; 2006, 431-432). Il défend cette thèse épistémologique parce que les études actuelles en neurosciences (voir Kandel et coll., 2000 cités dans Bickle 2003, 3) le persuadent que la *notion* d'esprit est réductible aux *concepts* représentant les processus biomoléculaires dans le cerveau (Bickle 2003, 3). Pour cette raison, *expliquer* l'esprit, les émotions et les comportements (et, par voie de conséquence, les pathologies mentales) consiste, en dernière instance, à donner l'histoire biochimique à l'œuvre dans le système cérébral. En plus de cette thèse épistémologique réductionniste, Bickle défend ce que je nomme une thèse « neurocausale ». Celle-ci est basée sur deux constantes expérimentales qui, selon

¹⁰ Pour connaître ces principes, voir l'appendice A de ce mémoire.

Bickle, démontrent la supériorité de la force causale des événements neuraux sur celle des événements environnementaux (Bickle 2008b, 472-473). Mais cette thèse empirique neurocausale et l'épistémologie réductionniste qu'adopte Bickle sont problématiques, comme je le démontrerai. En effet, principalement, elles sont contredites par plusieurs faits scientifiques relatifs à l'étiologie des troubles de santé mentale (voir, par exemple, Kendler 2005, 436-437; Kirmayer et Gold 2012; McNally 2011). Pour cela, je soutiens qu'en médecine mentale il faut abandonner le modèle mécaniste et réductionniste de l'explication de Bickle. C'est pourquoi je présente et soutiens, ensuite, le modèle mécaniste et pluraliste de l'explication en psychiatrie qu'a développé le psychiatre et chercheur Kendler (2001, 2005, 2008a, 2012; Kendler et Campbell 2009; Kendler, Zachar et Craver, C. 2011). Le modèle de Kendler repose sur huit principes pluralistes¹¹. Ce dernier a construit son modèle à partir de données scientifiques que révèlent ses propres recherches (voir spécialement Kendler et Prescott 2006) et conformément à la littérature philosophique concernant l'explication en médecine mentale (Schaffner 1994, 2008). Il s'est aussi inspiré des modèles mécanistes (Bechtel 2008; Bechtel et Richardson 2010 [1993]; Craver 2007; Glennan 2002, 2005) et du pluralisme intégratif (Mitchell 2003, 2008a, 2008 b, 2009). Enfin, Kendler a intégré à son modèle de l'explication le modèle interventionniste (MI) de la causalité de James Woodward¹². Le MI, contrairement à celui de Bickle, n'est pas caractérisé par un fondationnalisme causal (Woodward 2008a, 146) et ne rencontre pas le problème de la pertinence que rencontre la théorie nomologique de la causalité de Hempel.

Avant de passer au premier chapitre, j'aimerais faire quelques remarques au sujet des notions d'explication, de réduction et de causalité. Il est effectivement important

¹¹ Pour connaître ces principes, consulter l'appendice B de ce mémoire.

¹² Ce modèle de la causalité s'inscrit dans la foulée des critiques ayant été faites en regard de la *conception de la causalité* que l'on retrouve dans le modèle hempelien que je présenterai dans le premier chapitre (section 1.2) de mon mémoire.

de noter que *chacune* de ces notions peut être étudiée sans nécessairement faire référence aux autres notions. Par exemple, on peut aborder le concept d'explication sans s'attarder sur le concept de réduction (voir, entre autres, Hempel 1965c; Salmon 1971; Van Fraassen 1980, chap. 5; Woodward 2003) ou on peut travailler le concept de réduction sans aborder le concept de causalité (par exemple, Dizadji-Bahmani, Frigg et Hartmann 2010; Nagel 1961, chap. 11). On peut aussi travailler le concept d'explication sans aborder le concept de causalité puisqu'il existe, comme le souligne Max Kistler (2011, 109), «[...] des explications scientifiques qui *ne sont pas* causales [...]». Dans le même ordre d'idées, mentionnons que les explications en mathématique elles aussi ne sont pas de nature causale.

Je prends le temps de noter cela parce que, dans mon mémoire, ces notions d'explication, de réduction et de causalité sont mises en relation par moi, Bickle et Kendler. Pour ma part, comme je l'ai dit, je formule une interprétation *réductionniste* du modèle de l'*explication (causale)*¹³ de la loi de couverture. Quant à Bickle, il conçoit la *réduction* d'un concept psychologique comme une *explication réductive* et *causale* de ce même concept. En d'autres termes, pour Bickle, réduire un concept psychologique consiste à donner l'histoire biomoléculaire *causale* responsable de la fonction cérébrale décrite par ce concept psychologique. Quant à Kendler, il conçoit son modèle de l'*explication* comme un modèle de l'*explication causale*. De fait, ce dernier appréhende la notion de relation causale selon le MI formulé par Woodward.

¹³ En effet, le modèle des lois de couverture implique, implicitement, la théorie nomologique de la causalité (voir section 1.2).

CHAPITRE I

L'EXPLICATION ET LA RÉDUCTION SELON LES REPRÉSENTANTS DE L'EMPIRISME LOGIQUE

1.1 La conception syntaxique des théories scientifiques en guise d'introduction au premier chapitre

Carl Gustav Hempel, Paul Oppenheim et Ernest Nagel sont parmi les représentants de l'empirisme (ou du positivisme)¹⁴ logique les plus connus. Avant d'aborder les modèles de l'explication que ceux-ci ont développés, on doit avoir en tête que ces penseurs conçoivent les théories scientifiques comme des structures linguistiques exprimées de façon axiomatique dans le langage de la logique des prédicats du premier ordre (voir surtout les points 10 à 12 ci-dessous). On a qualifié de « syntaxique » cette conception des théories scientifiques. Les points (1) à (14) suivants rendent compte de ce qui me semble être l'essence de la conception des

¹⁴ Comme l'explique Marion Vorms (2011, 68, note 9), les expressions « positiviste » et « empiriste » sont pratiquement équivalentes. Toutefois, affirme-t-elle, « on peut supposer, comme me l'a suggéré Pierre Jacob, que l'on choisit l'une ou l'autre selon que l'on cherche à insister davantage sur l'ancrage anti-métaphysique (positiviste) ou sur l'héritage de l'empirisme classique (représenté en particulier par David Hume). Il existe également, [ajoute-t-elle], une nuance temporelle entre ces deux notions : l'expression "empirisme logique" désigne souvent la version plus tardive libéralisée du programme positiviste logique » (ibidem, 68, note 9).

théories (et du monde) que soutiennent les défenseurs de cette école de pensée¹⁵ et à l'intérieur de laquelle il faut s'imaginer l'idée d'explication relative aux phénomènes naturels et aux théories (voir le point 14) que je présenterai dans les prochaines sections de ce premier chapitre¹⁶.

- 1) Selon les représentants de l'empirisme logique, nous n'avons pas accès à la *réalité en soi* ou à la *structure fondamentale de l'univers* (voir, par exemple, Nagel 1961, 364).
- 2) Suivant (1), aucun langage ne peut nous informer sur la *réalité en soi* ou sur la *structure fondamentale de l'univers*.
- 3) (1) et (2) forme une thèse épistémologique *antiréaliste*. En effet, une conception épistémologique *réaliste* se traduirait par la thèse selon laquelle la *réalité en soi* est connaissable (quoique partiellement, probablement).
- 4) Cependant, nous avons accès à la *réalité empirique*.
- 5) La *réalité empirique* consiste en un complexe de sensations que les sens permettent de percevoir. (C'est la thèse phénoménaliste ou descriptiviste de l'empirisme logique.)

¹⁵ Les points (1) à (14) correspondent à la conception des théories scientifiques d'aucun tenant de l'empirisme logique *en particulier*. Carnap, Nagel et Hempel, par exemple, comme l'a montré Vorms (2011, chapitre 2), ne s'entendaient pas sur tout concernant ce que devait être une théorie scientifique. Cependant, ces derniers s'entendaient sur certains grands principes. Il me semble que les points (1) à (14) sont compatibles avec ces principes et auraient été acceptés par la plupart des membres de l'empirisme logique.

¹⁶ Ces points (1) à (14) consistent en ma compréhension personnelle des principales thèses se rapportant à la conception des théories (et du monde) des tenants de l'empirisme logique. J'ai cru qu'il serait utile d'exposer de cette manière les idées des philosophes appartenant à ce courant de pensée parce que, dans ce mémoire, je réfère souvent à des points précis de cette section 1.1. Je dois noter, par ailleurs, que Frederick Suppe a formulé lui aussi de manière schématique (et plus technique) cette conception syntaxique des théories scientifiques (voir Suppe 1974, 50-51 ou, pour la traduction française de cette formulation, voir Vorms 2011, 124-125). Suppe, toutefois, dans sa formulation, ne mentionne pas explicitement les corollaires ontologiques et épistémiques que l'on peut tirer, me semble-t-il, des thèses caractérisant la conception syntaxique des théories scientifiques (voir les points 13 et 14 de ma présentation).

- 6) Il existe un langage L (a) identique à l'ensemble des théories scientifiques connues et à connaître et (b) qui rend compte de ce que les humains perçoivent de la *réalité empirique*¹⁷.
- 7) Conformément à (6), à l'exception des termes logiques et mathématiques (voir le point 8a suivant), le domaine d'interprétation ou le cadre sémantique de L est la *réalité empirique*.
- 8) L contient des termes de trois types : (a) les *termes logiques et mathématiques* (b) les *termes théoriques* et (c) les *termes observationnels*.
- a. Intrinsèquement, les *termes logiques et mathématiques* n'ont pas de contenu empirique, c'est-à-dire qu'ils ne réfèrent pas à la réalité empirique. Par exemple, l'opérateur de la relation d'identité « \equiv » ne réfère à rien dans la réalité empirique.
- b. Les *termes théoriques* « font référence à des choses ou à des processus inobservables » (Vorms 2011, 77). Idéalement, il faut donner une interprétation empirique (observationnelle) à ces termes théoriques. Par exemple, le terme théorique « $2E/3$ », à savoir l'énergie cinétique moléculaire moyenne que l'on ne peut pas voir directement, peut être interprété en termes observationnels (voir le sous-point 8c suivant). On peut donner une interprétation empirique à « $2E/3$ » si on réussit à construire une *règle de correspondance* ayant la forme logique $Tx \equiv Ox$ où « T » est un terme théorique et « O » un terme observationnel (Suppe 1974, 12). Par exemple, pour le terme « $2E/3$ », il y a la règle de correspondance « $kT = 2E/3$ » où l'on identifie la température « T » en Kelvin (d'où « K »), un terme observable (à l'aide d'un thermomètre, par exemple) à l'énergie cinétique moléculaire moyenne « $2E/3$ ».

¹⁷ Il me semble qu'il pourrait exister d'autres langages L_1, L_2, \dots, L_i construit (a) par d'autres espèces vivantes que les humains et (b) conformément à ce que ces derniers perçoivent de la *réalité empirique*. De plus, conformément à (2), en regard de la *réalité en soi*, je crois que l'on ne peut pas dire si L est plus ou moins vrai que L_1, L_2, \dots, L_i . Par ailleurs, j'aimerais dire aussi que, comme L est une construction humaine, alors L change avec le temps T . Ainsi, quand on « consulte » L , il faut toujours avoir conscience qu'il s'agit de L au temps T .

- i. Note : les termes mathématiques et théoriques, selon les tenants de l'empirisme logique qui furent fortement influencés par l'ouvrage d'Alfred North Whitehead et Bertrand Russell *Principia Mathematica* (1910-913), sont tous traduisibles dans le langage de la logique des prédicats du premier ordre (voir Suppe 1974, 12).
 - c. Les *termes observationnels* font référence à « des entités, des propriétés, des relations ou des processus observables » (Vorms 2011, 77). Par exemple, les termes « chaud », « solide » ou « jaune » peuvent être directement vérifiés par nos sens, c'est-à-dire qu'ils sont directement vérifiables dans la réalité empirique.
- 9) *L* contient des *énoncés* de deux types :
- a. les *énoncés théoriques* composés d'au moins deux termes théoriques (par exemple, « Dans cette chambre à bulles, *un photon gamma vient de se transformer en un électron et un positron.* » [Kistler 2007, 2, mes *italiques*]) ou, au moins, d'un terme observationnel et d'un terme théorique (les règles de correspondance comme « $kT = 2E/3$ » sont des énoncés théoriques);
 - i. C'est grâce aux règles de correspondance que l'on peut donner une interprétation empirique à la théorie (voir point 12).
 - b. Les *énoncés observationnels* composés uniquement de termes observationnels (par exemple, « La table est jaune »).
- 10) Toutes les théories scientifiques connues et à connaître sont contenues dans *L* (voir le point 6a) et forment un ensemble ou un réseau axiomatique d'énoncés théoriques et observationnels (voir le point 9).
- 11) Ces théories (et les énoncés qu'elles contiennent) sont, *fondamentalement*, une structure axiomatique exprimée dans le langage de la logique des prédicats du premier ordre et sont, pour cela, tout comme les énoncés qui les composent, déductibles les unes des autres.

12) *Idéalement*, tous les termes et énoncés des théories scientifiques (à l'exception des termes et énoncés logiques et mathématiques) doivent être *justifiés* ou *vérifiés*, c'est-à-dire *interprétés empiriquement* par le biais des *règles de correspondances* (voir points 7 et 8 b). *Interpréter empiriquement* un énoncé consiste à identifier cet énoncé à un nombre fini d'énoncés observationnel.

- a. J'écris « idéalement » parce que, dans les faits, les théories scientifiques — ou, plus précisément, les théories empiriques — sont toujours partiellement interprétées (Carnap 1995 [1966], 237). Il en est ainsi parce que, entre autres, « les lois de la nature¹⁸ ont une forme universelle, et ne peuvent pas par conséquent être logiquement déduites d'un nombre fini d'énoncés : l'inférence de "tous les corbeaux que j'ai observés jusqu'à présent sont noirs" à "tous les corbeaux sont noirs" n'est pas logiquement rigoureuse » (Vorms 2011, 76 note 21).
- b. Quand l'on ne donne pas d'interprétation empirique à *une théorie*, celle-ci se réduit à ce qu'elle est fondamentalement, c'est-à-dire à une structure logique non interprétée ou, comme je le disais dans l'introduction de cette section, à une structure linguistique exprimée de façon axiomatique dans le langage de la logique des prédicats du premier ordre.

13) Le *corollaire ontologique* des points (1) à (12) est qu'exister ne peut avoir qu'une signification empirique. En effet, dans le cadre de l'empirisme logique, « exister » signifie « être reconnu comme une entité de *L* ». C'est ce qu'explique Rudolf Carnap, un représentant éminent de l'empirisme logique, lorsqu'il affirme que

[t]he concept of reality [...] is an empirical scientific non-metaphysical concept. To recognize something as a real thing or event means to succeed in incorporating it into the system of things at a particular space-time position so that it fits together with the other things as real, *according to the rules of the framework* [(c'est-à-dire selon les règles de *L*)]. (Carnap 1950, mes *italiques*)

¹⁸ Je présente la conception des lois de la nature que mettent de l'avant les représentants de l'empirisme logique à la section 1.2 (voir la note de bas de page numéro 21).

14) Le corollaire épistémique des points (1) à (12) est que, grosso modo, dans le cadre de *L*

- a. l'explication *d'un phénomène naturel* consiste à déduire ou induire logiquement l'énoncé observationnel décrivant le phénomène que l'on veut expliquer d'énoncés théoriques — ici, de lois universelles ou statistiques — et d'énoncés observationnels décrivant les conditions physiques dans lequel ce phénomène a lieu (ou aurait lieu, si l'on faisait une prédiction) (voir Hempel 1965c et Hempel et Oppenheim 1948).
- b. l'explication — on pourrait dire, aussi, comme on le verra, la réduction — *d'une théorie* consiste à déduire les lois et postulats d'une théorie de ceux d'une autre théorie dite plus fondamentale ou dont la force explicative est plus grande (voir Hempel et Oppenheim 1948; Nagel 1949; 1961, chapitre 11 et 1970).

Aux chapitres un et deux, j'explique le point (14). Plus précisément, dans le premier chapitre, je présente les modèles de l'explication déductive-nomologique (DN) et de l'explication inductive-statistique (IS) (section 1.2). Puis, je fais connaître le concept de réduction/explication interthéorique (sections 1.3 et 1.4). Tout en présentant ces notions d'explication et de réduction, j'explique à la section 1.3 en quoi consisterait, selon moi, une explication réductionniste des phénomènes naturels dans le cadre de la conception syntaxique des théories que je viens de présenter. Dans le deuxième chapitre, je montre que ce type linguistique, formel et axiomatique¹⁹ d'explication réductionniste n'est pas praticable en psychiatrie. Ultimement, il n'est pas praticable parce que (a) la réduction interthéorique, dans le cadre logico-syntaxique, pose un problème important (section 2.5); parce que (b) les modèles DN

¹⁹ J'emploie ici le terme « axiomatique » dans le sens « d'axiomatiser selon la logique des prédicats du premier ordre ». En effet, une axiomatisation n'est pas nécessairement écrite dans ce langage logique. On peut, par exemple, axiomatiser une théorie selon la théorie des ensembles et des modèles comme l'a fait, entre autres, Patrick Suppes (1957) (voir aussi la section 3.2.1 de ce mémoire pour un exemple d'axiomatisation selon la logique des prédicats et la théorie des ensembles).

et IS posent des problèmes en neurosciences et, par voie de conséquence, en psychiatrie (sections 2.6.2 et 2.6.3) et parce que (c) ces mêmes modèles ne nous donnent pas une représentation adéquate des troubles de santé mentale (section 2.6.4).

1.2 L'explication scientifique selon Carl Gustav Hempel (et Paul Oppenheim)

Il est communément admis que, chez les tenants de l'empirisme logique, Carl Gustav Hempel (1965a [1942], 1965 b, 1965c; Hempel et Oppenheim 1948) est le philosophe qui a formulé deux des plus influents modèles de l'explication scientifique. Ce dernier a développé, en effet, le modèle déductif-nomologique (modèle DN) et le modèle inductif-statistique (modèle IS) de l'explication qui sont devenus des classiques de la littérature en philosophie des sciences.

Selon le modèle DN, et conformément à la conception syntaxique des théories scientifiques (section 1.1), une explication présente la structure logique d'un *argument déductif*. Ce dernier est composé d'un *explanans* (ce qui explique) et d'un *explanandum* (ce qui est à expliquer). L'*explanans* est constitué (1) d'énoncés linguistiques empiriquement vérifiables²⁰ que Hempel nomme les conditions initiales C_1, C_2, \dots, C_k et (2) d'énoncés linguistiques référant à des lois universelles L_1, L_2, \dots, L_r , c'est-à-dire à des énoncés généraux décrivant des *régularités dans la nature*²¹. Quant à l'*explanandum*, il est un énoncé linguistique référant à l'événement

²⁰ Hempel, en effet, affirme que « [...] les propositions qui constituent une explication scientifique doivent pouvoir se prêter à des tests empiriques » (Hempel 1972 [1966], 76).

²¹ Chez les empiristes logiques, on considère généralement que ces lois universelles ont la forme des conditionnelles universelles que l'on exprime dans le langage formel de la logique des prédicats de la manière suivante : $\forall x(Fx \rightarrow Gx)$. Cependant, Hempel ajoute que, pour être des lois universelles, ces énoncés généraux doivent aussi supporter les énoncés conditionnels contraires aux faits et les énoncés conditionnels subjectifs. Par exemple, affirme Hempel (1965c, 339), la loi des gaz de Boyle-Mariotte

E que l'on veut expliquer. Dans le cadre conceptuel du modèle DN, expliquer scientifiquement quelque chose consiste à *déduire* de la conjonction de C_1, C_2, \dots, C_k et L_1, L_2, \dots, L_r l'événement E . En d'autres termes, ce type d'explication des faits, selon Hempel, peut être identifié à une déduction où *l'explanandum* est déduit par subsomption sous des lois générales (Hempel, 1972, 79). Hempel note que lorsque nous faisons cette opération déductive, nous répondons à la question suivante : pourquoi est-ce que l'événement E décrit par *l'explanandum* était prévisible? (Hempel 1965c, 337). Hempel et Oppenheim représentent le modèle DN par la structure logique suivante (Hempel et Oppenheim 1948, 138) :

$$\begin{array}{l} C_1, C_2, \dots, C_k \text{ (énoncés des conditions initiales)} \\ L_1, L_2, \dots, L_r \text{ (énoncés des lois universelles et déterministes de la nature)} \\ \hline E \text{ (description du phénomène empirique qui est à expliquer)} \end{array}$$

G. Randolph Mayes (2005, section 2) propose cette explication hempelienne de la maladie génétique qu'est la trisomie 21 ou le syndrome de Down :

supporte l'énoncé conditionnel contraire aux faits suivant : si l'oxygène dans ce cylindre avait été chauffé (ou était chauffé) sous une pression constante, alors il aurait pris (ou prendrait) de l'expansion. À l'inverse, l'énoncé conditionnel universel « tous les membres du conseil de l'école Greenbury, en 1964, étaient chauves » ne supporte pas le conditionnel subjectif suivant : « Si Robert Crocker avait été un membre du conseil de l'école Greenbury en 1964, il aurait été chauve. » Mais, en dernière instance, selon James Woodard (2009), il semble que Hempel conçoive les lois comme des énoncés généraux décrivant des régularités dans la nature. De fait, Woodward écrit :

In his (1965[c]) he [(Hempel)] considers a number of familiar proposals having this character and finds them all wanting, remarking that the problem of characterizing the notion of law has proved "highly recalcitrant" (1965[c], 338). It seems fair to say, however, that his underlying assumption is that, at bottom, *laws are just exceptionless generalizations describing regularities that meet certain additional distinguishing conditions that he is not at present able to formulate* (Woodward 2009, section 2.2, mes italiques).

Ce fœtus possède 3 copies du chromosome 21. (**Condition initiale**)

Toute personne possédant 3 copies du chromosome 21 souffrira du syndrome de Down.²² (**Loi universelle**)

Lorsqu'il naîtra, ce fœtus souffrira du syndrome de Down. (**Description du phénomène empirique qui est à expliquer**)

Le modèle IS est très semblable au modèle DN. Toutefois, la loi ou les lois qu'il présente, contrairement à ce dernier, ne sont pas des lois universelles ($\forall x [Fx \rightarrow Gx]$), mais des lois statistiques présentant la forme statistique/probabiliste suivante : $P(G|F)=r$.²³ (Hempel 1965c, 376). Ainsi, contrairement au modèle DN où l'explication a la forme d'un argument déductif, dans le modèle IS, celle-ci présente la forme d'un argument inductif. Par voie de conséquence, « l'explanans implique l'explanandum non avec "une certitude déductive" [(comme dans le modèle DN)], mais avec une quasi-certitude ou encore avec une forte probabilité » (Hempel 1972 [1966], 91, mes *italiques*). On peut représenter le modèle IS selon cette forme générale :

C_1, C_2, \dots, C_k (énoncés des conditions initiales)

L_1, L_2, \dots, L_r (énoncés des lois statistiques de la nature)

===== [r]

E (description du phénomène empirique qui est à expliquer)

Hempel illustre ce type d'explication, entre autres, avec le cas suivant (Hempel 1972 [1966], 91) :

²² Je dirais, cependant, que cet énoncé ou cette loi de la nature présente une forme biconditionnelle $\forall x (Fx \leftrightarrow Gx)$ plutôt qu'une forme conditionnelle $\forall x (Fx \rightarrow Gx)$. En effet, toutes les personnes possédant 3 copies du chromosome 21 présentent le syndrome de Down et vice versa.

²³ En mot, cela signifie que la probabilité de G étant donné F est de r où, dans le cadre de la théorie des probabilités, la valeur de r peut se situer entre 0 et 1. Mais, dans le cadre d'une explication IS, la valeur de r doit être élevée parce que l'explanandum doit être le cas avec une grande probabilité.

Pierre a été en contact avec une personne atteinte de rougeole (**Condition initiale**)
 La probabilité pour que des gens en contact avec une personne atteinte de rougeole l'attrapent est élevée (**loi statistique**)

===== [r]

Pierre a attrapé la rougeole (**Description du phénomène empirique qui est à expliquer**)

Par ailleurs, notons que les explications de type DN et IS susmentionnées peuvent aussi être interprétées comme représentant une *relation causale* entre C_1, C_2, \dots, C_k et l'événement E que décrit l'*explanandum*. En effet, à la lumière de ces modèles, on peut affirmer que C_1, C_2, \dots, C_k cause l'événement E parce qu'il existe des lois pertinentes L_1, L_2, \dots, L_r en vertu desquelles l'occurrence des conditions antécédentes C_1, C_2, \dots, C_k est nomologiquement suffisante [(ou presque suffisante si c'est une loi statistique)] à l'occurrence de l'événement E (Psillos 2002, 223). De fait, écrivent Hempel et Oppenheim²⁴,

If E describes a particular event, then the antecedent circumstances described in the sentences C_1, C_2, \dots, C_k may be said jointly to "cause" that event, in the sense that there are certain empirical regularities, expressed by the laws L_1, L_2, \dots, L_r , which imply that whenever conditions of the kind indicated by C_1, C_2, \dots, C_k occur, an event of the kind described in E will take place. (Hempel et Oppenheim 1948, 139-140, mes *italiques*)

Comme on peut le constater, la relation causale, dans cette perspective, n'est pas appréhendée selon son essence. En effet, à l'instar de Rudolf Carnap (Carnap 1928, 35-36 dans Psillos 2002, 216), Hempel ne s'intéresse pas au problème métaphysique de la nature profonde de la causalité. Autrement dit, il ne fait pas d'hypothèse sur la

²⁴ Hempel et Oppenheim, de même que Stathis Psillos, donnent cette interprétation causale en regard du modèle DN. Selon moi, et c'est pourquoi j'en fais mention, on peut aussi donner cette interprétation pour le modèle IS.

relation ontique nécessaire qui pourrait exister entre une cause et son effet. Il s'intéresse plutôt à la relation causale d'un point de vue épistémique. Or, selon cette conception que l'on pourrait qualifier d'humienne de la causalité, une relation causale n'est rien de plus qu'une corrélation — c'est-à-dire une fonction mathématique — existant entre des énoncés linguistiques (C_1, C_2, \dots, C_k et l'événement E que décrit l'*explanandum*) qui réfèrent à des objets de la *réalité empirique* (voir section 1.1, point 5) et qui sont subsumés sous une ou des lois générales L_1, L_2, \dots, L_r . Westley Salmon semble voir dans cette conception de la causalité l'essence de l'explication scientifique telle qu'elle est comprise dans les modèles de Hempel et qu'il nomme, pour sa part, la « prévisibilité *nomique* », à savoir l'idée selon laquelle la prévisibilité d'un événement E s'explique par le fait que des lois universelles ou statistiques L_1, L_2, \dots, L_r « connectent » E avec certitude ou avec une grande probabilité à d'autres événements décrits par C_1, C_2, \dots, C_k (Salmon 1989, 57 cité dans Woodward 2009, section 2.4)²⁵.

Je finis cette exposition des modèles de Hempel avec les cinq conditions que doivent respecter, selon Salmon, les explications DN et IS (Salmon 1999, 26, ma traduction)²⁶ :

1. L'explication doit être un argument ayant une forme logique adéquate (déductive ou inductive).
2. L'*explanans* doit contenir au moins une loi générale (universelle ou statistique); et cette loi doit être requise pour la déduction²⁷ de l'*explanandum*.

²⁵ Selon cette perspective hempelienne (et humienne) de la causalité, relativement au syndrome de Down (voir l'exemple plus haut), on pourrait dire qu'étant donnée la loi de la nature selon laquelle pour tout fœtus, si un fœtus présente 3 copies du chromosome 21, alors ce fœtus deviendra une personne présentant une trisomie 21, on peut affirmer qu'avoir 3 copies du chromosome 21 *cause* le développement du syndrome de Down.

²⁶ On retrouve évidemment ces conditions dans les textes de Hempel. Cependant, il m'a semblé que le résumé de Salmon était plus pertinent parce que, ce dernier, contrairement à Hempel, modifie la formulation de ces conditions afin qu'elles s'appliquent aux deux types de modèles (DN et IS).

3. *L'explanans* doit avoir un contenu empirique; il doit être possible, au moins en principe, de le tester par une expérience ou une observation.
4. Les énoncés qui composent l'*explanans* doivent être vrais²⁸.
5. L'explication doit satisfaire l'exigence de spécificité maximale (« requirement of maximal specificity »).

Avant d'explicitier ces clauses, puisque la cinquième est plus difficile à appréhender que les autres qui sont facilement compréhensibles, j'aimerais m'attarder sur elle un instant. D'abord, il faut dire que l'exigence de spécificité maximale concerne essentiellement l'explication IS. (Elle est automatiquement respectée, en effet, dans une explication DN. J'expliquerai pourquoi bientôt). Comme le fait savoir Salmon (1999, 26)²⁹, conformément à cette exigence, lorsqu'on construit une explication IS, nous devons y inclure *toutes les informations disponibles pertinentes qui nous sont connues*. Ainsi, contrairement à l'explication DN, comme on peut le constater, l'explication IS « est empreinte d'une "relativité épistémique" par rapport à un système particulier de connaissances » (Barberousse, Kistler et Ludwig 2000, 121). Une explication DN, en effet, ne présente pas de relativité épistémique parce qu'elle contient nécessairement une loi universelle qui vaut en tout temps et en tout lieu et qui n'admet aucune exception. Dans ce cas, ce que connaît le sujet qui formule une explication DN n'a aucune incidence sur la probabilité de l'*explanandum*. Cette probabilité, effectivement, dans une explication DN, est toujours de 1 : si Socrate est un homme, et si tous les hommes sont mortels, *peu importe ce que je sais d'autres* sur les hommes et Socrate, il est certain que Socrate est mortel. À l'inverse, et là je

²⁷ Salmon utilise le mot anglais « derivation ». Je traduis avec le mot « déduction » parce que, dans les modèles des lois de couverture, comme me l'a dit Serge Robert, on ne dérive pas l'*explanandum* de l'*explanans*, on le déduit.

²⁸ Hempel et Oppenheim, toutefois, affirment qu'il serait plus juste d'affirmer que l'*explanans*, plutôt que d'être *vrai*, doit être *confirmé* par ce que nous dit la science de l'époque (Hempel et Oppenheim, 137; voir aussi la note de bas de page 17 [section 1.1] de ce mémoire). En effet, un énoncé peut être corroboré par les faits à un stage S_1 de la science et être invalidé ultérieurement à un stage S_2 .

²⁹ Voir Hempel 1965c, 397-403 pour l'explication originale plus technique de l'exigence de spécificité maximale.

paraphrase un exemple de Hempel (1965c, 394), pour être à peu près certain que Jean qui présente une infection à streptocoques guérira s'il prend de la pénicilline, il n'est pas suffisant de savoir que la probabilité de guérir d'une infection à streptocoques est élevée si on est soigné à la pénicilline. En effet, le respect de l'exigence de spécificité maximale commande que l'on s'assure de savoir, entre autres, si Jean présente une infection résistante à la pénicilline. *Si nous le savons*, en effet, l'exigence de spécificité maximale requiert que l'on tienne compte de ce fait dans la construction de cette explication IS et que l'on formule une loi statistique qui tiendra compte de ce fait. En ce cas, cette loi statistique prévoira avec quasi-certitude *l'explanandum* selon lequel « Jean n'a pas guéri de son infection résistante à la pénicilline ». Toutefois, *si nous ne savons pas* que Jean présente une souche résistante à la pénicilline, alors ce ne sera pas une faute de ne pas inclure cette information dans l'explication IS et de ne pas changer la loi statistique en conséquence.

Maintenant que j'ai jeté de la lumière sur l'exigence de spécificité maximale, j'explique toutes ces cinq conditions en prenant pour exemple le cas mentionné plus haut où l'on explique la raison pour laquelle Pierre a attrapé la rougeole.

Premièrement, conformément au premier point, on remarque que la forme logique de cet argument inductif est adéquate. De fait, on y retrouve des prémisses et une conclusion. La deuxième condition est aussi observée. Effectivement, cet argument inductif contient une loi statistique, laquelle, d'ailleurs, est nécessaire au respect de la première clause. De fait, si *l'explanans* ne contenait pas de lois statistiques, la forme logique de cette explication ne serait pas adéquate et, conséquemment, on ne pourrait pas effectuer ce raisonnement inductif. La troisième condition est aussi respectée : tous les énoncés de *l'explanans* sont empiriquement vérifiables. La quatrième condition est de même observée. En effet, il est parfaitement plausible de penser qu'il est vrai que (a) Pierre a été en contact avec une

personne atteinte de rougeole et que (b) la probabilité que des gens en contact avec une personne atteinte de rougeole l'attrapent est élevée. Enfin, on peut supposer que l'exigence de spécificité maximale est prise en compte parce qu'il n'y a pas d'autres faits pertinents que nous connaissons (qu'il faudrait citer dans *l'explanans*) et qui, par voie de conséquence, altéreraient la loi statistique en présence. Dans cet exemple, on suppose qu'il n'est pas le cas, par exemple, que la personne atteinte de rougeole et avec laquelle Jean fut en contact avait une infection déclinante et très peu contagieuse. Si tel avait été le cas, *et si on l'avait su*, afin de respecter l'exigence de spécificité maximale, on n'aurait pas eu le droit de formuler de cette façon cet argument inductif. Il aurait fallu, en effet, tenir compte du fait que la personne avec laquelle Jean fut en contact avait une rougeole peu contagieuse et, en conséquence, formuler une loi statistique selon laquelle la probabilité que des gens en contact avec une personne atteinte d'une rougeole déclinante ne *contractent pas* la rougeole est élevée.

Jusqu'ici, j'ai exposé la conception syntaxique des théories scientifiques dans le cadre de laquelle il faut concevoir l'explication selon les empiristes logiques. Ensuite, j'ai présenté et montré sommairement comment fonctionnent les modèles DN et IS Hempelien. J'ai alors appliqué ces modèles à des *phénomènes naturels* à expliquer. J'ai aussi fait voir que ces modèles contenaient une *théorie de la causalité* qui mettait de l'avant l'idée de Salmon de prévisibilité nomique selon laquelle des lois universelles ou statistiques « connectent » *l'explanandum* avec certitude ou avec une grande probabilité à d'autres événements décrits dans *l'explanans*. Dans la prochaine section, je souhaite faire voir que le modèle DN peut avoir une autre fonction, à savoir celle d'expliquer les lois et les théories scientifiques. J'introduirai aussi, dans cette section, mon interprétation réductionniste des modèles de l'explication DN et IS appliqués aux phénomènes naturels.

1.3 L'explication des lois et théories scientifiques selon le modèle DN et mon interprétation réductionniste des modèles DN et IS appliqués aux phénomènes naturels

Le modèle DN de Hempel (et Oppenheim) peut être utilisé pour expliquer autre chose que des phénomènes naturels produits par des relations causales. De fait, on peut utiliser le modèle DN pour expliquer les *lois de la nature* (Hempel et Oppenheim 1948, 136) et, corollairement, les *théories* et les *disciplines*. Expliquer une théorie consistera alors à montrer qu'il est possible de déduire les principes et les lois de la théorie que l'on veut expliquer (*l'explanandum*) des principes et des lois de la théorie qui explique (*l'explanans*). Si on réussit cette entreprise explicative, on dira que nous avons *réduit* une théorie moins fondamentale T_H dite « de haut niveau » à l'aide d'une théorie plus fondamentale T_B dite « de bas niveau », la théorie T_B étant jugée plus fondamentale que la théorie T_H dans la mesure où T_B est capable de prévoir plus de phénomènes que T_H , et ce, à partir d'un plus petit nombre de faits observables que T_H (ibidem, 164)³⁰.

Chez les tenants de l'empirisme logique, lesquels raisonnent selon la conception syntaxique des théories scientifiques (section 1.1), cette façon de s'imaginer l'explication des théories par voie déductive et réductive est porteuse de l'idéal selon lequel plus nous réussissons à réduire les théories $T_{H1}...T_{Hn}$ par les théories $T_{B1}...T_{Bn}$, plus nous faisons un progrès en science. Nous faisons une avancée parce que, selon les représentants de l'empirisme logique, quand nous réussissons cette entreprise, nous trouvons des théories $T_{B1}...T_{Bn}$ ayant un pouvoir explicatif supérieur à celui que

³⁰ Évidemment, dans ce cas, ces explications ne seront pas de nature causale. De fait, comme le dit Hempel, « [...] the explanation of a general law by deductive subsumption under theoretical principles is clearly not an explanation by causes » (Hempel 1965c, 352). De fait, il serait inadéquat d'affirmer, lorsqu'on réduit/explique une théorie T_H (*l'explanandum*) à/par une théorie T_B (*l'explanans*), que T_B cause T_H .

présentent les théories réduites $T_{H1...T_{Hn}}$.³¹ Pour les empiristes logiques, le but ultime, dans cette entreprise, serait de réduire de manière déductive-nomologique et selon une logique transitive toutes les théories de haut niveau $T_{H1...T_{Hn}}$ à la microphysique T_B . Par exemple, comme le fait remarquer Jonathan Opie, si la biologie cellulaire est réductible à la chimie, et que la chimie est réductible à la physique atomique, alors, en vertu du principe de transitivité, la biologie cellulaire est réductible à la physique atomique (Opie 2010, 271). Putnam et Oppenheim ont explicité cet idéal réductionniste logico-syntaxique ou, d'un autre point de vue, cette hypothèse de travail, dans leur article célèbre « Unity of Science as a Working Hypothesis » (1958)³².

Dans ce qui suit, je présente cette idée d'explication des théories par voie déductive/réductive telle qu'Ernest Nagel, suite à Hempel et Oppenheim, l'a conçue. Je fais connaître ce concept de réduction interthéorique parce que, selon moi, si l'exercice d'opérer une telle réduction est plausible et applicable, alors cela implique que nous devons adopter une *conception réductionniste* des modèles DN et IS de

³¹ Hempel, à mon avis, adopte cet optimisme lié à l'idéal réductionniste quand, pour la biologie, il vante les mérites d'une heuristique mécaniste et montre la stérilité scientifique des doctrines vitalistes. De fait, il affirme que

[...][le mécanisme] enjoint à l'homme de science de persister dans sa recherche de théories physico-chimiques fondamentales des phénomènes biologiques, au lieu de se résigner à une conception d'après laquelle les concepts et les principes de la physique et de la chimie seraient incapables de rendre compte de façon satisfaisante des phénomènes de la vie. (Hempel 1972, 165)

De plus, poursuit-il, « [l']adhésion à cette maxime [heuristique] s'est certainement révélée très fructueuse dans la recherche biophysique et biochimique — et c'est là un titre auquel ne peut prétendre la conception vitaliste » (ibidem, 165).

³² Même si le projet réductionniste d'Oppenheim/Putnam (1958) ne procède pas par réduction interthéorique, mais par réduction méréologique, il me semble que les propos de leur article expriment le cœur de cet idéal réductionniste (ou, du moins, il a été interprété ainsi). Si ce n'avait pas été le cas, d'ailleurs, Jerry Fodor n'aurait pas pris pour cible cet article quand, dans son texte de 1974 « Special Sciences : Or the Disunity of Science as a Working Hypothesis », il critique la réduction interthéorique.

l'explication des *phénomènes naturels*. Je m'explique : comme je viens de le noter, plus une théorie est de bas niveau, plus elle est fondamentale et, par voie de conséquence, les lois qu'elle contient fourniront, en principe, de meilleures explications que les lois provenant de théories de haut niveau. Conséquemment, me semble-t-il, si le projet réductionniste est faisable, alors une explication DN ou IS appliquée aux *phénomènes naturels* devra être *réductionniste*, c'est-à-dire qu'elle ne devra contenir, dans *l'explanans*, que des lois de bas niveau $L_{B1}...L_{Bn}$ tirées d'une théorie de bas niveau T_B . En psychiatrie, ce type d'explication réductionniste DN ou IS, s'il est plausible, ne devrait contenir que des lois neurobiologiques (et, idéalement, biomoléculaires) et, par voie de conséquence, en vertu de la conception de la causalité inhérente aux modèles DN et IS (voir section 1.2), les seules relations causales que l'on devrait retenir opéreraient au niveau neurobiologique.

Cette interprétation réductionniste des modèles DN ou IS appliquée aux *phénomènes naturels* est-elle viable? Après avoir présenté la conception de la réduction d'Ernest Nagel, je tenterai de montrer, au deuxième chapitre, qu'elle ne l'est pas. De fait, nous verrons, entre autres, que la réduction de la psychiatrie par la neurobiologie, dans ce cadre syntaxique des théories scientifiques, n'est pas possible et, ipso facto, que l'interprétation réductionniste des modèles DN et IS appliqués aux *phénomènes naturels* que je viens de présenter ne l'est pas non plus.

1.4 La réduction selon Ernest Nagel³³

Au chapitre XI de son livre *The Structure of Science* (1961), Ernest Nagel a construit, dans le cadre de la conception syntaxique des théories scientifiques (voir section 1.1), un modèle de la réduction interthéorique que l'on pourrait qualifier de « DN ». En effet, ce philosophe des sciences, quoique sans y faire explicitement référence, développe les idées hempelienues et oppenheimiennes concernant la réduction/explication des théories et des lois dont j'ai brièvement parlées dans la section précédente (section 1.3).

1.4.1 Les modèles de Hempel/Oppenheim et Nagel sont semblables

Hempel, Oppenheim et Nagel partagent plusieurs idées en regard de l'explication des théories et du réductionnisme.

D'abord, disons que, pour Nagel, comme pour Hempel et Oppenheim (section 1.3), la réduction interthéorique ou l'entreprise consistant à réduire une théorie T_H par une théorie T_B est un idéal qui guide les scientifiques dans leur recherche d'une science dont les principes unifieraient toutes les disciplines scientifiques. Comprise de cette façon, la réduction de théories par d'autres théories témoigne du progrès de la science dans l'histoire. De fait, Nagel écrit que

³³ Pour construire et structurer cette présentation du modèle de Nagel, en plus de m'être basé sur les textes classiques de Nagel à ce sujet (surtout Nagel 1961, chap.11 et Nagel 1970), je me suis inspiré de plusieurs résumés faits par des philosophes à ce sujet. Les exposés du modèle nagelien de Robert Batterman (2006; 2012, section 1) et de Max Kistler (2007, la meilleure présentation que j'aie pu trouver), par exemple, m'ont particulièrement influencé dans l'articulation de ma réflexion.

[...] [l'idéal de la réduction] continues to leaven current scientific speculation; and, in any case, the phenomenon of relatively autonomous theory becoming absorbed by, or reduced to, some other more inclusive theory is an undeniable and recurrent feature of the history of modern science. (Nagel 1961, 336-337)

Par ailleurs, comme Hempel et Oppenheim (section 1.3), Nagel considère la *réduction* des théories comme une *explication* où une théorie T_H est expliquée par une théorie plus fondamentale T_B . En effet, il écrit,

Reduction [...] is the explanation of a theory or a set of experimental laws established in one area of inquiry, by a theory usually though not invariably formulated for some other domain (ibidem, 338, mes italiques)

Enfin, comme dans le modèle de Hempel et Oppenheim, la réduction/explication nagelienne est de nature épistémique et linguistique, c'est-à-dire qu'elle consiste en une mise en relation linguistique et logique de T_H et T_B et non en une mise en relation des entités du monde auxquelles réfèrent ces théories. En d'autres termes, l'explication/réduction interthéorique, pour Nagel, n'est pas de nature ontologique. Ce dernier soutient une telle thèse parce qu'il pense, comme les tenants de l'empirisme logique, que nous n'avons pas accès à la nature des constituants élémentaires des choses (voir section 1.1, point 1) et, par voie de conséquence, à ce que ces dernières impliquent ou n'impliquent pas. De fait, il pose que

[...] the «natures» of things, and in particular of the «elementary constituents» of things, are not accessible to direct inspection and [...] we cannot read of by simple inspection what it is they do or do not imply. (ibidem, 364)

Pour cette raison, Nagel pense qu'une réduction/explication de type ontologique serait hautement spéculative et, pour tout dire, impraticable. À ce sujet, il écrit,

[...] the supposition that, in order to reduce one science to another, some properties must be deduced from certain other properties or «nature» converts what is eminently a logical and empirical question *into a hopelessly irresolvable speculative one.* (ibidem, 365, *mes italiques*)

Bref, comme l'affirme Sahotra Sarkar, «reduction is for Nagel purely an epistemological issue with no necessary ontological commitment» (Sarkar 1992, 172-173). Comme nous le verrons plus loin (section 2.4), certains philosophes, comme Kenneth F. Schaffner, soutiendront, contrairement à Nagel, qu'une réduction implique à l'occasion une « simplification ontologique » (Schaffner 1967, note 9; 1969a, 285)³⁴. Pour l'instant, toutefois, je ne m'attarde pas davantage sur cette question, mais continue plutôt d'examiner le modèle nagelien de la réduction interthéorique.

1.4.2 Le modèle interthéorique de la réduction nagelienne et ses concepts

Nous avons vu dans la section précédente que la réduction/explication nagelienne est de nature épistémique, qu'elle ne requiert aucun engagement ontologique et qu'elle consiste en une heuristique idéale en science. Dans ce qui suit, je fais connaître les concepts que contient son modèle de manière plus détaillée. Commençons par voir en quoi consiste celui que l'on nomme « énoncé de balises physiques » (« statement of boundary conditions »). J'exposerai ensuite les autres notions nageliennes relatives à la réduction interthéorique.

³⁴ En cela, on peut dire que, bien que son modèle de la réduction des années 1960 et 1970 que je présenterai à la section 2.4 s'inscrive dans la tradition de l'empirisme logique, Schaffner se distancie de ce cadre de pensée en introduisant cette thèse ontologique.

1.4.3 Les énoncés de balises physiques

Comme nous l'avons vu, il est communément admis par les philosophes des sciences qu'il existe des théories de bas niveau (T_B) et des théories de haut niveau (T_H) (voir section 1.3). De plus, comme je l'ai déjà mentionné, on peut dire que les théories de haut niveau $T_{H1}...T_{Hn}$ sont celles que l'on veut réduire/expliciter et que les théories de bas niveau $T_{B1}...T_{Bn}$ sont celles qui réduisent/expliciter les théories $T_{H1}... T_{Hn}$. De plus, on peut affirmer que les théories de bas niveau $T_{B1}...T_{Bn}$ capturent un plus grand éventail de phénomènes physiques que les théories de haut niveau $T_{H1}... T_{Hn}$.³⁵ Pour cela, quand on voudra réduire T_H par T_B , il faudra tenir compte des *conditions physiques* nécessaires à l'avènement des *phénomènes physiques* que capturent les théories de haut niveau $T_{H1}...T_{Hn}$ par des *énoncés représentant ces conditions physiques*. William Bechtel, suivant Nagel, affirme qu'il faut référer à ces conditions physiques par des *énoncés de balises physiques* ou, en anglais, par des « statements of boundary conditions » (Bechtel et Hamilton 2007, 385). En postulant ces conditions physiques à l'aide de ces énoncés — et c'est ce qu'on doit faire quand on opère une réduction — on circonscrit ou balisera (d'où le terme anglais « boundary ») l'application de T_B au domaine de référence de T_H . Par exemple, comme l'explique Hempel et Oppenheim (1948, 136), si nous voulons réduire la théorie galiléenne du mouvement des corps en chute libre T_H par la théorie de la gravitation universelle de Newton T_B , il faudra postuler le référentiel terrestre parce que T_H ne peut que décrire la trajectoire des objets évoluant sur la Terre contrairement à T_B qui s'applique non seulement aux mouvements des objets sur la Terre, mais aussi à la trajectoire de la Terre elle-même, aux planètes et aux étoiles.

³⁵ En ce sens, suivant la définition hempelienne/oppenheimienne que j'ai présentée à la section 1.3 où j'ai défini ce qu'était le pouvoir explicatif des théories, on peut affirmer que les théories de bas niveau, parce qu'elles expliquent davantage de phénomènes, sont plus fondamentales que les théories de haut niveau.

Dans la prochaine section, j'explique en quoi, pour Nagel, consiste une *réduction homogène*. Ensuite, je présente la *réduction hétérogène* et les *hypothèses de liaison*, lesquelles sont des nouveautés nageliennes relativement au modèle hempelien et oppenheimien. Enfin, avant de clore ce chapitre, j'analyserai les *conditions informelles* et *formelles* que suppose le modèle de la réduction de Nagel.

1.4.4 Les deux types de réduction et les hypothèses de liaison

Nagel pense qu'il existe deux types de réduction, soit la réduction homogène et la réduction hétérogène. La réduction homogène est la plus simple des deux. Effectuer une réduction de ce genre est possible lorsque T_H et T_B présentent un vocabulaire similaire ou identique et que l'on peut déduire directement de T_B la théorie T_H . L'exemple que nous venons de voir de Hempel et Oppenheim concernant la réduction de la théorie galiléenne du mouvement des corps en chute libre (T_H) par la théorie de la gravitation universelle de Newton (T_B) est une réduction de ce genre parce que, comme l'écrit Nagel, « [...] the laws of the secondary science [T_H] employ no descriptive terms that are not also used with approximately the same meaning in the primary science [T_B] » (Nagel 1961, 339). Bref, comme les termes de T_H et T_B sont homogènes, la réduction de T_H par T_B l'est aussi.

Comme on peut s'en douter, cependant, il se peut que la théorie T_H que l'on souhaite réduire/expliciter contienne des termes qui réfèrent à des entités ou des propriétés qui n'apparaissent pas dans la théorie réductrice T_B . En ce cas, la réduction de T_H par T_B sera qualifiée d'*hétérogène* par Nagel. La réduction de la thermodynamique qui contient des termes comme « température », « pression » et « entropie » par la mécanique statistique qui ne contient pas ces termes est l'exemple

— maintenant classique — dont se sert Nagel pour illustrer ce type de réduction. Ultimement, ce dernier montre que, dans le cadre de la conception syntaxique des théories scientifiques (voir section 1.1, surtout les points 10 à 12), l'on peut déduire la loi de Boyle-Mariotte — présente dans la thermodynamique (T_H) — de la théorie cinétique des gaz — présente dans la mécanique statistique (T_B) — en formulant des *règles de correspondance* (voir section 1.1, points 8b et 9 b) qui relient les *termes observationnels* (voir section 1.1, point 8c) de la thermodynamique que l'on ne retrouve pas dans la mécanique statistique (par exemple, celui référant à la température) avec des *termes théoriques* (voir section 1.1, point 8 b) de la mécanique statistique (par exemple, celui référant à l'énergie cinétique moléculaire).

Nagel illustre cette réduction par une expérience de pensée intéressante (Nagel 1961, 343-345). Premièrement, ce dernier postule un certain nombre *d'énoncés de balises physiques* (voir section 1.4.3). Il suppose, en effet, qu'un *gaz idéal* est contenu dans un contenant de volume V . Un *gaz idéal* consiste en une vue de l'esprit — ou, plus précisément, en un modèle thermodynamique — et se comporte conformément à la loi ou l'équation empirique de Boyle-Mariotte $pV = kT$.³⁶ Le *gaz idéal* dont cette loi rend compte est un gaz qui présente le comportement général que tous les *gaz réels* tendent à avoir à *basse pression* et à *température élevée*. De fait, la relation que décrit la loi de Boyle-Mariotte $pV = kT$ entre la pression P , le volume V et la température T est *indépendante de la nature du gaz* (Wikipédia 2011).

Deuxièmement, parce que Nagel veut faire un lien entre (a) le concept référant au phénomène macroscopique qu'est *la pression* (qu'il identifiera éventuellement à la température) que décrit la thermodynamique et (b) les *principes de la mécanique*, il

³⁶ Cette équation se lit de la manière suivante : « le produit de la pression et du volume d'un gaz est proportionnel à sa température » (k est la constante de Boltzmann que l'on multiplie avec la quantité de température : la température est donnée en Kelvin [d'où « K »]).

fait l'hypothèse que ce gaz parfait est composé d'un nombre gigantesque de molécules et que ces dernières sont sphériques, parfaitement élastiques et éloignées les unes des autres. De postuler qu'elles sont éloignées les unes des autres lui permet de ne pas considérer les interactions ou forces électrostatiques (qui dépendent justement de la nature des gaz) et d'affirmer que « [...] les seules forces qui s'exercent sur elles sont dues aux chocs entre molécules et aux chocs avec les parois du récipient » (Kistler 2007, 9). Enfin, Nagel postule que ces molécules sont dans un système isolé (Nagel 1961, 344). Ayant formulé ces prémisses, Nagel peut légitimement prétendre expliquer les grandeurs thermodynamiques (pression, température) par les comportements des molécules décrits par les lois de la mécanique.

Cependant, comme les molécules du gaz parfait sont en nombres astronomiques, les coordonnées de vitesse et de position de chacune des molécules ne peuvent pas être données de manière exacte (ibidem, 344). C'est pourquoi, explique Nagel, il faut introduire une *hypothèse statistique* selon laquelle « [...] les vitesses et les positions des molécules sont distribuées de manière uniforme sur l'espace des vitesses et des positions possibles » (Kistler 2007, 9) d'un espace abstrait — décrit en langage mathématique — appelé « espace de phase » qui est découpé en « [...] un très grand nombre de “cellules de phase”, toutes de la même taille » (ibidem, 9). Ensuite, comme l'explique Kistler,

[l]'hypothèse statistique cruciale consiste à supposer que la probabilité qu'une molécule donnée occupe une cellule donnée est la même pour toutes les cellules, et que cette probabilité est indépendante de l'état des autres molécules. (ibidem, 9)

Nagel fait cette hypothèse, je suppose, pour s'éviter, dans ce modèle statistique, d'avoir à considérer les forces électrostatiques et de ne tenir compte que des forces

cinétiques. Ainsi, dans ce cadre théorique probabiliste, les lois de la mécanique seront les seules à déterminer les mouvements des molécules. Enfin, de dire toujours Kistler,

[g]râce à cette hypothèse, jointe aux lois de la mécanique, *on peut déduire que la pression macroscopique est une fonction de l'énergie cinétique moyenne des molécules $E : p = 2E/3V$. (ibidem, 9, mes italiques)*

Après avoir formulé l'équation $p = 2E/3V$ qui, au fond, est un théorème déduit des principes de la mécanique statistique, Nagel fait le raisonnement suivant.

L'équation $p = 2E/3V$, fait-il remarquer, peut être réécrite de la manière suivante : $pV = 2E/3$ (Nagel 1961, 344). Or, si on compare l'équation³⁷ $pV = 2E/3$ avec la loi de Boyle-Mariotte $pV = kT$, on constate que ces dernières expriment la même chose (mais différemment) et que, pour cette raison, on peut faire l'hypothèse, pense-t-il, que la loi de Boyle-Mariotte pourrait être déduite d'une quelconque façon de $pV = 2E/3$ si l'on pouvait *relier* la température avec l'énergie cinétique moyenne des molécules (ibidem, 344). C'est pourquoi Nagel, en faisant un raisonnement logico-transitif, se permet d'introduire le postulat $kT = 2E/3$ selon lequel — en mot — « [...] la valeur numérique de la température T est identique à une fonction de l'énergie moléculaire moyenne [...] » (Kistler 2007, 9, mes italiques).

Nagel appellera l'équation $kT = 2E/3$ « *hypothèse de liaisons* » (ou, et ce sont des concepts identiques, « règle de correspondance », « principe pont » ou « loi pont ») parce qu'elle met en relation le *terme observationnel* « T » de la thermodynamique avec le *terme théorique* « $2E/3$ » de la mécanique statistique. C'est grâce à ces hypothèses de liaison que Nagel pense que nous pourrions effectuer une réduction de

³⁷ Cette équation signifie que le produit de la pression (p) et du volume (V) d'un gaz est proportionnel à l'énergie cinétique moyenne des molécules ($2E/3$) qui composent ce gaz.

T_H par T_B . Selon Nagel, on peut concevoir ces règles de correspondance de plusieurs manières. Premièrement, ces principes ponts peuvent être appréhendés comme des *énoncés analytiques* (Nagel 1961, 354), c'est-à-dire comme des énoncés dont on peut déduire la vérité essentiellement en faisant l'analyse de la signification des termes qu'ils contiennent. Deuxièmement, on peut considérer les règles de correspondance comme des énoncés exprimant des *conventions linguistiques ou conceptuelles* où sont mis en relation *de manière arbitraire* un terme observationnel de la théorie à réduire T_H et un terme théorique de la théorie réductrice T_B (ibidem, 354). Troisièmement, les hypothèses de liaison peuvent être interprétées comme des *hypothèses synthétiques ou factuelles* ou en anglais des « physical hypotheses » (Nagel 1970, 125-128; Nagel 1961, 354-355) ayant soit la forme d'un énoncé conditionnel [$\forall x (Ax \rightarrow Bx)$] ou celle d'un énoncé biconditionnel [$\forall x (Ax \leftrightarrow Bx)$] (Nagel 1961, 354-355 et 355, note 5). En dernière instance, pour l'exercice de la réduction interthéorique, Nagel rejette la première et la seconde interprétation et choisit d'adopter la conception selon laquelle les hypothèses de liaison sont des *hypothèses synthétiques*. Cependant, Nagel n'a pas été clair quant à la forme logique que devaient présenter ces *hypothèses synthétiques*. En effet, bien que ce dernier ait dit que seulement les énoncés de forme biconditionnelle, d'un point de vue logique, assuraient la réduction (Nagel 1961, 355 note 5)³⁸, il n'a pas rejeté la pertinence des énoncés conditionnels³⁹. Plusieurs

³⁸ Dans une note de bas de page maintenant célèbre pour ceux qui travaillent le concept de réduction (Nagel 1961, 355 note 5), Nagel introduit l'idée selon laquelle seulement une hypothèse de liaison ayant la forme d'un énoncé biconditionnel assure la déduction ou la réduction de T_H de T_B ou, et c'est la même chose, assure le respect de la *condition de la déductibilité*. En effet, note-t-il, d'un point de vue *strictement logique*, si pour tous les termes A de la théorie T_H que l'on veut réduire il existe un terme B auquel on peut relier ces termes dans la théorie réductrice T_B et si ces termes sont reliés par un lien biconditionnel (c'est-à-dire si on a un lien de la forme « A si et seulement si B » ou, en notation logique, $A \leftrightarrow B$), alors, par voie de conséquence, A peut être remplacé par B dans n'importe laquelle des lois L de la théorie T_H que l'on veut réduire et où on retrouve A . En d'autres termes, si nous avons une loi $A_1 \rightarrow A_2$ de la théorie T_H que l'on veut réduire et qu'on peut relier les termes A_1 et A_2 de cette loi de manière biconditionnelle aux termes B_1 et B_2 de la loi $B_1 \rightarrow B_2$ de la théorie de bas niveau T_B , on pourra toujours réduire $A_1 \rightarrow A_2$ par $B_1 \rightarrow B_2$. De fait, si nous avons les hypothèses de liaison $A_1 \leftrightarrow B_1$ et $A_2 \leftrightarrow B_2$, il s'en suit que, dans tous les cas, selon un raisonnement transitif, $A_1 \rightarrow A_2$ est réductible à $B_1 \rightarrow B_2$. Donc, comme on peut le constater, l'hypothèse de liaison, quand elle a la forme d'un énoncé biconditionnel, garantit la réduction. Cependant, comme l'explique Nagel, le lien entre les termes A et

philosophes, comme Kenneth F. Schaffner et Lawrence Sklar, toutefois, contrairement à Nagel, ont soutenu que le seul type pertinent d'hypothèse de liaison pour effectuer une réduction interthéorique est celui ayant la forme d'un énoncé biconditionnel, c'est-à-dire celui qui exprime une *relation d'identité* entre deux *prédicats* (Schaffner 1967, 144; Sklar 1967, 124) ou entre deux *classes* d'entités (Schaffner 1967, 144; Sklar 1967, 120). Selon cette conception des hypothèses de liaison, par exemple, si nous avons une règle de correspondance $\forall x ((\text{est dans un état de peur})x \leftrightarrow (\text{présente une propriété physique } P \text{ activée}) x]$, alors l'activation de P et le fait que x ait lieu sont une seule et même chose de la même façon que $2+2=4$ ou que l'eau $=\text{H}_2\text{O}$. Pour le reste de ce travail, j'adopterai, comme Schaffner et Sklar, cette dernière interprétation qui, nous le verrons à la section 2.5.1, pose un problème sérieux — celui de la réalisation multiple — à l'exercice de la réduction interthéorique.

B n'a pas toujours la forme d'un énoncé biconditionnel. En effet, ce lien peut avoir la forme d'un énoncé conditionnel. Dans ce cas, on aurait un énoncé de la forme « si A , alors B ». Or, comme le souligne Nagel, dans cette éventualité, A n'est pas remplaçable par B . En effet, $A \rightarrow B$ n'implique pas que $B \rightarrow A$. Conséquemment, la théorie T_H que l'on veut réduire ne pourra pas être déduite de la théorie réductrice T_B . En somme, comme on peut le voir, le respect de la *condition de la connectivité* n'assure pas le respect de la *condition de la déductibilité* (ces deux notions sont définies dans la section 1.4.5). (Curieusement, l'exemple paradigmatique de Nagel pour exemplifier l'exercice de la réduction, à savoir l'hypothèse de liaison $kT = 2E/3$, ne présente pas la forme d'un énoncé biconditionnel, mais plutôt celle d'un énoncé conditionnel. En effet, la science d'aujourd'hui nous montre que le *prédictat* « température » ne peut pas être relié de manière biconditionnelle au *prédictat* « énergie cinétique moléculaire moyenne » parce que, entre autres, « [...] il existe des objets qui ont une température mais dans lesquels la température ne se réduit pas à l'énergie cinétique moléculaire » [Kistler 2007, 13]. Par exemple, la température « [...] d'un plasma ne peut être définie en termes d'énergie moléculaire puisqu'un plasma ne contient pas de molécules, mais seulement des atomes ionisés » [Poirier 2000, 48]).

³⁹ Nagel, en effet, dans son article de 1970, semble toujours penser que les hypothèses de liaison qui ont la forme d'un *énoncé conditionnel* doivent être considérées dans l'exercice de la réduction. De fait, il écrit, en parlant des règles de correspondance mettant en lien deux *prédicats* :

...bridge laws of the sort being considered may specify conditions for the occurrence of an attribute which are necessary as well as sufficient; in other cases the conditions specified may be sufficient without being necessary; and in still other cases, the conditions stated may be only necessary. (Nagel 1970, 126, mes italiques)

Dans la prochaine partie, avant de résumer ce qu'est le modèle nagelien et clore ce chapitre, je complète ma présentation du modèle nagelien en faisant connaître les deux *conditions formelles* et les trois *conditions informelles* qu'a formulées Nagel et que l'on doit respecter lorsqu'on fait une réduction hétérogène (et homogène). Je poursuivrai ensuite au second chapitre en montrant qu'il n'est pas possible (1) d'effectuer la réduction hétérogène mettant en jeu la psychiatrie T_H et la neurobiologie T_B et, conséquemment, (2) de donner essentiellement des explications réductionnistes DN ou IS des maladies en médecine mentale.

1.4.5 Les deux conditions formelles de la réduction hétérogène

Suivant l'expérience de pensée de Nagel que j'ai présentée précédemment (section 1.4.4) concernant la réduction de la thermodynamique par la mécanique statistique, ce dernier formule deux *conditions formelles* que l'on doit respecter lorsqu'on effectue une réduction hétérogène, à savoir la *condition de la connectivité* et la *condition de la déductibilité* (Nagel 1961, 353-354).

Nagel introduit la *condition de la connectivité* de la manière suivante :

(1) Assumptions of some kind must be introduced which postulate suitable relations between whatever is signified by «A» [«A» étant un terme présent dans T_H mais pas dans T_B] and traits represented by theoretical terms already present in the primary science [T_B]. (Nagel 1961, 353-354)

Comme nous l'avons vu dans l'exemple de la réduction de la thermodynamique (T_H) à la mécanique statistique (T_B), la théorie réduite T_H contient des termes qui ne sont pas compris par la théorie réductrice T_B . C'est pourquoi il faut introduire, comme l'a fait Nagel, des *hypothèses de liaison* qui relient un terme de T_H à un terme de T_B .

Comme nous l'avons vu, dans l'exemple classique de Nagel, une de ces hypothèses de liaison est l'équation $kT = 2E/3$.

Après avoir traité de la *condition de connectivité*, Nagel introduit la *condition de déductibilité* dans le passage suivant :

(2) With the help of these additional assumptions [les hypothèses de liaison], all the laws of the secondary science [T_H], including those containing the term 'A' [par exemple, le terme «température»], must be logically derivable from the theoretical premises and their associated coordinating definitions in the primary discipline [T_B]. (ibidem, 354)

Autrement dit, dans ce passage, Nagel affirme que l'on doit pouvoir déduire logiquement les lois et postulats de T_H des lois et postulats de T_B augmentée des hypothèses de liaisons. Dans notre exemple, cela signifierait que $pV = kT$ peut être déduit de $pV = 2E/3$ et $kT = 2E/3$.⁴⁰

Dans la prochaine section, avant de résumer en quoi consiste le modèle nagelien, je présente les trois conditions informelles que l'on doit respecter lorsqu'on opère une réduction interthéorique.

⁴⁰ On doit ajouter que cette déduction doit tenir compte des *énoncés de balises physiques* décrivant les conditions physiques pour laquelle la loi de Boyle-Mariotte fonctionne, à savoir, par exemple, les énoncés faisant référence à une température élevée et à une base pression (voir section 1.4.4).

1.4.6 Les trois conditions informelles

Nagel a formulé des *conditions informelles* dont on doit tenir compte quand l'on fait une réduction hétérogène (et également homogène, selon moi) (Nagel 1961, 358-364). Toutefois, Nagel, me semble-t-il, ne les a pas énumérées clairement. À mon avis, il y en a au moins trois. La première consiste en l'idée selon laquelle les hypothèses théoriques de la science primaire ou réductrice doivent être supportées par des données empiriques qui possèdent un certain degré de pertinence (ibidem, 358). Deuxièmement, je suppose que, étant donné la première condition, il s'en suit une seconde condition selon laquelle la science secondaire ou à réduire doit elle aussi être supportée par des preuves empiriques et avoir un intérêt heuristique et scientifique. En effet, il ne vaudrait pas la peine de vouloir réduire une théorie T_H qui, à l'époque où on tente d'effectuer une réduction, n'a aucun intérêt scientifique (disons, au 21^e siècle, la physique aristotélicienne ou la doctrine de l'harmonie préétablie de Leibniz⁴¹) à une théorie T_B qui est scientifiquement pertinente (disons les théories cinétiques de la physique actuelle ou les neurosciences de la conscience du 21^e siècle). Une des raisons qui me font penser que Nagel juge que la science secondaire est scientifiquement digne d'intérêt et empiriquement justifiée est son idée selon laquelle la science primaire doit aider la science secondaire à se développer et enrichir les lois qui composent cette dernière (Nagel 1961, 360). Or, je ne vois pas comment, par exemple, on pourrait enrichir les théories dualistes classiques concernant la relation corps/esprit par les neurosciences d'aujourd'hui. Mais je conçois parfaitement que les sciences cognitives de la mémoire puissent être enrichies par les recherches en

⁴¹ Je prends cet exemple de Jaegwon Kim (2000, 17) pour montrer que ce dernier ignore cette condition informelle nagelienne quand il affirme que « [...] la réductibilité nagelienne du mental est non seulement compatible avec ces théories dualistes, mais qu'elle est en fait impliquée par plusieurs d'entre elles » (ibidem, 17). Je ne pense pas, de fait, suivant ces conditions informelles, que Nagel aurait jugé pertinent de tenter de réduire la théorie leibnizienne de l'esprit ou toutes autres formes de théories dualistes par les neurosciences d'aujourd'hui.

neurobiologie de la mémoire ou, dans le cas qui nous préoccupe présentement, que la thermodynamique puisse être complétée par la mécanique statistique (voir Sklar 1999 à ce sujet). Enfin, il y a la troisième condition qui vaut aussi pour la réduction homogène et selon laquelle la théorie de bas niveau doit être plus fondamentale que la théorie de haut niveau. En d'autres termes, comme nous l'avons vu (section 1.3), la théorie de bas niveau doit être capable d'expliquer et prédire un plus grand nombre de phénomènes que la théorie de haut niveau, et ce, selon un plus petit nombre de données que cette théorie de haut niveau (Hempel et Oppenheim 1948, 164). Cette clause est nécessaire parce qu'elle indique la direction de la réduction. Par exemple, on réduit la théorie galiléenne du mouvement des corps en chute libre à la mécanique newtonienne et non l'inverse.

Dans la prochaine section, je résume les idées nagéliennes concernant la réduction interthéorique.

1.4.7 Résumé des idées nagéliennes sur la réduction

En somme, si on respecte les *conditions informelles* formulées par Nagel, on peut affirmer que la réduction d'une théorie T_H (l'*explanandum*) par une théorie T_B (l'*explanans*) est faisable (a) lorsqu'il est possible, en tenant compte des *énoncés de balises physiques* (« statement of boundary conditions »), de déduire toutes les lois et principes de T_H de celles et ceux de T_B (il s'agit de la *condition de la déductibilité*) après que — si c'est une réduction hétérogène — (b) les termes apparaissant dans la théorie à déduire (T_H) soient reformulés selon le langage de la théorie réductrice (T_B) (il s'agit de la *condition de la connectivité*) (Sarkar 2008, 426).

Pour finir, suivant William Bechtel (2008, 131), j'illustre schématiquement le modèle de la réduction interthéorique nagelien :

Lois de bas niveau L_B ou théorie T_B de bas niveau (Loi que l'on retrouve dans la science fondamentale qui réduit) [**explanans**]

Hypothèses de liaison [**explanans**]

Les conditions physiques — décrites par les *énoncés de balises physiques* — pour lesquelles la loi L_H ou la théorie T_H s'applique (« statement of boundary condition ») [**explanans**]

Lois de haut niveau L_H ou théorie T_H de haut niveau (Loi que l'on retrouve dans la science que l'on veut réduire) [**explanandum**]

Comme on peut le constater, à la lumière de ce schéma et de ce que nous venons de voir jusqu'ici, mis à part le fait que les *conditions initiales* (voir section 1.2) soient remplacées par les *conditions physiques* (voir section 1.4.3) et que Nagel ait élaboré l'idée de réduction hétérogène et l'idée des hypothèses de liaison qui lui est associée, il me semble clair, comme l'affirme Jaegwon Kim, et comme Nagel l'a lui-même souligné (1970, 120 et 123), que « [...] le modèle nagelien de la réduction n'est rien d'autre que le modèle hempelien de l'explication scientifique nomologique déductive [(DN)] appliqué aux contextes interthéoriques » (Kim 2006 [1998], 58).

1.5 Conclusion du premier chapitre

Dans ce chapitre, premièrement, j'ai d'abord présenté en quoi consistait la conception syntaxique des théories scientifiques des empiristes logiques (section 1.1). J'ai présenté cette conception parce que les modèles DN et IS de Hempel (et Oppenheim) (section 1.2 et 1.3) et le modèle de la réduction interthéorique de Nagel (section 1.4) s'inscrivent dans ce cadre théorique. J'ai ensuite montré en quoi consistaient les modèles DN et IS de l'explication scientifique (section 1.2 et 1.3). Nous avons vu que ces modèles pouvaient être interprétés de deux façons. En effet, on peut concevoir ces modèles comme (1) une théorie de l'explication des phénomènes empiriques ou comme (2) une théorie de la causalité⁴² (section 1.2). Nous avons vu, aussi, que le modèle DN pouvait être conçu comme un modèle de l'explication des lois et théories scientifiques par voie déductive et réductive (section 1.3). Pour faire suite à cette idée selon laquelle le modèle DN de Hempel (et Oppenheim) s'applique aux théories, j'ai présenté le modèle de la réduction interthéorique de Nagel (section 1.4). En effet, ce dernier a manifestement développé son modèle — maintenant classique — en s'inspirant des idées qu'Hempel et Oppenheim eurent au sujet de l'explication des théories. À la section 1.3, j'ai suggéré que cette conception de la réduction interthéorique adoptée par les tenants de l'empirisme logique commande que nous adoptions une *conception réductionniste* des modèles DN et IS de l'explication des *phénomènes naturels*. Cette *conception réductionniste* des modèles DN et IS est-elle viable? Je pense que non. En fait, je pense que les modèles DN et IS sont intenables de manière générale. Pour cela, je crois que l'on ne peut pas les utiliser en psychiatrie. J'explique pourquoi dans le deuxième chapitre.

⁴² Cette théorie est problématique (voir, par exemple, Bromberger 1966 et Salmon 1971). Je lui préfère la théorie interventionniste de la causalité de James Woodward (2003) que je présenterai aux sections 3.8 et 3.8.1.

CHAPITRE II

LA CONCEPTION SYNTAXIQUE DES THÉORIES SCIENTIFIQUES EST UN CADRE INADÉQUAT POUR FORMULER DES EXPLICATIONS RÉDUCTIONNISTES OU PLURALISTES EN PSYCHIATRIE

2.1 Introduction du deuxième chapitre

Dans le premier chapitre, j'ai présenté les modèles de Hempel/Oppenheim et Nagel parce que, comme je l'ai dit, je veux vérifier, dans ce deuxième chapitre, s'il est possible d'adopter un modèle exclusivement réductionniste de l'explication DN ou IS des *phénomènes naturels*. Comme je l'ai déjà expliqué (voir section 1.3), en effet, dans la mesure où le modèle nagelien et le projet réductionniste des empiristes logiques sont adéquats, je pense que le modèle hempelien de l'explication des *phénomènes naturels* doit nécessairement être un modèle réductionniste. En effet, si le modèle nagelien et l'idéal explicatif réductionniste en *regard des théories* qu'il suppose sont viables, alors cela implique, comme je l'ai déjà dit (a) qu'une explication hempelienne de style DN ou IS en *regard des phénomènes naturels* doit contenir, dans *l'explanans*, pour gagner en pouvoir explicatif et en valeur, des lois de la nature L_1, L_2, \dots, L_r qui soient la plus fondamentale possible et, corollairement (b) que les meilleures explications causales en regard des phénomènes naturels soient celles que nous fournissent ces mêmes lois. En psychiatrie, ces modèles

réductionnistes DN et IS ne présenteraient que des lois neurobiologiques. Est-il possible d'adopter cette interprétation réductionniste des modèles DN et IS? Je pense que non parce que le modèle de la réduction de Nagel et même le modèle néonagelien que Schaffner a élaboré et que je présenterai à la section 2.4 se butent à un problème qui rend l'exercice de la réduction interthéorique impraticable. Afin de bien faire comprendre cette idée, dans les prochaines sections, premièrement, j'expose à quoi ressemblerait, de manière schématique, une réduction interthéorique qui mettrait en jeu la psychiatrie et la neurobiologie (section 2.2). Deuxièmement (section 2.3), je fais connaître deux problèmes découlant des *conditions formelles de connectivité* et de *déductibilité* (vues à la section 1.4.5) que présente le modèle de Nagel. Troisièmement, j'explique comment Schaffner résout ces problèmes (section 2.4). Quatrièmement, je montre pourquoi, en dépit de la solution schaffnerienne, la réduction interthéorique n'est pas possible. Le modèle schaffnerien, en effet, présente un problème insurmontable, à savoir celui de la réalisation multiple (section 2.5). Puis, même si cette réduction interthéorique avait été possible, cinquièmement, je fais voir que les modèles de l'explication DN et IS des *phénomènes naturels*, parce qu'ils présentent eux aussi des problèmes, n'auraient pas pu être utilisés en psychiatrie (sections 2.6.2 et 2.6.3). Enfin, sixièmement, j'explique qu'ultimement l'explication de type DN ou IS ne peut être utilisée en psychiatrie parce qu'elle *représente* pauvrement ce qu'est l'explication (réductionniste ou pluraliste) d'un trouble de santé mentale (section 2.6.4). En dernière instance, ce deuxième chapitre me sert à montrer que la conception syntaxique des théories scientifiques est un cadre inadéquat pour formuler des explications DN ou IS réductionnistes ou pluralistes en psychiatrie.

2.2 La réduction de la psychiatrie par la neurobiologie moléculaire

On pourrait dire que les tenants du pluralisme conçoivent la psychiatrie comme une discipline hybride mettant en relation des notions biopsychosociales et des lois tirées, entre autres, de la biologie humaine, de la psychologie et de la sociologie et qui servent à expliquer les troubles de santé mentale. À l'inverse, comme l'affirme Ian Gold (2009, 506), les tenants du réductionnisme pensent que la psychiatrie devrait être identifiée à une théorie des troubles de santé mentale essentiellement ou en grande partie biologique. Pour cela, explique Gold (ibidem, 506), les défenseurs de l'explication réductionniste en psychiatrie soutiennent qu'il est souhaitable de réduire la médecine mentale (c'est-à-dire toutes les théories psychologiques et sociologiques quelle comprend) à la neurobiologie (spécialement cellulaire et moléculaire) parce que (1) ce faisant, conformément aux idéaux réductionnistes des représentants de l'empirisme logique (voir section 1.3), nous aurons à notre disposition des théories neurobiologiques concernant les troubles de santé mentale qui présenteront une force explicative plus puissante que celle des théories psychiatriques actuelles et parce que (2) grâce à ces connaissances théoriques approfondies et plus fondamentales, nous pourrons élaborer des traitements plus efficaces pour traiter les affections de l'esprit.

Dans le cadre du modèle de la réduction interthéorique nagelien (voir section 1.4), en principe, la réduction de la psychiatrie (T_H) par la neurobiologie (T_B) consisterait à déduire — avec l'aide de règles de correspondance de forme biconditionnelle [$\forall x (Ax \leftrightarrow Bx)$] — toutes les théories de la psychiatrie (T_H) de l'ensemble des théories qui forment la neurobiologie (T_B). Cette réduction présenterait alors la forme suivante (je m'inspire de Gold 2009, 508) :

Ensemble de lois **neurobiologiques** L_B de la neurobiologie (idéalement moléculaire) T_B [explanans]

Diverses **Hypothèses de liaisons** liant les termes de la neurobiologie à ceux de la psychiatrie [explanans]

Les **conditions physiques** — décrites par les *énoncés de balises physiques* — pour lesquelles la loi L_H ou la théorie T_H s'applique (« statement of boundary condition ») [explanans]

Toutes les lois **psychiatriques** L_H de la psychiatrie T_H [explanandum]

Évidemment, pour que cette réduction soit opérable, il faut supposer — même si ce n'est pas le cas — que la psychiatrie et la neurobiologie sont axiomatisées selon la logique des prédicats du premier ordre. J'imagine un tel cadre épistémique parce que je veux vérifier si, dans ce cadre logico-syntaxique fictif, cette réduction interthéorique nagelienne est faisable. L'est-elle, en effet? Il faut répondre à cette question parce que, comme je l'ai expliqué, afin de vérifier si mon *interprétation réductionniste* des modèles hempelien DN et IS de l'explication des *phénomènes naturels* est valable, il faut vérifier si le modèle nagelien de la réduction est viable ou, de manière plus générale, si l'on peut réduire une théorie T_H par une théorie T_B dans le cadre syntaxique des théories scientifiques. En effet, quand on utilisera le modèle hempelien de l'explication des *phénomènes naturels*, nous ne pourrions favoriser les lois de la neurobiologie (L_B) (tirées de T_B) aux lois psychiatriques (L_H) (tirées de T_H) qu'à condition de montrer que ces dernières sont réductibles aux premières. Corollairement, conformément à la théorie nomologique de la causalité mise de l'avant par Hempel et Oppenheim (section 1.2), nous n'aurons le droit d'affirmer que les seules forces causales fondamentales causant les maladies mentales résident aux niveaux neurobiologiques plutôt qu'au niveau psychologique ou environnemental qu'après avoir fait cet exercice de réduction.

Dans la prochaine section, je fais connaître deux problèmes que rencontre le modèle nagelien et qu'ont formulés Karl Popper et Paul Feyerabend. On pourrait penser que ces problèmes rendent inopérant le modèle nagelien et qu'il est impossible, pour cela, de réduire la psychiatrie par la neurobiologie. Mais ce n'est pas le cas. En effet, comme on le verra à la section 2.4, Kenneth F. Schaffner, un étudiant de Nagel, a développé un modèle de la réduction de style nagelien qui ne rencontre pas ces problèmes.

2.3 Les problèmes que pose la caractéristique DN du modèle nagelien

Les critiques qu'ont élaborées les philosophes Karl Popper et Paul Feyerabend relativement au modèle de la réduction interthéorique montrent que les conditions de *déductibilité* et de *connectivité* (voir section 1.4.5) qu'a formulées Nagel sont inapplicables et que, par voie de conséquence, son modèle ne l'est pas plus. Dans ce qui suit, j'explique pourquoi.

2.3.1 L'argument de Popper

Si l'on considère l'histoire des sciences, nous remarquons qu'il existe très peu de théories scientifiques qui furent expliquées selon le critère cardinal du modèle nagelien, c'est-à-dire de manière *strictement déductive*. Par exemple, selon Popper, et contrairement à ce que dit Nagel, la théorie galiléenne de la chute des corps ne peut pas être déduite de la mécanique newtonienne et des lois de la gravitation (Popper (1991[1979], 306-308).⁴³ On ne peut pas faire cette déduction parce que dans la loi⁴⁴

⁴³ Je tire ces idées du livre de Popper *La connaissance objective* paru en 1979 (en anglais). Cependant, j'aimerais souligner que ces idées de Popper proviennent du texte « Le but de la science »

de Galilée $d = gt^2/2$, l'accélération g est une constante alors que dans la théorie de Newton elle n'en est pas une (ibidem, 308). Cependant, comme l'explique Kistler,

[...] la théorie newtonienne permet de dériver une *autre* loi qui constitue une bonne approximation de la loi de Galilée à condition que la distance d soit petite en comparaison du rayon de la Terre. (Kistler 2007, 18, mes *italiques*)

Aussi, comme l'ajoute immédiatement Kistler,

[l]e fait que la loi réduite constitue une bonne approximation de la loi galiléenne explique pourquoi cette dernière est approximativement vraie et pourquoi elle a été confirmée par l'observation *bien qu'elle soit, à strictement parler, fausse*. (ibidem, 18, mes *italiques*)

Cette remarque de Kistler (suivant Popper) selon laquelle ce qui est déduit de la mécanique newtonienne est une *approximation* des lois galiléennes de la chute des corps pose un problème au modèle nagelien parce qu'il rend impossible le respect de la *condition de la déductibilité* qu'a formulée Nagel. En effet, si l'on considère (a) que la loi galiléenne de la chute des corps est fausse *comparativement* à celle de Newton et (b) qu'il est requis que les valeurs de vérité soient conservées lors d'une déduction (ou d'une réduction, selon le modèle nagelien), alors (c) on ne peut pas, si on veut respecter les lois de la logique, déduire la théorie galiléenne de la théorie de Newton. De fait, il est logiquement impossible de déduire une théorie fautive d'une théorie vraie. En d'autres mots, comme le dit Markus Eronen, « [...] since logical deduction is truth-preserving, the new reducing theory cannot both be true and logically entail a false theory » (Eronen 2010, 16).

qui fut originellement écrit en anglais par ce philosophe autrichien en 1957 et publié dans la revue *Ratio*, vol.1, n° 1, p.24-35 sous le titre « The Aim of Science ».

⁴⁴ Il faut lire cette loi de la manière suivante : « [...] la distance d parcourue par un corps en chute libre est proportionnelle au carré du temps t de la chute [...] » (Kistler 2007, 18).

Suivant cet argument de Popper, et si on fait l'hypothèse que la psychiatrie et la neurobiologie sont axiomatisées selon la logique des prédicats du premier ordre et contiennent des lois générales de la nature (ce n'est pas le cas, cependant), on peut dire que, si la médecine mentale contient des lois fausses relativement à celles contenues dans la neurobiologie, l'ensemble des théories constituant la psychiatrie ne peut pas être déduit de cette dernière. Par exemple, bien que la théorie psychanalytique freudienne, en raison des intuitions sur la psyché humaine qu'elle contient, suscite depuis quelques années l'intérêt de beaucoup de chercheurs en neurosciences qui tentent d'en comprendre les bases biologiques avec les outils technologiques actuels (voir Solms 2006), elle contient des principes qui sont empiriquement faux, comme celui du complexe d'Œdipe (voir Bunge 2004, 221-222). Pour cette raison, on ne pourrait pas directement déduire cette théorie fautive — que certains, comme Eric Kandel, incluent dans la psychiatrie (Kandel, 2002 [1998], 35-38)⁴⁵ — de la neurobiologie (que l'on suppose vraie).

2.3.2 L'argument de Feyerabend

Bien que Popper soutienne que la théorie de la chute des corps de Galilée n'est pas déductible de la théorie de Newton, il ne pense pas qu'il n'existe aucune manière de mettre en correspondance ces deux théories. Je me permets de penser que Popper croit qu'il existe une certaine continuité entre les théories scientifiques en général parce qu'il affirme que la nouvelle théorie [celle qui réduit, selon Nagel] *contient l'ancienne* [celle qui est réduite] *de manière approximative* (Popper (1991[1979], 312,

⁴⁵ Kandel affirme, en effet, qu'il « [...] serait malheureux, et même tragique, si les riches découvertes qui sont venues de la psychanalyse devaient être perdues dans le rapprochement entre la psychiatrie et les sciences biologiques » (Kandel, 2002 [1998], 35).

mes *italiques*). Or, ce n'est pas ce que pense Feyerabend (1962). En effet, celui-ci ne croit pas que certaines théories soient reliées par quelques *liens logiques* que ce soit (même pas par un lien d'inclusion, comme chez Popper). De fait, Feyerabend affirme que

[...] [l]es domaines de référence de certaines théories sont incomparables, en ce sens qu'aucune des relations logiques habituelles (l'inclusion, l'exclusion, l'intersection) ne peut être établie entre eux. (Feyerabend 1979 [1975], 246)

Feyerabend pense une telle chose parce que, selon lui, lors d'une réduction d'une théorie T_H par une théorie T_B , l'ontologie de T_H est remplacée par celle de T_B . En conséquence, les *mêmes énoncés d'observation* des différentes théories présentent une signification différente. De fait, il écrit,

[...] [w]hat happens when transition is made from a restricted theory T_H to a wider theory T_B (which is capable of covering all the phenomena which have been covered by T_H) is something much more radical than incorporation of the *unchanged* theory T_H into the wider context of T_B . What happens is rather a *complete replacement* of the ontology of T_H by the ontology of T_B , and a corresponding change in the meanings of all descriptive terms of T_H (provided these terms are still employed). (Feyerabend 1962, 59)

Ces idées de Feyerabend furent à l'origine de la thèse relative à *l'incommensurabilité* de l'ontologie des théories (ou du domaine de référence des théories)⁴⁶. Feyerabend fait reposer cette thèse largement sur l'idée selon laquelle la signification de tous les termes que nous utilisons est déterminée par le contexte théorique [— ou ontologique, suivant la pensée de Feyerabend —] dans lequel ces derniers sont utilisés (Feyerabend 1965, 180, je traduis).

⁴⁶ Chez Feyerabend, à mon avis, l'expression « ontologie de la théorie » réfère ou peut être retraduite par l'expression « domaine de référence de la théorie ». C'est pour cela que je mets entre parenthèses, dans le corps du texte, l'expression « ou du domaine de référence des théories ».

Par exemple, selon Feyerabend (1962, 80; 1979 [1975], 308), interprétées de manière réaliste, la théorie de la mécanique classique et la théorie de la relativité sont incommensurables, c'est-à-dire que leur domaine de référence et, par voie de conséquence, leur signification, n'ont rien en commun. Il en est ainsi parce que, dans le cadre théorique de la première, les propriétés des objets physiques comme la forme, la masse et le volume sont définies comme « des parties inhérentes des objets [...] [ne changeant] qu'à la suite d'une interférence physique directe » (1979 [1975], 308) alors que selon l'interprétation qu'Einstein et Bohr donnèrent de la seconde, ces mêmes propriétés ne sont pas inhérentes aux objets, mais « des relations entre les objets physiques et des systèmes coordonnés qui peuvent changer, sans aucune interaction physique, quand nous remplaçons un système de coordonnées par un autre » (1979 [1975], 308). Par voie de conséquence, comme le dit Alan F. Chalmers, suivant Feyerabend,

[...] tout énoncé d'observation se référant à des objets physiques au sein de la mécanique classique aura une signification différente d'un énoncé d'observation d'aspect similaire en théorie de la relativité. (Chalmers 1987, 220)

Ces idées de Feyerabend, si on les tient pour vraies, rendent impossible la formulation d'hypothèses de liaison parce qu'elles rendent impossibles le respect de la *condition de la connectivité* qu'a formulée Nagel. De fait, si deux théories quelconques T_H et T_B sont incommensurables au sens de Feyerabend, c'est-à-dire que si leurs postulats fondamentaux théoriques (ou leurs domaines de référence) sont incomparables, alors on ne peut pas les mettre en relation à l'aide d'hypothèses de liaison. Par exemple (Feyerabend 1962, 76-78), pour reprendre un cas que nous avons beaucoup analysé jusqu'ici, puisque la thermodynamique n'est fondamentalement pas une science statistique et que la théorie cinétique des gaz en est une, les concepts « températures », « entropie » et « pression » contenus dans chacune de ces théories

présentent des significations radicalement différentes et, en conséquence, il est impossible de formuler l'hypothèse de liaison $kT = 2E/3$, c'est-à-dire de définir un énoncé observationnel de la thermodynamique (la température) en termes statistiques (l'énergie cinétique moléculaire moyenne). Donc, si formuler des hypothèses de liaison n'est pas possible, on ne peut pas déduire la thermodynamique de la mécanique statistique suivant la méthode que prescrit le modèle nagelien. C'est ce qu'explique Robert Batterman, suivant Feyerabend (1962, 76-78), dans le passage suivant :

Consider the term "temperature" as it functions in classical thermodynamics. This term is defined in terms of Carnot cycles and is related to the strict, nonstatistical second law as it appears in that theory. The so-called reduction of classical thermodynamics to statistical mechanics, however, fails to identify or associate *nonstatistical* features in the reducing theory, statistical mechanics, with the nonstatistical concept of temperature as it appears in the reduced theory. How can one have a genuine reduction, if terms with their meanings fixed by the role they play in the reduced theory get identified with terms having entirely different meanings? Classical thermodynamics is not a statistical theory. The very possibility of finding a reduction function or bridge law that captures the concept of temperature and the strict, nonstatistical, role it plays in the thermodynamics seems impossible. (Batterman 2012, section 1)

Suivant cette thèse relative à *l'incommensurabilité* de l'ontologie des théories, on peut penser qu'il serait impossible de réduire la médecine mentale par la neurobiologie, puisque la première, selon les tenants du pluralisme en psychiatrie, entre autres choses, est une théorie « irrévocablement enracinée dans le mental, c'est-à-dire dans l'expérience à la première personne » (Kendler 2005, 433, je traduis) et que la seconde ne l'est pas. Effectivement, par exemple, les processus biomoléculaires, ayant lieu dans l'amygdale du cerveau d'un individu énonçant une croyance de peur malade, n'ont pas peur.

2.3.3 La conclusion générale de Popper et Feyerabend

En raison de ce genre de problématique, c'est-à-dire pour des raisons logiques selon Popper et en raison de l'incommensurabilité des domaines de références des théories selon Feyerabend, on doit conclure qu'une approche strictement logico-syntaxique de la réduction (ou de l'explication des théories) est impraticable. En effet, Popper affirme que, suivant l'exemple de la prétendue déduction de la théorie de Galilée⁴⁷ de celle de Newton,

[...] ni une *inférence déductive* ni une inférence inductive ne peuvent jamais faire passer de prémisses consistantes à une conclusion qui contredise formellement les prémisses dont on est parti. (Popper (1991[1979], 307, mes *italiques*)⁴⁸

Quant à lui, au début de son célèbre texte « Explanation, Reduction, and Empirism » (1962), Feyerabend écrit : «[t]he main contention of the present paper is that formal account of reduction and explanation is impossible for general theories» (Feyerabend 1962, 28). Mais, heureusement pour les tenants du réductionnisme, comme nous le verrons dans la section suivante, Schaffner a modifié le modèle nagelien de telle sorte qu'il ne rencontre pas ces problèmes.

⁴⁷ Et de celle de Kepler, mais je n'en ai pas traité.

⁴⁸ Je cite tout le passage de Popper même si, dans mon texte, je n'ai traité que de l'inférence *déductive*.

2.4 La solution de Schaffner

Schaffner (1967, 1969b, 1974)⁴⁹ tentera de résoudre ces problématiques logiques et celles reliées à l'incommensurabilité des domaines de références des théories en créant un modèle de la réduction interthéorique que l'on pourrait qualifier de « néonagelien ». Schaffner acceptera et inclura dans son modèle de la réduction l'idée selon laquelle la théorie réduite T_H' est une *approximation qui corrige* la théorie T_H que l'on veut réduire. Il soutiendra aussi que T_H' — la théorie réduite corrigée — n'est pas déduite de T_B — la théorie réductrice de base originale —, mais de T_B' , laquelle est T_B enrichie d'hypothèses auxiliaires et d'énoncés de balises physiques (voir Dizadji-Bahmani, Frigg et Hartmann 2010, 398 et Schaffner 1969b, 332; 1993, 427). Par ailleurs, il dira de la théorie corrigée T_H' qu'elle est fortement *analogue* à la théorie que l'on veut réduire T_H et qu'elle fait de *meilleures prédictions* que T_H (c'est pourquoi Schaffner affirme qu'elle corrige T_H) (Schaffner 1967, 144). Selon la lecture que fait Rasmus Grønfeldt Winther des propos de Schaffner, T_H et T_H' font partie d'un ensemble K que l'on peut définir comme l'ensemble des théories semblables à T_H , soit $K = \{T_H, T_H', T_H'', \dots\}$ (Grønfeldt Winther 2009, 123). Par ailleurs, Schaffner soutiendra que, lors d'une réduction, les théories T_H' et T_B' sont reliées par des hypothèses de liaison — construites à partir de T_B' — qu'il nommera *fonction de réduction*. Celles-ci, selon lui, comme je l'ai mentionné plus haut (section 1.4.4), ont la forme logique d'énoncés biconditionnels et mettent en relation soit des *classes d'entités*, soit des *prédicats* (Schaffner 1967, 144). Schaffner dit de ces fonctions qu'elles représentent une *identité synthétique* (factuelle) parce qu'elles mettent en relation *deux concepts différents* (soit des classes d'entités ou des prédicats) *qui ont la même extension*, c'est-à-dire qui réfèrent à la même entité ou à la même propriété

⁴⁹ Schaffner peaufinera son modèle dans son livre *Discovery and Explanation in Biology and Medicine* (1993, chap.9).

d'une entité physique. Par exemple, de dire Schaffner, « [...] the term 'gene' can be understood to refer to the same entity which is named by a sequence of nucleotides of DNA (or RNA in some special cases involving viruses) » (Schaffner 1974, 614). Aussi, il est pertinent d'ajouter que ces identités sont dites « *synthétiques* » parce que, quand elles ont été formulées, elles étaient confirmées empiriquement et non par une analyse linguistique (en ce cas, on aurait dit d'elles qu'elles sont des identités *analytiques*) (ibidem, 614). En effet, remarque Schaffner, « [ge]nes were not discovered to be DNA via the analysis of meaning; important and difficult empirical research was required to make such an identification » (ibidem, 614-615).

Il est important de noter que, à l'inverse de Nagel qui ne s'intéressait pas à la métaphysique, selon Schaffner, concevoir les fonctions de réduction comme des identités synthétiques leur donne une portée ontologique. Effectivement, comme le souligne Schaffner, « [i]t is only by use of synthetic identities that reduction can decrease the ontology of the universe, without imputing "unreality" to some theoretical entities » (Schaffner 1967, note 9). Par exemple, comme l'explique Schaffner, c'est parce que, à un certain moment dans l'histoire des sciences, on n'arrivait plus à identifier l'éther à quoi que ce soit de réel ou de physique qu'on jugea que l'éther n'existait pas ou ne faisait pas partie de la constitution du monde (Schaffner 1969a, 285). Cette idée de simplification ontologique implique que certaines théories qui ne décrivent pas adéquatement la réalité doivent être remplacées — et non réduites — par d'autres théories. Par exemple, la théorie selon laquelle les maladies sont produites par des démons malicieux fut remplacée par la théorie selon laquelle les maladies sont produites par des germes délétères (Schaffner 1993, 428). En psychiatrie, dans cette optique, on pourrait dire que les théories psychanalytiques de l'autisme ont été remplacées (et non réduites) par des théories biologiques.

En somme, suivant ce qui précède, il faut retenir que le modèle de Schaffner contient les éléments suivants :

- a) T_B : la théorie réductrice
- b) T_B' : la théorie réductrice corrigée
- c) T_H : la théorie à réduire
- d) T_H' : la théorie à réduire corrigée (déduite de T_B')
- e) Les fonctions de réduction
- f) Les énoncés de balises physiques

Si on résume, pour réduire T_H par T_B , selon Schaffner, il faut, premièrement, reconstruire, à partir de T_B , une théorie T_B' . Deuxièmement, il faut déduire de T_B' , à l'aide des fonctions de réduction construites dans T_B' , une théorie T_H' qui est fortement analogue à T_H . Le schéma suivant illustre bien cette procédure :

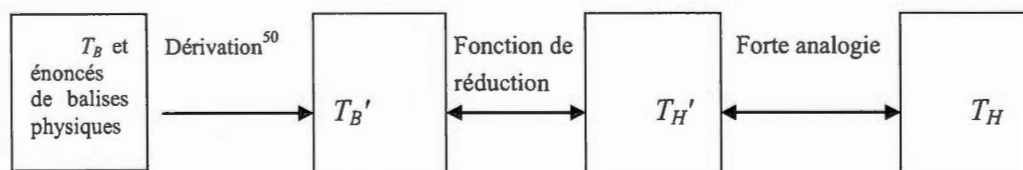


Figure 1 Le modèle de la réduction de Schaffner (schéma tiré de Dijadji-Bahamani, Frigg et Hartman 2010, 399).

Comme on peut le constater, le modèle de la réduction de Schaffner réussit à absorber les critiques de Popper et Feyerabend. D'une part, en incorporant le fait que la théorie réduite et la théorie réductrice changent lors d'une réduction, Schaffner

⁵⁰ Dizadji-Bahmani, Frigg et Hartmann (2010, 399) utilisent le terme anglais « derivation ». Je suppose que ce terme fait référence au fait que l'on *reconstruit* T_B' à partir de T_B et des énoncés de balises physiques de cette dernière.

absorbe la critique de Popper. De fait, c'est T_H' et non T_H qui est déduite de T_B' (et non de T_B). Concrètement, pour reprendre l'exemple de Popper, le modèle schaffnerien incorpore l'idée selon laquelle une théorie galiléenne *corrigée* (T_H'), laquelle ne contient pas d'énoncés faux *relativement* à une théorie mécanique de Newton enrichie d'hypothèses auxiliaires et d'énoncés de balise physique (T_B'), est déduite de T_B' . D'autre part, on peut dire que le modèle de Schaffner absorbe la critique de Feyerabend pour les mêmes raisons. En effet, en acceptant l'idée selon laquelle la réduction s'opère entre une version corrigée T_H' et la théorie réductrice modifiée T_B' , Schaffner établit, à l'aide de son modèle, un *lien de commensurabilité* (Winther 2009, 123) qui permet de mettre en relation les deux domaines de référence des théories en présence.

En psychiatrie et en biologie, à ma connaissance, ne serait-ce que parce que ces disciplines ne sont pas axiomatisées selon la logique du premier ordre et ne contiennent pas de grands principes et de lois universelles, de telles réunifications *réductrices* logico-syntaxiques schaffneriennes n'existent pas encore.⁵¹

Il est maintenant temps, dans la prochaine section, de faire connaître un problème épistémique majeur que rencontre le modèle néonagelien de Schaffner.

⁵¹ Mais on retrouve d'autres formes *non réductrices* d'harmonisation théorique intéressante. En effet, en science, il s'opère des *fusions de champs disciplinaires*. Par exemple, comme le dit Mario Bunge, l'intégration de la psychologie, de la neurobiologie, de l'endocrinologie, de l'immunologie, de la neurologie et de la psychiatrie forme ce qu'on appelle les *neurosciences cognitives* ou la *psychobiologie* (Bunge 2004, 118-119). Comme l'explique ce dernier, « [...] le but de cette synthèse nouvelle [...] [consiste] à "mettre en correspondance le cerveau et l'esprit" (*mapping the mind onto the brain*) » (Bunge 2004, 119). De plus, les neurosciences cognitives peuvent être elles-mêmes enrichies et fusionnées à d'autres disciplines comme l'informatique et les mathématiques. Des travaux extrêmement intéressants doivent être mentionnés dans ce champ de recherches des neurosciences mathématisées. Je pense ici aux travaux où on montre qu'en traitant des images par résonance magnétique fonctionnelle (IRMf) avec des modèles informatiques reposant sur l'interprétation bayésienne de la théorie des probabilités, il est possible de voir sur un écran les représentations mentales advenant à l'intérieur de la tête des gens (voir Koch 2012 pour une vulgarisation de ces travaux).

2.5 Le problème épistémique que rencontre le modèle de Schaffner

Comme je l'ai mentionné plusieurs fois déjà, la psychiatrie et la neurobiologie sont loin d'être axiomatisées selon la logique des prédicats du premier ordre et ne présentent pas de lois universelles comme en physique (Eroren 2010, 19-20; Kitcher 1984; Sarkar 1998). Conséquemment, en principe, il semble impossible d'utiliser le modèle néonagelien de Schaffner pour réduire la médecine mentale à la neurobiologie. Mais supposons que, dans un avenir hypothétique, nous réussissions à axiomatiser selon la logique des prédicats du premier ordre la psychiatrie et la neurobiologie et à leur trouver de grands principes et des lois, resterait-il encore des difficultés? Il semble que oui. En effet, il resterait l'obstacle à la réduction interthéorique qu'a mis en lumière Jerry Fodor (1974) et qui découle du fait que les entités postulées par les théories de haut niveau sont réalisées de façon multiple au niveau microphysique (ou biomoléculaire dans notre cas). Je présente cet argument dans ce qui suit⁵².

2.5.1 L'argument antiréductionniste de la réalisation multiple

Jerry Fodor (1974), dans un article maintenant célèbre, a formulé l'argument de la réalisation multiple qui réfute la thèse selon laquelle les sciences spéciales, c'est-à-dire les sciences autres que la physique — et donc la science psychiatrique — sont

⁵² La section 4.2 de l'article « Reductionism in Biology » (2012) d'Ingo Brigandt et d'Alan Love dans *The Stanford Encyclopedia of Philosophy* est particulièrement éclairante relativement à l'argument de Fodor. Je m'inspire fortement du contenu de cette section et du texte de Fodor (1974) pour la présentation de cet argument.

réductibles aux théories et aux lois de théories dites plus fondamentales. Voici son argument.

Premièrement, imaginons une loi hypothétique de la science psychiatrique qui présente la forme d'un énoncé conditionnel universel⁵³ reliant deux espèces naturelles⁵⁴, soit Sx et Tx . On peut exprimer formellement cette loi de la manière suivante :

$$\forall x (Sx \rightarrow Tx)$$

Cet énoncé nomologique peut être lu comme ceci : « Tous les S sont des T » ou « Tous les événements de type S engendrent des événements de type T . » Dans le cadre de la psychiatrie, on pourrait illustrer concrètement cette formule par l'exemple suivant : « Toutes les hallucinations de type S engendrent des comportements anormaux de type T . »

Deuxièmement, supposons que nous voulons réduire cette loi de la psychiatrie à une loi de la neurobiologie moléculaire. Est-ce possible? Il semble que non. En effet,

⁵³ Le quantificateur universel n'est pas approprié aux lois des sciences spéciales parce que celles-ci ne valent pas universellement. Mais, pour les besoins de l'argumentaire, pour l'instant, on fera comme-ci ces lois étaient universelles.

⁵⁴ Alexander Bird et Emma Tobin, dans le passage qui suit, expliquent bien en quoi consiste une espèce naturelle :

Les données que l'on étudie selon les diverses disciplines scientifiques sont divisées en espèces à propos desquelles on théorise. Or, dire d'une espèce qu'elle est « naturelle » revient à dire que cette espèce correspond à un regroupement ou à un ordonnancement qui ne dépend pas de manipulations humaines. On suppose trouver ces espèces en faisant de la science. Selon le réalisme scientifique, quand tout va bien, les classifications et les taxonomies utilisées en science font référence à des espèces réelles de la nature. L'existence de ces espèces réelles et indépendantes [de nous] justifie nos inférences et pratiques scientifiques. (Bird et Tobin 2008, je traduis librement)

une espèce naturelle de haut niveau — par exemple une hallucination chez un individu — n'est jamais réalisée par les mêmes espèces naturelles du niveau biomoléculaire. En effet, ce ne sont pas les mêmes entités et processus biomoléculaires qui provoquent une même hallucination. Or, comme l'expliquent Ingo Brigandt et Alan Love (2012), si l'espèce naturelle psychologique S que l'on retrouve dans la loi précédente est réalisée par les espèces naturelles biomoléculaires M_1, M_2, \dots, M_i , alors l'hypothèse de liaison (ou la fonction de réduction, comme la nommerait Schaffner) que nous formulerons prendra une forme disjonctive comme celle-ci :

$$\forall x (Sx \leftrightarrow (M_1x \vee M_2x \vee \dots \vee M_ix))$$

De plus, comme le soulignent Brigandt et Love (2012), l'espèce naturelle T de la précédente loi serait elle aussi multi réalisée par des espèces biomoléculaires N_1, N_2, \dots, N_j . C'est en accord avec ce fait empirique que Fodor (1974, 109) formule la thèse selon laquelle la base microphysique (ici, biomoléculaire) sur laquelle est fondée la loi de la science spéciale $\forall x (Sx \rightarrow Tx)$ consiste en différentes lois biomoléculaires (réduisant les différentes instances de x décrit par cette même loi de haut niveau) que l'on peut représenter formellement comme ceci :

$$\forall x (M_1x \rightarrow N_3x), \forall x (M_2x \rightarrow N_1x), \forall x ((M_3x \vee M_4x) \rightarrow N_2x), \dots$$

Ainsi, comme l'expliquent Brigandt et Love (2012), alors qu'au niveau de la réalité décrite selon le domaine de la science spéciale nous avons une unique loi $\forall x (Sx \rightarrow Tx)$, la traduction de cette dernière en termes moléculaires consiste en un énoncé disjonctif comme celui-ci :

$$\forall x ((M_1x \vee M_2x \vee \dots \vee M_ix) \rightarrow (N_1x \vee N_2x \vee \dots \vee N_jx))$$

Comme le disent Brigandt et Love (2012), cet énoncé disjonctif [potentiellement composé d'une infinité d'espèces biomoléculaires] n'a pas l'apparence d'une véritable loi. En effet, « law of nature cannot have the form of a disjunction, let alone an open-ended one: Laws cannot be disjunctive » (Dizadji-Bahmani, Frigg et Hartmann 2010, 401). Pour cette raison, selon Brigandt et Love (2012), cette loi de haut niveau ne peut pas être réduite à une seule et unique loi de bas niveau.

En somme, vu cela, dans ce cadre syntaxique des théories scientifiques, il est impossible de réduire, par voie déductive et interthéorique, les lois de la psychiatrie par celles de la neurobiologie moléculaire. Corollairement, on doit conclure que, dans ce cadre logico-syntaxique formulé par les tenants de l'empirisme logique, *on ne peut pas retenir et utiliser essentiellement l'interprétation réductionniste des modèles DN et IS de l'explication des phénomènes naturels* (voir section 1.3). Autrement dit, ne serait-ce que pour cette raison épistémique, en psychiatrie, dans le cadre de la conception syntaxique des théories scientifiques (voir section 1.1), on ne peut pas faire appel essentiellement aux lois de bas niveau dans les modèles de lois de couverture.

Dans la prochaine section, j'explique que même si cette réduction interthéorique s'avérait faisable, il n'aurait pas été possible d'utiliser les modèles hempelien DN et IS (voir section 1.2) pour expliquer les phénomènes naturels ayant lieu dans le cerveau et, en conséquence, puisque les maladies mentales impliquent des processus

cérébraux⁵⁵, les phénomènes naturels étudiés dans le domaine de la psychiatrie. En effet, comme nous le verrons, ces modèles DN et IS présentent des problèmes.

2.6 Les problèmes que les modèles DN et IS rencontrent dans l'explication du fonctionnement du cerveau

Le cerveau humain est un système biologique d'une complexité fabuleuse capable d'émotions, de cognitions et de production de comportements. Pour comprendre ces fonctions (et les dysfonctions de ces fonctions), en neurosciences, « l'une des approches les plus classiques consiste en l'analyse des constituants du système nerveux » (Bear, Connors et Paradiso 2002, 13). C'est la dimension de ces constituants qui détermine le niveau de l'analyse qu'on utilisera pour étudier ces derniers. Dans les sciences du cerveau, actuellement, on conçoit cinq niveaux d'analyse (ibidem, 13-14). Premièrement, il y a le niveau d'analyse biomoléculaire où on étudie les diverses molécules dont est composé le cerveau et qui sont responsables du fonctionnement cérébral. Deuxièmement, il y a le niveau d'analyse cellulaire où on étudie comment toutes ces molécules confèrent à l'unité fonctionnelle de base du système nerveux — la cellule nerveuse qu'est le neurone — ses propriétés particulières (ibidem, 13). Troisièmement, il y a le niveau d'analyse intégrée où on étudie les systèmes de neurones qui sont à la base de différentes fonctions du cerveau comme, par exemple, celle de la vision (on étudiera alors le système visuel). Quatrièmement, il y a le niveau d'analyse comportementale où on étudie les systèmes de systèmes neuronaux à la base de nos comportements. Il s'agit de voir, par exemple, dans quelle partie du cerveau agit l'alcool et pourquoi cette substance altère les

⁵⁵ Attention. Je ne dis pas que les maladies mentales sont nécessairement *causées* par l'activité moléculaire, cellulaire et des systèmes cérébraux. J'énonce seulement le fait trivial que ces maladies *résultent* de ces activités, de la même façon que lever un bras ou dire les mots « joie et tristesse » en résulte.

fonctions de l'esprit et l'humeur. Cinquièmement, il y a le niveau d'analyse cognitif où on tente d'appréhender les bases neurobiologiques « responsables des plus hauts niveaux de l'activité mentale chez l'homme, tels que la conscience, les représentations mentales et le langage » (ibidem, 14).

En psychiatrie, un modèle explicatif adéquat devra minimalement être capable de rendre compte des phénomènes naturels décrits selon les différents niveaux d'analyse susmentionnés. Ainsi, pour être utiles en psychiatrie, les modèles DN et IS doivent être capables de s'acquitter de cette tâche. Le sont-ils? Il semble que non, comme l'a montré Carl Craver (2007, 34-40). Pour faire valoir son point de vue, celui-ci utilise un exemple — celui de la transmission synaptique — tiré des neurosciences cellulaires et biomoléculaires. Avant d'expliquer pourquoi les modèles des lois de couverture sont incapables de rendre compte des explications données en neurosciences (et, donc, en psychiatrie), je présente brièvement certains éléments fondamentaux du système nerveux que l'on doit connaître pour comprendre les arguments de Craver.

2.6.1 Éléments fondamentaux du système nerveux

Le système nerveux est une machine biologique dont une des principales fonctions est de traiter l'information provenant de l'environnement extérieur et du corps humain. Cette information, quand elle est traitée par le cerveau, prend la forme d'influx nerveux qu'on peut aussi appeler « potentiels d'action ». Les neurones sont les voies par lesquelles passe cette information nerveuse. Les composantes principales des neurones sont le corps cellulaire, le noyau qui contient l'information génétique qui détermine le fonctionnement du neurone (Feldman, Giroux et Cauchy 1994, 85), l'axone — par lequel voyage les potentiels d'action —, la gaine de

myéline faite de cellules gliales assurant la conduction nerveuse, les terminaisons axonales qui transmettent l'influx nerveux à d'autres neurones et les dendrites qui reçoivent les influx provenant d'autres neurones. En plus d'être fait de neurones (la matière grise), le cerveau est constitué des cellules gliales (la matière blanche ou la glie). Ces dernières forment l'environnement des neurones et contribuent au bon fonctionnement de ces derniers en les nourrissant, en les supportant, en les protégeant, en éliminant les déchets causés par la mort des neurones (l'apoptose) et « en agissant comme gaine isolante [— c'est la gaine de myéline susmentionnée —] de certains axones » (Le cerveau à tous les niveaux 2012a). Bref, comme on peut le constater, le cerveau est un organe dont la structure et la constitution assurent, entre autres, la bonne conductance — essentielle pour notre bien-être mental — des influx nerveux.

Ces influx nerveux ou ces potentiels d'action présentent une double caractéristique. En effet, ils sont à la fois électriques et chimiques. Quand ils voyagent dans l'axone vers les terminaisons axonales, les influx nerveux sont de nature électrique. Ceux-ci résultent d'une dépolarisation de la membrane cellulaire, c'est-à-dire du passage transitoire du potentiel électrique de membrane négatif vers un état positif (Wikipédia 2012a). Cette dépolarisation est causée par une montée de la concentration d'ions de sodium (Na^+) dans le milieu intracellulaire du neurone et par une montée de la concentration d'ions de potassium (K^+) dans le milieu extracellulaire. Les potentiels d'action vont toujours dans la même direction.

Pour aller dans les autres neurones, les potentiels d'action doivent traverser l'espace existant entre les neurones et que l'on appelle l'espace synaptique. Pour ce faire, puisque les influx nerveux ne peuvent pas traverser cet espace sous leur forme électrique, ils doivent être « métamorphosés », si l'on peut dire, en substance chimique ou, plus précisément, en neurotransmetteurs. Cette transformation a lieu

quand un potentiel d'action atteint la terminaison axonale. À ce moment, en effet, ce potentiel d'action accompagné d'une hausse de la concentration de Na^+ dans la membrane intracellulaire cause une élévation de la polarité de la membrane, ce qui cause l'ouverture des canaux à ions de calcium (Ca^{2+})⁵⁶. Ce faisant, des ions de Ca^{2+} entrent dans la terminaison axonale et provoquent une cascade de réactions chimiques qui engendre (1) le déplacement puis la fusion de vésicules contenant les neurotransmetteurs avec l'extrémité de la membrane plasmique de l'axone (il s'agit de l'exocytose des vésicules) et (2) la libération des neurotransmetteurs dans l'espace synaptique. Il est important de noter que ce n'est pas dans tous les cas que le potentiel d'action et l'entrée de Ca^{2+} provoquent la décharge de neurotransmetteurs. En effet, le cerveau étant une machine biologique fonctionnant intrinsèquement de manière probabiliste, seulement 10 % à 20 % des potentiels d'action (et d'influx de Ca^{2+}) causent cette libération (Craver 2007, 26).

⁵⁶ J'aimerais préciser que ce n'est pas l'élévation de Na^+ qui cause l'ouverture des canaux calciques et l'entrée des ions de calcium Ca^{2+} lesquels, suite à des processus chimiques, causent la libération des neurotransmetteurs. Nous savons cela parce que, comme l'explique Craver (2007, 22-23) qui s'appuie sur les expériences de Bernard Katz et Ricardo Miledi (1967), il est possible de provoquer une dépolarisation (donc un potentiel d'action) autrement que par l'augmentation de Na^+ dans l'espace intracellulaire. Et cette dépolarisation qui n'est pas causée par une élévation de Na^+ , mais seulement par une charge électrique administrée par les expérimentateurs (Katz et Miledi, 1967, 535), provoque elle aussi l'ouverture des pompes à calcium et, ultimement, la décharge de neurotransmetteurs (d'acétylcholine, dans le cadre de l'expérience susmentionnée de Katz et Miledi). Donc, dans les faits, ce qui fondamentalement cause l'ouverture des pompes à calcium et ensuite la décharge de neurotransmetteurs est le potentiel d'action et non l'élévation d'ions de Na^+ dans la membrane cellulaire. La présence du calcium, cependant, est indispensable à cette libération. En effet, s'il y a un potentiel d'action, mais qu'il n'y pas de calcium dans l'espace intracellulaire, il n'y aura pas de libération de neurotransmetteurs (Craver 2007, 23).

2.6.2 Le modèle des lois de couverture, le potentiel d'action et le problème de la pertinence

Craver pense que le modèle de lois de couverture ne convient pas à la formulation d'explication en neuroscience parce que ce modèle permet de construire des explications qui contiennent des informations non pertinentes, et ce, même si cette explication respecte les cinq conditions caractérisant les modèles de loi de couverture (voir section 1.2). Pour faire valoir son point de vue, Craver présente l'exemple suivant :

[...] consider an experiment in which neurons are first blessed by an ordained parson (for example, by sprinkling them with isotonic holy water) and then stimulated with a 10-nanoampere (nA) current for one second. This experiment would no doubt confirm the generality that blessed neurons produce action potentials when stimulated with a 10 nA current. From this regularity, one could conclude that when a given cell has been blessed and stimulated with a 10 nA current, it will generate an action potential. (Craver 2007, 36)

Je pense que Craver a en tête la structure argumentative suivante lorsqu'il formule cet exemple⁵⁷ :

Tous les neurones bénis (**B**) stimulés par un courant de 10 nanoampères pendant une seconde (**S**) produisent des potentiels d'action (**P**). [$\forall x (Bx \wedge Sx \rightarrow Px)$]

Les neurones ont été bénis (cela signifie qu'ils ont été aspergés d'eau bénite isotonique). (**B**)

Les neurones ont été stimulés par un courant de 10 nanoampères pendant une seconde. (**S**)

Les neurones ont produit un potentiel d'action. (**P**)

⁵⁷ Cet argument, comme le dit Craver, a une structure semblable à l'argument que présente H. Kyburg (1965) dans son article classique où il est question non pas de neurones bénis, mais de sel enchanté.

Aussi surprenant que cela puisse paraître, il s'agit d'une explication DN adéquate. Toutefois, nous savons très bien que le fait que les neurones présentent la propriété d'être bénis ne cause pas la production du potentiel d'action. Il s'agit d'une information non pertinente. De fait, c'est seulement la stimulation qui cause la production du potentiel d'action. Ainsi, comme on peut le voir, dans le cadre du modèle des lois de couverture, l'ajout de prémisses non pertinentes — celles concernant le fait que les neurones soient bénis — n'a pas d'incidence sur la force de l'argument. Comme le dit Craver, en effet, «[t]he strength of an argument is not diminished one bit by the addition of any number of irrelevant premises» (ibidem, 37)⁵⁸.

Bref, dans ce cas, on peut dire que le modèle des lois de couverture permet de formuler des explications douteuses au niveau de l'analyse cellulaire du cerveau (voir section 2.6 pour la définition de ce niveau). Or, qu'un modèle de l'explication permette qu'il y ait des informations non pertinentes dans une explication sans rendre invalide cette dernière n'est pas souhaitable en science. Il s'agit de la première raison qui m'incite à rejeter les modèles DN et IS en psychiatrie.

2.6.3 La libération des neurotransmetteurs et le problème lié à la thèse de la prévisibilité nomique

Comme nous l'avons vu (section 1.2), selon Salmon, l'élément au cœur des modèles DN et IS est l'idée de prévisibilité nomique⁵⁹. De fait, dans le cadre des modèles des

⁵⁸ À la section 3.8.1, je montre que le modèle interventionniste de Woodward ne présente pas ce problème.

⁵⁹ De fait, Salmon écrit, au début de son célèbre article « Statistical Explanation » (1971) :

Ever since his classic paper with Paul Oppenheim, « Studies in the Logic of Explanation », first published in 1948, Carl G. Hempel has maintained that an « explanatory account [of a

lois de couverture, l'idée selon laquelle *l'explanans* nous permet de prévoir *l'explanandum* avec certitude ou quasi-certitude est centrale. C'est ce que dit clairement Hempel dans le passage suivant :

Une réponse rationnellement acceptable à la question « pourquoi l'événement X s'est-il produit? » doit nous donner des informations qui montrent que l'on devait s'attendre à X — sinon de manière certaine, comme dans le cas d'explication DN, du moins avec une probabilité raisonnable. Ainsi l'information explicative doit donner de bonnes raisons de bases pour croire que X s'est bien produit; sinon cette information ne nous donnerait aucune raison adéquate pour dire « Voilà l'explication — cela montre pourquoi X s'est produit » (Hempel 1965c, 367-368, traduction de Denis Bonnay dans Barberousse, Bonnay et Cozic 2011, 20)

Considérant le fonctionnement intrinsèquement probabiliste du cerveau, comme le fait remarquer Craver (2007, 39-40), cette exigence de prévisibilité nomique pose un problème. En effet, dans le cadre du modèle des lois de couverture, le fait qu'un potentiel d'action cause la libération de neurotransmetteurs que dans 10 à 20 pour cent des fois *ne peut pas* être retenu comme une explication valable de cette libération. Cette probabilité, de fait, est trop faible pour convenir à l'exigence de prévisibilité nomique formulée par Hempel. Donc, le modèle de Hempel ne réussit pas à appréhender ce qui, dans les faits, est une explication. Cela est problématique en science. En conséquence, un modèle de l'explication ne doit pas reposer sur un tel principe. En effet, comme le dit Salmon, « [i]n spite of its enormous intuitive appeal, I believe that this precept is incorrect and that it has led to an unsound account of scientific explanation » (Salmon 1971, 29).

particular event] may be regarded as an argument to the effect that the event to be explained...*was to be expected* by reason of certain explanatory facts » (my italics). It seems fair to say that this basic principle has guided Hempel's work on *inductive* as well as *deductive* explanation ever since. (ibidem, 29)

En définitive, dans le cadre des modèles de lois de couverture (voir section 1.2), eux-mêmes formulés dans le cadre de la conception syntaxique des théories scientifiques (voir section 1.1), comme on peut le constater, on ne peut pas rendre compte de ce qui explique dans les faits la neurotransmission. En conséquence, on ne peut pas retenir ces modèles pour les neurosciences et pour la psychiatrie (dans la mesure où les maladies mentales résultent, comme tout ce que nous faisons et sommes, nécessairement de processus cérébraux).

J'aimerais finir ce chapitre en montrant que même s'il était formellement possible de construire sans heurts une explication réductionniste ou non réductionniste dans ce cadre syntaxique, il resterait un problème majeur. Je pense, en effet (suivant Eronen, 2010, 19-20), qu'au-delà des problèmes formels que présentent les modèles DN et IS ce type d'explication mettant en jeu des entités linguistiques évoluant dans un réseau nomique structuré selon la logique du premier ordre est une mauvaise façon de se *représenter* une explication en médecine mentale.

2.6.4 Le problème de la représentation de l'explication des troubles de santé mentale que posent les modèles DN et IS

Selon moi, les modèles IS et DN donnent une mauvaise *représentation* de l'explication — qu'elle soit réductionniste ou non réductionniste — d'un trouble de santé mentale. Afin de faire comprendre cette idée, faisons fi, pour les besoins de l'argumentaire, de tous les problèmes susmentionnés. Ensuite, supposons que nous avons réussi à réduire une théorie psychologique de la dépression à une théorie neurophysiologique selon la méthode Nagelo-Schaffnerienne. Suivant ce que j'ai dit à

la section 1.3, nous serions alors en droit de formuler une *explication réductionniste* IS de ce type⁶⁰ :

Explication IS 1

Presque toutes les personnes dont l'aire cérébrale 25 est hyperactive (**A**)
présentent une dépression majeure (**D**). (Loi probabiliste) [explanans]

L'aire 25 du cerveau de Jean (**j**) est hyperactive (**A**). (Condition initiale)
[explanans]

===== [r]

Jean (**j**) souffre d'une dépression majeure (**D**). (Description du phénomène
empirique qui est à expliquer) [explanandum]

Dans le cadre de la conception syntaxique des théories scientifiques (voir section 1.1), cet argument inductif se réduit à la structure logique suivante :

P (D|A)=r

A_j

===== [r]

D_j

Voyons maintenant une *explication non réductionniste ou pluraliste*⁶¹ de cette même dépression majeure⁶² :

⁶⁰ Je m'inspire de l'article de Thomas Insel (2010) pour formuler cette explication IS.

⁶¹ Je considère qu'une explication est non-réductionniste ou pluraliste si elle fait référence à autre chose qu'aux concepts de la neurobiologie du cerveau, à savoir à des facteurs psychologiques ou sociaux, par exemple. Ainsi, plus une explication en psychiatrie se situera essentiellement au niveau biomoléculaire, plus elle sera réductionniste.

⁶² Je m'inspire des propos de Mitchell (2009, 5-11) pour formuler cette explication IS.

Explication IS 2

Presque toutes les personnes présentant un profil génétique vulnérable à la dépression majeure (**G**) et évoluant dans un environnement stressant (**E**) développeront une dépression majeure (**D**). (Loi probabiliste) [explanans]

Jean (**j**) présente un profil génétique vulnérable au développement de la dépression majeure (**G**) (Condition initiale) [explanans]

Jean (**j**) évolue dans un environnement stressant (**E**) [explanans]

===== [r]

Jean (**j**) souffre d'une dépression majeure (**D**). (Description du phénomène empirique qui est à expliquer) [explanandum]

Dans le cadre de la conception syntaxique des théories scientifiques (voir section 1.1), cet argument inductif se réduit à la structure logique suivante :

$$\frac{P(D|G \wedge E) = r}{G_j \wedge E_j} = [r]$$

D_j

Est-ce que les praticiens et les chercheurs en psychopathologie se représentent l'explication d'un trouble de santé mentale selon ce langage de la logique du premier ordre? Selon ce que l'on observe dans la littérature scientifique, que l'explication de la maladie mentale soit réductionniste (par exemple, Insel 2010, 46) ou non-réductionniste (par exemple, Kendler et coll.1993 cités dans Murphy 2006, 136), il semble que non. De fait, les psychiatres et chercheurs conçoivent plutôt une maladie mentale selon un *modèle mécaniste* représentant un *mécanisme physique* évoluant dans le temps. Par exemple, un des représentants éminents du réductionnisme (selon Gold 2009), le directeur du *National Institute of Mental Health*, Thomas R. Insel, utilise un modèle mécaniste pour représenter (par la flèche jaune) le circuit de la dépression majeure (DM) dans le cerveau (Insel 2010, 46). Ce schéma, à mon avis,

peut être considéré comme une reformulation en termes mécanistes de l'explication *IS 1* susmentionnée. Voici le schéma d'Insel (2010, 46).

Explication 3 ou explication IS 1 selon un schème mécaniste



Figure 2 La dépression majeure (schéma tiré d'Insel 2010, 46).

Dans ce schéma, l'aire 25 est représentée par la boule jaune, l'amygdale est représentée par l'entité mauve, l'hypothalamus est représenté par la boule verte, l'hippocampe est représenté par la structure bleue et le cortex insulaire est représenté par la structure grisâtre.

Si je résume les propos d'Insel, celui-ci veut faire comprendre que l'hyperactivité de l'aire 25, laquelle est située dans le cortex préfrontal, entraîne un désordre neuronal systémique dans le système nerveux. Dans cette perspective réductionniste (seul le cerveau est en jeu, dans cette explication, en effet), selon Insel, c'est ce désordre — causé fondamentalement, on le suppose, par une anomalie génétique qui a produit cette aire 25 plus petite que la normale et défectueuse —, qui serait à l'origine des symptômes principaux de la DM. Premièrement, ce dérèglement de l'aire 25 produirait un taux d'anxiété et de stress plus élevé. En effet, on suppose que l'aire 25 défectueuse qui est en réseau avec l'amygdale et avec l'hypothalamus

provoque une hyperactivation de la première qui produit l'anxiété et dérègle la seconde qui est une structure impliquée dans les réponses au stress. Deuxièmement, ce désordre causerait une déformation de la réalité. De fait, comme l'amygdale et l'hypothalamus échangent des signaux avec l'hippocampe — une structure impliquée dans les processus de consolidation mnésique — et le cortex insulaire — une structure qui traite les perceptions sensorielles et les émotions —, il est plausible de penser que — et là j'interprète les propos d'Insel qui ne dit pas cela explicitement — l'enregistrement des souvenirs se fait inadéquatement chez l'individu déprimé et que, en conséquence, celui-ci finit par se représenter le monde selon une vision biaisée remplie de croyances négatives selon lesquelles il est impuissant et selon lesquelles la vie est anxiogène et/ou sans intérêt. Troisièmement, cette hyperactivité de l'aire 25 — par le biais du désordre systémique qu'elle engendre — causerait d'autres symptômes comme une baisse d'énergie, une altération de l'humeur, un mauvais sommeil, etc. Pour ces raisons, Insel pense que, pour guérir une dépression majeure, il faudrait manipuler par voie pharmacologique⁶³ l'activité de l'aire 25 de telle sorte qu'elle fonctionne normalement et que les activités neurales des structures cérébrales qu'elle influence redeviennent normales elles aussi (Insel 2010, 47).

Le représentant du pluralisme Kenneth S. Kendler, quant à lui, s'explique et se représente la DM selon le modèle suivant :

⁶³ Les propos qu'Insel tient dans le passage suivant m'incitent à penser que ce dernier privilégierait le traitement pharmacologique (ou même un vaccin [voir Insel 2010, 51]) : « If area 25 can cause the brain, like a computer, to get stuck in loop of abnormal activity, then the goal of treatment might be akin to "rebooting" a computer that has become frozen » (Insel 2010, 47).

Explication 4 ou explication IS 2 selon un schème mécaniste

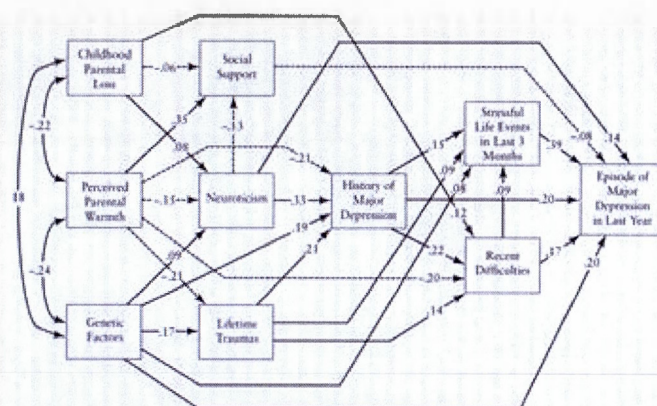


Figure 4.1. Statistical relationships among risk factors for depression in women (Kendler et al. 1993)

Figure 3 La dépression majeure chez les femmes (schéma tiré de Kendler et coll. 1993 cités dans Murphy 2006, 136).

Si ce schéma n'est pas exactement un modèle mécaniste de la dépression (au sens de Glennan 2002, 2005 [voir section 3.2.2 de ce mémoire]), Kendler voudrait idéalement que l'on réussisse à construire de tels modèles pour chacun des troubles de santé mentale (voir Kendler 2008a et Kendler, Zachar et Craver, C. 2011, 1146). Ce dernier, en effet, pense que l'on doit tenter de modéliser les troubles de santé mentale lesquels sont, selon lui, des «complex and multi-level causal mechanisms that produce, underlie and sustain psychiatric syndromes» (Kendler, Zachar et Craver, C. 2011, 1146). Dans ce qui suit, pour les besoins de mon argumentation, je me situe dans une perspective idéale où le schéma apparaissant ci-haut est véritablement un modèle mécaniste.

Comme on peut le constater, contrairement à celui qu'a construit Insel, ce modèle mécaniste de la DM présente une explication pluraliste de cette dernière. En effet, il met en jeu non seulement des facteurs causaux biologiques (les facteurs

génétiques)⁶⁴, mais aussi des facteurs causaux environnementaux et psychologiques. Ainsi, dans le cadre de ce modèle, comme l'explique Murphy, la probabilité qu'un individu souffre d'une DM est fonction de ces neuf variables causales qui interagissent entre elles (Murphy 2006, 135, mes *italiques*)⁶⁵. On pourrait aussi penser, il me semble, que ce modèle représente les réseaux causaux potentiels qui mènent à une DM. En effet, les individus ne développent pas une DM en raison des mêmes facteurs causaux. Par exemple, on peut supposer qu'un individu a développé une DM parce qu'il présentait un profil génétique vulnérable à la DM et qu'il a vécu des événements stressants au cours du dernier mois (disons un divorce). Mais on peut aussi penser qu'un individu a développé une dépression parce qu'il a intériorisé l'idée selon laquelle ses parents ne l'aimaient pas, a souffert d'isolement social et a perdu récemment son frère — son seul ami — dans un accident de voiture.

Ce schéma de Kendler, à mon avis, peut être considéré comme une reformulation (je l'admets bonifiée de plusieurs facteurs autres qu'« environnement stressant » et « profil génétique vulnérable ») en termes mécanistes de l'explication *IS 2*.

Selon moi, les explications *IS 1* et *IS 2* sont moins adéquates que les explications mécanistes 3 et 4. En effet, comme on peut le constater, les représentations/explications des troubles de santé mentale 3 et 4, en plus de mieux rendre compte de ce qui se fait en recherche en psychiatrie, sont beaucoup plus

⁶⁴ Mentionnons que ces facteurs biologiques sont tirés d'études où sont évaluées les histoires de dépression dans des familles. En effet, ces facteurs biologiques ne font pas référence à des gènes spécifiques corrélés avec la dépression ou à l'activité anormale d'aires cérébrales ou à la quantité de certains neurotransmetteurs chez les gens dépressifs. Cependant, Kendler ne s'objecterait pas, il me semble, à ce que soit ajouté à ce modèle mécaniste des facteurs biologiques de ce genre. D'ailleurs, dans un article récent (Kendler 2012, 12-14), Kendler fait référence à ce type de facteurs quand il donne une explication pluraliste et mécaniste du trouble de la dépendance à l'alcool.

⁶⁵ Murphy considère que ces variables sont des variables causales. De fait, il affirme que « [...] 9 predictor variables (3 independent, and 6 in three levels of intermediate dependency) represent the interactive *causal factors* [of major depression] » (Murphy 2006, 135, mes *italiques*).

informatives que les explications *IS 1* et *IS 2* susmentionnées concernant la dépression majeure. C'est aussi ce que soutiendrait Eronen, il me semble. En effet, celui-ci affirme que

[r]eductive explanations [(and non-reductive explanation, in my opinion)] in psychology and neuroscience [(and psychiatry, I think)] cannot be conceived as logical derivations. Instead, these disciplines typically look for descriptions of *mechanisms* that can serve as explanations for patterns, effects, capacities or phenomena. (Eronen 2012, 19)

Ultimement, en psychiatrie, la raison pour laquelle les explications sont représentées par des *modèles mécanistes* plutôt que par des *structures linguistiques* réside dans le fait que l'essence des théories n'est pas de nature syntaxique dans cette discipline et dans les sciences biologiques en général. Plus précisément, dans ces sciences, les théories ne consistent pas, fondamentalement, en un réseau de propositions formulées dans le langage de la logique des prédicats du premier ordre⁶⁶, mais en une *famille de modèles*. En effet, affirme Schaffner,

⁶⁶ Patrick Suppes va encore plus loin dans ce constat. En effet, il pense que la logique des prédicats du premier ordre est tout simplement insuffisamment puissante pour axiomatiser la plupart des théories scientifiques. De fait, il écrit,

Theories are ordinarily said to have a standard formalization when they are formulated within first-order logic. Roughly speaking, first-order logic is just the logic of sentential connectives and predicates holding for one type of object. Unfortunately, when a theory assumes more than first-order logic, it is neither natural nor simple to formalize it in this fashion. For example, if in axiomatizing geometry, we want to define lines as certain sets of points, we must work within a framework that already includes the ideas of set theory. To be sure, it is theoretically possible to axiomatize simultaneously geometry and the relevant of set theory, but this is awkward and unduly laborious.

Theories of more complicated structure like quantum mechanics, classical thermodynamics, or a modern quantitative version of learning theory, need to use not only general ideas of set theory but also many results concerning the real numbers. Formalization of such theories in first order logic is utterly impractical. Theories of this sort are similar to the theories mainly studied in pure mathematics in their degree of complexity. (Suppes 1967, 58)

A close examination of a wide variety of other biological theories in genetics, immunology, physiology, embryology, and the neurosciences suggests that the typical theory in the biomedical sciences is a structure of overlapping interlevel, causal, and temporal prototypical *models*. (Schaffner 2008b, 1009, *mes italiques*)⁶⁷

Cette notion de *modèle* est au cœur de la *conception sémantique* des théories scientifiques que nous verrons au début du troisième chapitre.

2.7 Conclusion du deuxième chapitre

En somme, ce deuxième chapitre m'a permis de montrer que (a) la réduction interthéorique n'était pas praticable (section 2.5) et que (b) conséquemment, l'interprétation réductionniste des modèles de l'explication DN et IS en regard des *phénomènes naturels* n'était pas viable (voir section 1.3) et que (c) ces modèles en eux-mêmes ne l'étaient pas non plus (sections 2.6.2 et 2.6.3). Je suis même allé plus loin dans ce constat d'échec. En effet, j'ai fait voir que même s'il était formellement possible de construire sans heurts une explication IS⁶⁸ dans ce cadre syntaxique, il resterait un problème. En effet, au-delà de ces difficultés que rencontrent les modèles DN et IS, j'ai montré qu'une explication mettant en jeu des entités linguistiques évoluant dans un réseau nomique est une mauvaise façon de se *représenter* les pathologies en médecine mentale (section 2.6.4). Cette discipline, de fait, contient des descriptions de mécanismes plutôt que des théories formelles constituées de propositions linguistiques. En effet, les explications des maladies mentales, en psychiatrie, sont représentées par des *modèles mécanistes* et non par des arguments déductifs ou inductifs mettant en jeu des énoncés et dont la structure repose sur la

⁶⁷ Ces remarques de Schaffner s'appliquent aussi à la psychiatrie (voir Schaffner 1994, 281).

⁶⁸ Cela aurait aussi pu être une explication DN.

logique des prédicats du premier ordre. Il en est ainsi parce que les théories, en psychiatrie et en sciences biomédicales, ne sont pas des structures formulées dans le langage de la logique des prédicats du premier ordre, mais des familles de *modèles* (voir Schaffner 1994). Pour cela, dans le troisième chapitre, je présente la *conception sémantique* des théories scientifiques dans laquelle ont été développées la notion générale de *modèle* et la notion particulière de *modèle mécaniste* (Glennan 2002, 2005; Machamer, Darden et Craver 2000). Je présente aussi dans ce dernier chapitre le modèle de l'explication *réductionniste* et mécaniste de Bickle (2003, 2006, 2008b) et le modèle de l'explication *pluraliste* et mécaniste de Kendler (2005, 2008a, 2012; Kendler et Campbell 2009). Ultimement, je rejette le modèle de Bickle et je défends celui de Kendler.

CHAPITRE III

L'EXPLICATION RÉDUCTIONNISTE (BICKLE) ET PLURALISTE (KENDLER) DANS LE CADRE SÉMANTIQUE (ET MÉCANISTE) DES THÉORIES SCIENTIFIQUES

3.1 Introduction du troisième chapitre

Au premier chapitre, à la section 1.1, j'ai présenté la *conception syntaxique* des théories scientifiques. Selon cette dernière, une théorie scientifique forme un ensemble ou un réseau axiomatique d'énoncés théoriques (idéalement interprétés empiriquement) et observationnels. Ce réseau est une structure axiomatique exprimée dans le langage de la logique des prédicats du premier ordre. Cette façon de concevoir les théories implique, comme je l'ai montré au deuxième chapitre, une conception de l'explication (et de la réduction) qui pose des problèmes. Qui plus est, d'un point de vue représentationnel, en ce qui concerne l'explication, cette conception des théories ne convient pas à la psychiatrie. Dans ce troisième chapitre, je présente une façon d'appréhender les théories scientifiques qui s'inscrit au sein de la *conception sémantique*. Comme je l'ai dit à la fin du deuxième chapitre, je pense que cette manière de comprendre les théories est plus adéquate pour formuler des explications en médecine mentale. Afin d'explicitier cette thèse, j'expose, d'abord, la conception sémantique des théories scientifiques (sections 3.2). Puisque cette conception repose

sur le concept de modèle, j'expose les différentes manières d'appréhender cette notion suivant Patrick Suppes (1957), Bas C. Van Fraassen (1980, 1987) et Ronald N. Giere (1988, 1999a, 1999b, 2006 et 2010) (section 3.2.1). De ces interprétations du terme « modèle », je crois que celle qu'a développée Giere convient à la représentation des explications en psychiatrie. De fait, pour ce dernier, un modèle est une entité représentationnelle et il se trouve que cette façon de concevoir l'idée générale de modèle est compatible avec l'idée particulière de *modèle mécaniste* (section 3.2.2), lequel type de modèle me semble pertinent pour représenter une explication réductionniste ou pluraliste en psychiatrie (et en neuroscience). J'expose, après, deux *modèles de l'explication* en médecine mentale mettant à profit le concept de modèle mécaniste. D'abord, je présente le modèle mécaniste et réductionniste de Bickle (sections 3.3 et 3.4). Je veux soutenir que ce modèle ne convient pas pour le domaine de la psychiatrie (section 3.5). Je présente ensuite le modèle mécaniste et pluraliste de Kendler qui me semble plus adéquat pour représenter l'explication d'un trouble de santé mentale (sections 3.6 et 3.7). Je termine le chapitre en présentant le modèle interventionniste (MI) de la causalité qu'adopte Kendler (section 3.8).

3.2 La conception sémantique des théories scientifiques

Selon la conception sémantique, « une théorie scientifique doit être décrite comme une classe (ou famille) de modèles » (Vorms 2011, 128).⁶⁹ Plus précisément, selon cette conception, affirme Marion Vorms,

⁶⁹ Selon Frederick Suppe, à l'origine, on retrouve les idées relatives à la conception sémantique des théories scientifiques dans les travaux du mathématicien John von Neumann (Suppe 1989, 3) ainsi que dans ceux du logicien néerlandais Evert Beth (Suppe 1989, 6). Mais le père fondateur de cette conception, selon Marion Vorms (2011, 127), serait plutôt le philosophe des sciences Patrick Suppes (voir, par exemple, Suppes 1957, chapitre 12).

[...] une théorie est une entité extra-linguistique qui ne peut être identifiée à aucune de ses formulations particulières mais correspond à la classe des modèles qui satisfont ses différentes formulations linguistiques. (ibidem, 128)

Les défenseurs de la conception sémantique des théories diront, par exemple, que les formulations newtonienne, lagrangienne et hamiltonienne de la mécanique classique sont identiques parce que la classe des modèles qui les satisfont est la même (ibidem, 158)⁷⁰.

Comme on peut le constater, cette conception des théories scientifiques est très différente de celle qu'adoptent les tenants de l'empirisme logique et que j'ai exposée à la section 1.1 (voir spécialement les points 10 à 12). En effet, comme on l'a vu, ceux-ci soutiennent une conception linguistique des théories, contrairement aux représentants de la conception sémantique qui adoptent ce que l'on pourrait nommer, suivant C. Ulises Moulines (2006, 107), une conception « modélistique ».

3.2.1 Les différentes conceptions du terme « modèle »

Comme je l'ai dit (section 3.2), les avocats de la conception sémantique s'entendent sur l'idée selon laquelle une théorie doit être décrite comme une classe de modèles. Cependant, en plus de diverger sur plusieurs points que je ne discuterai pas ou peu⁷¹, ces derniers ne partagent pas tous les mêmes idées sur ce qu'est un modèle.

⁷⁰ Il est à noter qu'à l'inverse des avocats de la conception sémantique des théories, Vorms pense que ces formulations ne sont pas équivalentes (voir Vorms 2011, section 4.3).

⁷¹ Par exemple, épistémologiquement, certains adoptent le réalisme (par exemple, Giere 1999a; Suppe 1989) et d'autres l'antiréalisme (par exemple, Van Fraassen 1980).

Par exemple, Patrick Suppes (1957) propose que l'on conçoive un modèle comme une structure ensembliste (« set-theoretic structures ») telle que définie dans la théorie des ensembles. Plus spécifiquement, selon la lecture que fait Giere des propos de Suppes, « [...] a model [...] is a set-theoretical structure consisting of set of objects together with properties, relations, and functions defined over a set of objects » (1999b, 42). Dans cette perspective, toute entité qui satisfait une telle structure — laquelle est un prédicat ensembliste — est considérée comme un modèle de la théorie (Vorms 2011, 130-131). Afin de comprendre cette idée, suivant Moulines (2006, 112, note 1), formulons selon le langage de la logique standard et de la théorie des ensembles une mini-théorie des familles (traditionnelles). Voici les axiomes de cette dernière :

$F(X)$ si et seulement si : $\exists P, M, E$:

- (0) $X = \langle P, M, E \rangle$
- (1) $P \neq \emptyset \wedge \|P\| > 2$
- (2) $M \subseteq P \times P \wedge \forall x \forall y (\neg xMx \wedge (xMy \rightarrow yMx))$
- (3) $E \subseteq P \times P \wedge \forall x \forall y (xEy \rightarrow \neg yEx)$
- (4) $\exists x, y, z \in P (xMy \wedge xEz \wedge yEz)$.

Moulines utilise les abréviations « F » pour « ... est une famille »; « P » pour un ensemble de personnes, « M » pour la relation de mariage et « E » pour la relation d'engendrer. Ce dernier spécifie aussi que toute relation dyadique R sera écrite comme ceci : « xRy ».

Grosso modo, cette théorie signifie que X est une famille si X comporte un ensemble de personnes P , lequel n'est pas vide et comporte plus de deux personnes (axiome 1). De plus, cette théorie prétend que X est une famille si ces personnes dans P sont en relation de mariage (M) ou ont été engendrées (E) par les personnes mariées

(axiome 0). Cette théorie des familles suppose aussi que pour toute personne x dans P , x ne peut pas être mariée avec elle-même, mais seulement avec une autre personne y dans P , laquelle personne y est nécessairement mariée avec x (axiome 2). De plus, cette théorie nous indique que si x dans P engendre y dans P , alors il est impossible que y ait engendré x (axiome 3). En outre, cette mini-théorie montre que pour un ensemble de personnes x , y et z dans P , si x est marié avec y , il s'en suit que x a engendré z et y a engendré z (axiome 4). Enfin, selon les axiomes (2), (3) et (4) on peut déduire le théorème selon lequel les trois personnes x , y et z de l'ensemble P impliquées dans le quatrième axiome sont nécessairement différentes.

En somme, de ce qui précède, il faut comprendre que si X satisfait le prédicat ensembliste susmentionné constitué des axiomes (0) à (4), alors X est un modèle de cette théorie hypothétique des familles. Ainsi, selon cette conception sémantique, toutes les entités X satisfaisant le prédicat ensembliste susmentionné formeront la classe des modèles de cette théorie. Par exemple, le modèle représentant un groupe de personnes comportant une femme mariée à un homme ainsi qu'un enfant engendré par ces derniers serait un modèle de cette théorie ou de cette structure ensembliste⁷².

Bas C. Van Fraassen (1980, 1987)⁷³ propose une autre façon de concevoir les modèles. Selon lui, ces derniers sont « des “trajectoires” ou des “régions” dans un *espace d'états* » (Moulines 2006, 124). La notion d'espace d'état provient de la physique (ibidem, 124). Par exemple, en thermodynamique, « l'état d'un système est complètement déterminé par un triplet de nombres réels $\langle p, v, t \rangle$, qui représente respectivement des grandeurs pression, volume et température » (ibidem, 124)

⁷² Il est important de souligner que ce qui est le modèle de cette théorie des familles ou de cette structure ensembliste n'est pas les personnes réelles de cette famille, mais le *modèle représentant les personnes de cette famille* (à ce sujet, voir Suppes 1962).

⁷³ Frederick Suppe (1989) conçoit lui aussi les modèles selon l'approche des espaces d'états.

(souvenons-nous de la loi de Boyle-Mariotte $pV = kT$ vue à section 1.4.4). Aussi ajoute Moulines,

En général, dans les théories physiques, les états d'un système physique sont identifiés par des points dans un système physique de coordonnées qui a le même nombre de dimensions que le nombre de composantes déterminant l'état. À chaque type de système physique correspond ainsi un espace d'états, qui est l'ensemble de toutes les suites possibles de n éléments ($n=3$ pour l'exemple de la thermodynamique). Les états possibles du système physique sont des « points » dans cet « espace ». La tâche des axiomes propres à une théorie déterminée consiste alors à imposer des contraintes sur l'ensemble des suites logiquement possibles, en fixant les suites qui sont acceptables et celles qui ne le sont pas. Les suites acceptables représentent les modèles admis par la théorie. (Ibidem, 124)

Par exemple, selon les suites logiquement acceptables de l'espace d'état d'un système thermodynamique, entre autres choses, nous verrions que (1) si on augmentait la pression en gardant le volume constant, alors la température augmenterait et que (2) si on diminuait le volume en gardant la température constante, alors la pression augmenterait (Bais 2011, 24).

La notion d'espace d'état, selon Mario Bunge, pourrait être utilisée dans toutes les disciplines scientifiques (Bunge 2006, 17). Par exemple, selon lui, une maladie peut être conçue comme un processus ou une *séquence d'états* corporels ayant lieu dans un organisme et ayant un début et une fin (Bunge 2003a, 252). Selon cette conception des pathologies, les suites logiquement possibles du système physique (ici, un corps humain) qu'admet cette théorie de la maladie en question seraient l'espace d'état de cette maladie.

Une troisième conception des modèles est celle de Ronald N. Giere (1988, 1999a, 1999b, 2006). Celui-ci définit un modèle comme « une entité représentationnelle » (Giere 1999a, 5, ma traduction). Les modèles, en effet, sont des outils dont se servent

les scientifiques pour *représenter* le réel. La référence aux agents cognitifs que sont les scientifiques dans la phrase précédente est importante parce que, selon Giere, les modèles ne représentent rien en eux-mêmes (Giere 2006, 63). En effet, ce sont les scientifiques qui, utilisant ces modèles, opèrent l'acte de représenter (ibidem, 64). Cet acte de représenter consiste, pour les chercheurs, à déterminer certaines caractéristiques d'un modèle qui sont, selon eux, *similaires* à l'aspect du monde dont ils tentent de rendre compte. De fait, dit Giere,

One way scientist do this [(the representing)] is by picking out some specific features of the model that are then claimed to be similar in some specific respect to features of the designated real system. It is the possibility of specifying such similarities that makes possible the use of the model to represent the real system in this way. (ibidem, 64)

Il faut ajouter que cet acte de représenter sert toujours un objectif du ou des chercheurs (Giere 2010). En d'autres termes, l'acte de représenter implique l'intentionnalité d'agents cognitifs. Cette idée de modèle s'inscrit ainsi dans une conception intentionnaliste de la représentation scientifique (« intentional conception of scientific representation ») (ibidem, 274). Giere résume son propos à ce sujet de la manière suivante : « Agents (1) intend; (2) to use model, M; (3) to represent a part of the world, W; (4) for some purpose, P » (ibidem, 269). Pour finir, j'ajouterais que cette représentation du réel que nous donnent les modèles construits par des agents, à l'instar des cartes géographiques, est partielle (Giere 2006, 73).

Il existe plusieurs types de modèles représentationnels. D'une part, il y a les modèles matériels (Giere 1999b, 44-48). Parmi ceux-ci, on retrouve des objets comme la carte d'une ville, les diagrammes (par exemple, le diagramme d'un circuit électrique) et les maquettes (« scale model », par exemple, la maquette représentant la double hélice de l'ADN fait par Jim Watson).

D'autre part, il y a les modèles abstraits (ibidem, 48-50). Par exemple, l'équation $y = ax + b$ est un modèle abstrait (ibidem, 48). On pourrait évidemment dire qu'il est aussi un modèle matériel parce qu'il consiste en des lettres et des signes sur du papier ou sur un écran d'ordinateur. Mais, fondamentalement, ce modèle est abstrait parce que l'équation (qui est ce modèle) présente une relation abstraite. Giere affirme que celle-ci est abstraite parce qu'elle peut être exprimée de plusieurs manières. Par exemple, comme le dit celui-ci, on peut l'exprimer aussi comme ceci : $y - ax - b = 0$. En effet, la relation exprimée dans cette équation est identique à celle exprimée dans l'équation $y = ax + b$. Giere nomme ce genre de modèles abstraits les *modèles purement mathématiques* (« pure mathematical model ») (ibidem, 49). Comme l'explique ce dernier, on peut représenter le modèle abstrait $y = ax + b$ (une relation linéaire) par un modèle matériel comme celui-ci :

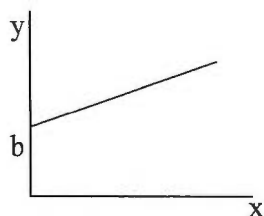


Figure 4 Modèle matériel représentant le modèle abstrait $y = ax + b$ (graphique tiré de Giere 1999b, 49).

Notons, par ailleurs, qu'il est possible d'utiliser la structure de ces *modèles purement mathématiques* pour en faire des *modèles mathématiques appliqués* (« applied mathematical models »). Pour ce faire, il faut remplacer les éléments mathématiques du modèle purement mathématique par des termes désignant des objets abstraits et désignant des relations entre ces objets. Par exemple, on peut transformer l'équation $y = ax + b$ en l'équation $d(t) = vt + d_0$ où y devient la distance parcourue par un point depuis un point fixe et où x devient le temps t que prend ce

point à se déplacer depuis un moment arbitraire de départ (qui peut être zéro) et où a devient la vitesse v de ce point mobile et où b devient la distance initiale (d_0) de ce point relativement à l'origine [de ce plan cartésien] (ibidem, 49). Comme le dit Giere, on pourrait rendre cette équation encore plus spécifique si l'on supposait qu'elle représente le mouvement linéaire d'une auto partant à un temps zéro d'une intersection à une vitesse v et à une distance d_0 de l'origine dans le plan cartésien. Suivant cela, ce *modèle mathématique appliqué* serait le *modèle d'une auto et d'une intersection* où cette auto, au sein du modèle, se déplace de manière parfaitement linéaire à une vitesse constante depuis une intersection (ibidem, 49). Pour n'importe quel temps t , la distance parcourue par l'auto depuis cette intersection idéale est effectivement donnée par l'équation $d(t) = vt + d_0$. Cependant, note Giere, cette équation n'est vraie que pour ce modèle. On ne peut pas dire, en effet, que cette équation est vraie pour une auto réelle. De fait, aucune auto réelle ne se déplace de manière parfaitement rectiligne à une vitesse constante. Pour cette raison, Giere propose que nous introduisions des marges d'erreur (des deltas) dans les termes de l'équation $d(t) = vt + d_0$. Suite à cet ajout, on aura alors l'équation suivante :

$$d \pm \delta_d d = (v \pm \delta_v v) (t \pm \delta_t t) + (d_0 \pm \delta_d d_0).$$

On pourrait dire de cette nouvelle équation comportant des marges d'erreur qu'elle peut être vraie d'une auto réelle, affirme Giere (ibidem, 50). Cependant, comme le note ce dernier, bien que cette façon d'interpréter les modèles abstraits en science soit acceptable, elle n'est pas la meilleure (ibidem, 50). En effet, fait-il remarquer, habituellement, les marges d'erreur n'apparaissent dans les calculs que lorsque les prédictions théoriques ont été comparées aux mesures réelles [(pas avant, comme dans l'exemple susmentionné)] (ibidem, 50). Cette façon de faire plus habituelle en science, poursuit-il, nous indique que l'équation originale, à savoir celle qui ne présente pas de marges d'erreur, ne réfère pas à une chose réelle, mais à un

modèle abstrait qui satisfait par définition cette équation (ibidem, 50). Or, quand vient le temps de comparer le modèle abstrait avec la réalité, suggère Giere, les deltas pourraient être interprétés comme spécifiant le degré de similarité (qu'il soit conjecturé ou réel) existant entre le modèle abstrait et ce système réel (ibidem, 50). Ainsi, dans cette optique, on peut dire que la modélisation mathématique, ultimement, consiste à construire un modèle idéalisé et abstrait que l'on peut ensuite comparer à un système réel afin de déterminer le degré de similarité que ce modèle entretient avec ce système (ibidem, 50).

Outre les *modèles purement mathématiques* et les *modèles mathématiques appliqués*, il existe ce que Giere nomme les *modèles théoriques*. Ceux-ci sont des modèles abstraits construits à l'aide de *principes théoriques*. Ces principes sont, par exemple, les lois de Newton en mécanique classique, l'équation de Schrödinger en mécanique quantique, les principes de la relativité, le principe de l'évolution par la sélection naturelle ou les lois mendéliennes en génétique (ibidem, 51). Selon Giere, il faut concevoir ces grands principes théoriques non comme des nécessités naturelles ayant lieu dans le réel, mais comme des définitions. Un modèle théorique sera alors un système idéal qui satisfait ces définitions. Par exemple, tous les systèmes idéels tridimensionnels où deux corps bougent selon les lois de Newton sont des modèles théoriques newtoniens. Les lois de Newton, dans cette perspective, ne décrivent aucunement des corps réels (ibidem, 51). On peut déduire de cela qu'un modèle théorique newtonien est vrai non parce qu'il est confirmé empiriquement, mais par définition. En effet, ce modèle est vrai dans la mesure où les entités abstraites qu'il contient évoluent conformément aux lois de Newton. Toutefois, on peut utiliser les modèles théoriques pour expliquer les phénomènes empiriques. Par exemple, comme l'explique Giere (ibidem, 51-52) on peut utiliser le modèle théorique newtonien pour expliquer des phénomènes empiriques si on identifie, par exemple, les corps théoriques de ce modèle à des corps réels comme la Lune et la Terre. Une fois que

ces identifications seront faites, on pourra formuler des hypothèses empiriques (suivant ce modèle théorique newtonien) et vérifier si la Lune et la Terre se comportent de *manière similaire* aux corps abstraits évoluant dans le modèle théorique newtonien.

De ces trois façons de concevoir les modèles, je pense que la plus appropriée pour la psychiatrie est celle qu'a élaborée Giere. Je soutiens cela parce qu'à l'heure actuelle, en pratique, comme on l'a vu à la section 2.6.4, les chercheurs en psychiatrie (et en neurosciences) s'expliquent la plupart du temps les troubles de santé mentale (et les processus cérébraux normaux comme anormaux) à l'aide de *représentations* — c'est-à-dire à l'aide de *modèles représentationnels* — de mécanismes physiques. En effet, donner une explication, dans ce domaine, ne consiste pas à élaborer, comme le ferait Suppes, un modèle de maladie mentale qui satisfait une structure ensembliste. Cela ne consiste pas, non plus, à déterminer, comme le ferait Van Fraassen ou Bunge, les suites logiquement possibles d'états d'une pathologie dans un espace mathématique abstrait représentant les espaces de ces états. En effet, donner une explication en psychiatrie consiste plutôt à construire un modèle qui *représente* selon un degré de *similarité* maximale une maladie réelle, c'est-à-dire le *mécanisme physique* ayant lieu dans le monde et dont résulte ce trouble de santé mentale.

À mon avis, ce modèle qui représente ce mécanisme physique pathologique ayant lieu dans le monde peut être décrit comme un *modèle mécaniste* tel que défini par Stuart Glennan (2002, 2005). Par ailleurs, il faut noter que ce *modèle mécaniste* n'est pas un *modèle théorique* au sens de Giere. En effet, contrairement aux explications que l'on retrouve en mécanique classique, par exemple, il n'existe pas de grands principes en médecine mentale sur lesquels on pourrait assoir nos explications psychiatriques. Il me semble plutôt que, suivant la typologie des modèles de Giere (1999b), les *modèles mécanistes* que l'on retrouve en psychiatrie sont souvent (a) des

diagrammes, lesquels représentent des mécanismes sans mettre à profit des outils mathématiques (voir, par exemple, à la section 2.6.4, Insel 2010) ou (b) des *modèles mathématiques appliqués* où les facteurs représentant les constituants du mécanisme sont reliés, par exemple, par des relations statistiques (voir, par exemple, à la section 2.6.4, Kendler 1993 et coll. cités dans Murphy 2006).

Dans la prochaine section, j'explique ce qu'est un mécanisme et ce qu'est un modèle mécaniste. Je décris, aussi, en quoi consisteraient, dans ce cadre mécaniste, une *explication réductionniste* et une *explication pluraliste* en psychiatrie.

3.2.2 De ce que sont un mécanisme et un modèle mécaniste

Chez les philosophes contemporains, les défenseurs de l'*explication mécaniste* sont nombreux (par exemple, voir Bechtel 2008; Bechtel et Richardson 2010 [1993]; Bunge 1998 [1967], 28-49, 2003; Craver 2007; Glennan 1996, 2002, 2005; Machamer, Darden et Craver 2000; Thagard 2006). De manière générale, selon eux, on obtient une *explication mécaniste* lorsqu'on réussit à construire le *modèle mécaniste* d'un *mécanisme*. Dans la littérature en philosophie des sciences, plusieurs auteurs ont tenté de définir ce qu'est un mécanisme. Comme le note Stuart Glennan (2008), les deux définitions suivantes sont souvent citées :

Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions. (Machamer, Darden et Craver 2000, 3)

A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations. (Glennan 2002, S344)

Selon ces définitions, les mécanismes partagent quatre caractéristiques communes (Glennan 2008, 377-378, ma traduction) :

- 1) *Les mécanismes produisent des phénomènes ou des comportements.* Il est à noter qu'un mécanisme peut produire plus d'un comportement. En conséquence, les parties, les activités des parties, les interactions entre les parties et les balises du système mécaniste que l'on identifie lors d'une recherche sont choisies en fonction du phénomène (ou du comportement) que l'on décide d'expliquer ou, dit autrement, de modéliser (ibidem, 377). Prenons un exemple pour illustrer cette idée : un système cérébral produit plusieurs propriétés comme les fonctions cognitives mnémoniques et langagières. Aussi, il génère parfois des propriétés que l'on juge anormales comme, par exemple, les symptômes d'une dépression majeure ou d'une schizophrénie. Or, on décomposera le cerveau en plusieurs parties et on ciblera l'activité de ces parties et les interactions qu'elles entretiennent en fonction de la propriété que l'on voudra expliquer ou modéliser. Ainsi, on ne décomposera pas le cerveau de la même façon selon que l'on décide de rendre compte des fonctions mnémoniques ou des symptômes de la schizophrénie.

- 2) *Les mécanismes sont composés de parties qui forment une structure.* Les parties des mécanismes peuvent aussi être appelées « entités » ou « objets ». Les parties des mécanismes présentent des propriétés qui sont relativement stables dans le temps et qui peuvent être, *au moins en théorie*, l'objet de manipulation ou d'isolation du reste du système mécaniste (ibidem, 378). Par exemple, selon une perspective réductionniste, Insel (2010) formule l'hypothèse selon laquelle une manipulation de l'activité défectueuse de l'aire cérébrale 25 d'un individu souffrant d'une DM pourrait rétablir le mécanisme

normal du cerveau et, ce faisant, faire disparaître les symptômes de ce trouble de santé mentale. Par ailleurs, il est plausible de penser que Kendler, un défenseur du pluralisme, pourrait formuler l'hypothèse selon laquelle une augmentation du support social ⁷⁴ donné à une personne dépressive augmentera les chances de rémission de cette dernière. Enfin, il faut noter qu'un système mécaniste n'est pas identifiable à la somme des parties qui le composent, mais à la manière dont ces parties sont organisées et à la façon dont celles-ci se déploient dans le temps. En effet, affirme Glennan :

[...] [m]echanisms are individuated not simply by what parts they have, but by how those parts are organized. A heap of parts does not make a mechanism. Rather the characteristic spatial, temporal, and functional organization of the parts explains the behavior of the mechanism. (Glennan 2008, 378)

- 3) *Les mécanismes présentent un comportement régulier, mais pas parfaitement régulier.* Comme le fait remarquer Glennan (ibidem, 378), parce que les mécanismes présentent des parties organisées de manière stable, ces parties *interagiront* de manière stable et régulière et produiront avec régularité des comportements précis ou des propriétés précises. Par exemple, si j'appuie sur le bouton « démarrer » de mon ventilateur, les hélices de ce dernier se mettent à tourner. Glennan, suivant Woodward (2003), qualifie ce type d'interaction entre ces deux parties de *généralisation invariante* (Glennan 2002, S344)⁷⁵. Mais, remarque ce dernier (Glennan 2008, 378), ces interactions donnant lieu à ces propriétés ou comportements peuvent être altérées en raison de perturbations des parties du système mécaniste ou de l'environnement dans

⁷⁴ Comme Kendler, je fais l'hypothèse que le « support social » est une *partie* qui (a) présente des propriétés et qui (b) est une entité du mécanisme à l'origine de la dépression.

⁷⁵ Voir aussi la section 3.8.1 de ce mémoire au sujet du concept de généralisation.

lequel évolue ce dernier. En conséquence, les propriétés ou comportements engendrés par le système mécaniste subissant ces dérèglements seront altérés. Par exemple, un être humain (un système mécaniste biologique) manifeste en général ce que l'on pourrait appeler un élan vital ou un goût pour la vie, mais un épisode de stress extrême (par exemple, une humiliation publique, un divorce et la mort d'un proche en même temps) peut causer (au moins partiellement) une perte du goût de vivre chez un individu. De la même façon, les souris (des systèmes mécanistes biologiques) présentent normalement une fonction cognitive appelée « mémoire à long terme », mais cette mémoire à long terme disparaît lorsqu'on manipule le génome de ces souris de telle sorte qu'elles ne produisent plus de protéines CREB (à ce sujet, voir la section 3.3.1 de ce mémoire).

- 4) *Les mécanismes sont hiérarchiquement organisés*⁷⁶. Glennan (ibidem, 378) explique que les mécanismes sont hiérarchiquement organisés parce que (a) les parties d'un mécanisme sont souvent elles-mêmes des systèmes mécanistes et parce que (b) l'existence des activités entre ces parties peut impliquer l'activité d'autres mécanismes. Pour illustrer cette idée, considérons le système (mécaniste) cérébral. Illustrons d'abord l'énoncé (a) : le cerveau est composé de plusieurs parties interagissant et produisant des pensées, des émotions et des comportements; ces parties cérébrales sont elles-mêmes des

⁷⁶ Je traduis l'énoncé de Glennan (2008, 378) « Mechanisms are hierarchical » par l'énoncé « Les mécanismes sont hiérarchiquement organisés ». Cependant, je ne comprends pas pourquoi Glennan utilise le terme « hierarchical ». En effet, comme on le voit suivant Glennan dans ce quatrième point, il me semble que l'existence des mécanismes que l'on retrouve à un niveau d'organisation quelconque (par exemple, organique) suppose ou implique nécessairement des mécanismes ayant lieu à un niveau plus bas (par exemple, cellulaire) *et* des mécanismes ayant lieu à un niveau d'organisation plus haut (par exemple, l'environnement social). Par exemple, si le système cérébral doit son existence aux cellules qui le composent, il la doit aussi au monde qui l'entoure. Comme le dit Bunge (2006, 252), « [t]he brain could not exist without the world surrounding it, from which it draws not only nourishment but also stimulation required for its normal development and functioning ».

systèmes mécanistes formés de parties (entre autres les neurones) interagissant entre elles lesquelles sont elles-mêmes des systèmes mécanistes composés de parties (entre autres de protéines) interagissant les unes les autres lesquelles sont composées d'autres parties (entre autres les acides aminés) interagissant entre elles et ainsi de suite. Cette intégration, note Glennan (*ibidem*), peut aussi aller vers l'extérieur du système mécaniste considéré [(ici le cerveau produisant des pensées, émotions et comportements)]. En effet, on pourrait ajouter que le cerveau est une des parties du système mécaniste qu'est le corps humain, lequel fait partie d'un système environnemental et ainsi de suite. Suivant cela, l'idée (b) selon laquelle l'existence des activités entre ces parties (du cerveau) peut impliquer l'activité d'autres mécanismes peut être illustrée par le fait qu'un cerveau ne fonctionne adéquatement que dans un système environnemental dont la température (une propriété de ce système) est propice au fonctionnement cérébral. Dans un environnement trop froid, par exemple, le cerveau (et toutes ces parties) cesse d'opérer. Bref, comme nous pouvons le constater, (1) chacune des parties composant le cerveau sont des systèmes mécanistes occupant, si l'on peut dire, une position ou un certain niveau d'organisation de la réalité (par exemple, les niveaux organique, cellulaire, biomoléculaire, etc.) et (2) le cerveau est contenu dans d'autres systèmes occupant eux-mêmes un niveau d'organisation de la réalité (par exemple, le niveau individuel, le niveau environnemental). Je pense que c'est en ce sens que Glennan affirme que « les mécanismes sont hiérarchiquement organisés » (*ibidem*).

J'adopte la définition du terme « mécanisme » formulée par Glennan parce que la définition qu'ont élaborée Peter K. Machamer, Lindley Darden et Carl F. Craver (2000) est problématique. Elle est inadéquate, selon Paul Thagard (2006), parce que l'idée selon laquelle un mécanisme présente un début et une fin n'est pas exacte

(ibidem, 7). De fait, plusieurs mécanismes ne présentent pas un début et une fin précise. Pour faire valoir son idée, Thagard donne l'exemple simple du mécanisme constitué des parties « thermostat », « appareil de chauffage central » et « température ambiante ». Un thermostat, affirme-t-il, est un dispositif qui (a) détecte le moment où la température tombe sous un certain seuil et qui, à ce moment, (b) « signale » à l'appareil de chauffage central de se mettre en marche afin que la température ambiante augmente (ibidem). Pris individuellement, explique Thagard, le thermostat et l'appareil de chauffage central sont des machines unidirectionnelles qui peuvent être décrites par les règles simples suivantes : si la température est moins que X , alors un signal dit à l'appareil de chauffage central de se mettre en marche; si un signal « disant » de commencer est reçu, alors l'appareil de chauffage central se met en marche. Prises ensemble, cependant, ces parties constituent un mécanisme en boucle (un « feedback mechanism ») dont la fonction est de maintenir la température ambiante à un certain niveau. Dans cette perspective, explique Thagard (ibidem), ce mécanisme ne présente pas de début ou de fin, mais un processus continu de régulation thermique. Les systèmes mécanistes dont sont issus les symptômes des maladies mentales, à mon avis, présentent souvent ce type de dynamique. Par exemple, si une personne entretient des idées négatives et culpabilisantes, alors son estime d'elle-même diminuera. Et plus son estime d'elle-même diminuera, plus ses pensées négatives et culpabilisantes se maintiendront dans son esprit.

Maintenant que nous avons pris connaissance de la notion de mécanisme et que j'adopterai celle de Glennan, voyons en quoi consiste un *modèle mécaniste* : un modèle mécaniste — lequel est un *modèle représentationnel* au sens de Giere, il faut s'en souvenir — représente un mécanisme. Glennan définit plus précisément ce qu'est un modèle mécaniste dans le passage suivant :

A mechanical model is a description of a mechanism, including (i) a description of the mechanism's behavior; and (ii) a description of the mechanism which accounts for that behavior. (Glennan 2002, S347)

Si les philosophes, chercheurs et praticiens en médecine mentale s'entendent généralement sur (i), c'est-à-dire sur la description du comportement (ou du phénomène) résultant d'un mécanisme jugé pathologique, ils ne s'entendent pas sur (ii), c'est-à-dire sur les mécanismes (ou les facteurs) qui causent ce comportement. Par exemple, si la plupart d'entre eux s'entendent sur l'idée selon laquelle une dépression majeure se manifeste, chez un individu, par divers symptômes comme une perte d'entrain, une augmentation de l'anxiété et par de l'insomnie ou de l'hypersomnie, ils ne s'entendent pas sur les mécanismes causaux qui sont responsables de l'avènement de ce comportement dépressif. Autrement dit, ceux-ci s'expliquent une même maladie mentale selon des explications mécanistes différentes et, par voie de conséquence, par des modèles mécanistes différents. En effet, comme nous l'avons vu à la section 2.6.4, d'un côté, certains soutiennent, comme Insel (ou Kandel), que les seuls mécanismes qui causent les troubles de santé mentale sont essentiellement les mécanismes cérébraux et, ultimement, les mécanismes biomoléculaires. Ce type de *modèle mécaniste* que je nomme *modèle mécaniste réductionniste* est adopté, chez les philosophes, comme je l'ai dit, par Bickle (2003, 2006, 2007, 2008a, 2008b). Par ailleurs, certains avancent, comme Kendler (2005, 2008a), que les troubles de santé mentale sont causés par des mécanismes moléculaires et cérébraux, mais aussi par des mécanismes psychologiques et environnementaux. Pour cela, je dirais que Kendler adopte un *modèle mécaniste pluraliste*. Plusieurs philosophes défendent ce type d'explication en psychiatrie (Faucher 2006b; voir aussi Faucher, Poirier et Lachapelle, 2006; Gold 2009; Kirmayer et Gold 2012; Mitchell 2008a, 2009; Murphy 2006, 2008, 2009, 2010b; Schaffner 1994, 2008a; Thagard 1998, 2008; Thagard, et Findlay, 2012; et Woodward 2008a).

Dans les prochaines sections, je présente et rejette le modèle de l'explication mécaniste et réductionniste de Bickle et défends le modèle de l'explication mécaniste et pluraliste de Kendler.

3.3 Le modèle de l'explication mécaniste et réductionniste de Bickle

Les recherches de Bickle (2003, 2006, 2007, 2008a, 2008b) en philosophie des neurosciences qui ont mené à la formulation de son modèle mécaniste et réductionniste de l'explication ont été largement inspirées par les recherches actuelles en neurobiologie cellulaire et moléculaire de la mémoire. En effet, Bickle trouve dans ces travaux de la matière à penser qui lui permet de donner une nouvelle signification à la réduction, une signification qui, cette fois-ci, ne s'inscrit pas dans le cadre déductif et syntaxique formulé par Hempel, Nagel et Schaffner (voir chapitres 1 et 2), mais plutôt dans la pratique scientifique de laboratoire (Bickle 2003, 31). Selon Bickle, cette *conception de la réduction* que l'on peut aussi appréhender comme une *explication réductive* est particulièrement explicite dans ce passage du livre *Principles of Neural Science* (2000) où l'on apprend que les neuroscientifiques, grâce à leurs recherches et à leurs pratiques méthodologiques réductionnistes, prétendent être sur la voie d'expliquer — c'est-à-dire de réduire — l'esprit par les molécules :

[...] les sciences neuronales *cherchent à relier les molécules à l'esprit* — comment les protéines responsables de l'activité des cellules nerveuses prises individuellement sont liées à la complexité des processus neuronaux. Aujourd'hui, *il est possible de lier la dynamique moléculaire des cellules nerveuses prises individuellement aux représentations et aux actes moteurs et perceptuels du cerveau, et de relier ces mécanismes internes au comportement observable.* (Eric Kandel, James Schwartz et Thomas Jessell [2000] cités dans Bickle 2008a, p.183, traduction de Dominique Boucher, mes *italiques*)

Bickle retient des propos de Kandel et de ses collègues l'idée selon laquelle il est possible, ultimement, *d'expliquer* l'esprit et ses fonctions essentiellement et directement par des processus cellulaires et biomoléculaires. Selon lui, en effet, on doit idéalement chercher à donner une explication de l'esprit sans l'aide des langages de la psychologie, des sciences cognitives et des neurosciences cognitives. Ce n'est pas que ces sciences ne sont pas utiles. Selon lui, elles ont, en effet, un rôle *heuristique* parce qu'elles nous aident à générer et à tester des hypothèses neurobiologiques (Bickle 2003, 114). En dernière instance, cependant, en aucun cas ces sciences ne pourront nous fournir une *explication* adéquate de l'esprit. Seulement la biologie moléculaire peut le faire et, ultimement, la microphysique (Bickle 2003, 156-157; 2006, 431-432).

La marche à suivre réductionniste que tire Bickle des méthodologies et techniques d'investigation en neurosciences cellulaires et moléculaires et qu'il souhaite voir appliquer dans tous les domaines de recherches voulant expliquer/réduire l'esprit — donc en psychiatrie — est la suivante : (Bickle 2006, 425, ma traduction libre) :

- Effectuer une/des intervention (s) causale (s) au niveau de l'activité cellulaire ou [au niveau] des processus moléculaires ayant lieu dans des neurones spécifiques (les animaux génétiquement modifiés [c.-à-d. mutants], par exemple, [résultent de ce type d'intervention] [...]);
- Observer les effets de ces interventions, dans un cadre expérimental, en utilisant des protocoles acceptés en psychologie expérimentale et utilisés pour étudier des comportements.

Suivant cette méthodologie, si la quantité des comportements liés à une fonction cognitive est altérée de manière statistiquement significative suite à l'intervention (par exemple, si la fréquence des comportements augmente ou diminue de manière

significative ou devient nulle), Bickle dira que l'on a *expliqué*, *réduit* ou *fait connaître* le processus cellulaire et moléculaire responsable de cette fonction de l'esprit ou de cette espèce psychologique. De fait, ce dernier écrit :

This methodology constitutes an implicit condition on *explanation* in this field [(les neurosciences cognitives cellulaires et biomoléculaires)]. One only claims a successful *explanation*, a successful *search for a cellular or molecular mechanism*, or a successful *reduction*, of a psychological kind when one successfully intervenes at the lower level and then measures a statistically significant behavioral difference. (Bickle 2006, 425)

La figure suivante illustre ce réductionnisme impitoyable que défend Bickle (les flèches hachurées indiquent le lieu de l'intervention et la flèche pleine indique le lieu où l'on mesure les résultats de l'intervention [au niveau des comportements]) :

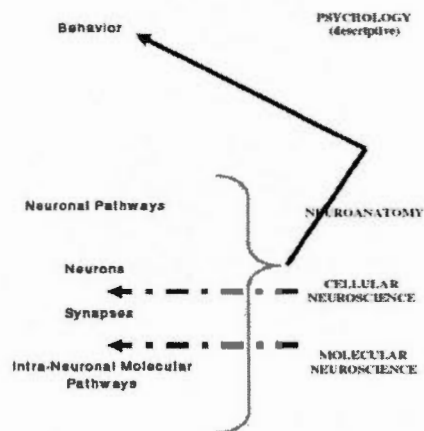


Figure 5 Le réductionnisme impitoyable (schéma tiré de Bickle 2006, 426 cité dans Eronen 2010, 22).

Après avoir fait une méta-analyse des pratiques expérimentales utilisées en cognition cellulaire et moléculaire, Bickle et le neurobiologiste Alcino J. Silva ont raffiné cette méthodologie réductionniste (Bickle 2008b, 470). En effet, ils ont tiré de

cette méta-analyse quatre conditions devant être respectées lorsqu'on donne une explication en termes de mécanismes moléculaires (et cellulaires) d'un phénomène cognitif (ibidem, 470). Voici ces conditions (ibidem, 470) :

- *Observation* : observed occurrences of the mechanism must be strongly correlated with those of the behavioral measures used to indicate the cognitive phenomenon in controlled experimental settings.
- *Negative alteration*: decreasing the probability of the hypothesized mechanism must reliably decrease the probability of the behavioral measures used to indicate the cognitive phenomenon.
- *Positive alteration*: increasing the probability of the hypothesized mechanism must reliably increase the probability of the behavioral measures used to indicate the cognitive phenomenon.
- *Theoretical integration*: available experimental evidence about the hypothesized mechanism must link up with available experimental evidence about the behavioral measures used to indicate the cognitive phenomena so that a plausible account exists that explains how the hypothesized mechanism could be part of the causal nexus producing those behavioral effects.

Un des cas dont se sert Bickle pour exemplifier cette méthodologie est celui où il montre qu'il est possible de réduire (et même remplacer) les modèles cognitifs de la consolidation de la mémoire sociale chez les souris par des modèles cellulaires et moléculaires (Bickle 2008b).

Dans la prochaine section, suivant Bickle (ibidem, 471), j'utilise des études portant sur la transmission sociale des préférences alimentaires (TSPA) chez les souris afin d'exemplifier trois des quatre conditions susmentionnées.

3.3.1 Le rôle des protéines CREB dans la transmission sociale des préférences alimentaires (TSPA) chez les souris

Nous savons plusieurs choses sur la mémoire. D'abord, des résultats d'études portant sur la limace de mer *Aplysia*, sur les drosophiles et sur les souris nous indiquent que l'activation des protéines CREB⁷⁷ est nécessaire à l'activation des processus moléculaires permettant la mémoire à long terme (Kogan et coll. 1997, 1). Par ailleurs, nous savons que les animaux apprennent et retiennent rapidement les indices sociaux olfactifs (Bickle 2008b, 471). Par exemple, les rongeurs développent une préférence naturelle pour les aliments qu'ils sentent dans l'haleine de leurs congénères (ibidem, 471). Tenant compte de cela, afin de valider à nouveau la thèse selon laquelle les protéines CREB présentes chez les souris sont essentielles à la consolidation mnésique, Kogan et ses collaborateurs ont, entre autres, fait une expérience sur la TSPA chez ces rongeurs (Kogan et coll. 1997, 5-7). L'expérience allait comme suit. Ces chercheurs ont logé dans une cage des souris « démonstratrices » qu'ils ont nourries avec de la nourriture sentant la cannelle ou le cacao. Dans une autre cage, ils ont logé des souris « observatrices » privées de nourriture. Ces souris observatrices étaient de deux types : des souris normales et des souris génétiquement modifiées. Le génome de ces dernières avait été altéré de telle sorte qu'elles ne puissent pas produire de protéine CREB. À un certain moment, les chercheurs ont introduit les souris observatrices dans la cage des souris démonstratrices — où il n'y avait alors plus de nourriture — afin qu'elles interagissent et que les premières sentent l'haleine des secondes. Puis, Kogan et ses collaborateurs ont retiré les souris observatrices de la cage des souris démonstratrices. Une minute ou vingt-quatre heures après cette interaction, ils ont présenté aux souris observatrices de la nourriture sentant la cannelle ou le cacao et de la nourriture ne

⁷⁷ L'abréviation « CREB » signifie « *cAMP response element binding proteins* ».

sentant pas ces odeurs. Ces chercheurs ont alors observé la chose suivante : (a) une minute après l'interaction, les souris observatrices normales *et* modifiées ont mangé deux fois plus de nourriture sentant la cannelle ou le cacao que de nourriture ne sentant pas ce parfum. Cependant, (b) vingt-quatre heures plus tard, alors que les souris observatrices normales ont mangé deux fois plus de nourriture sentant la cannelle ou le cacao que de nourriture ne sentant pas la cannelle ou le cacao, les souris observatrices modifiées n'ont pas affiché de préférence pour l'une ou l'autre des nourritures. Que peut-on conclure de cela? On peut inférer que la *mémoire à court terme* des souris observatrices modifiées fonctionne bien, mais que leur *mémoire à long terme* ne fonctionne pas. Ce résultat fait dire à Bickle que cette étude constitue une réussite expérimentale respectant la condition « *negative alteration* » (Bickle 2008b, 471) (voir section 3.3 pour la définition de cette condition). De fait, comme on peut le voir, quand les protéines CREB sont inopérantes, la fréquence des comportements liés à la TSPA devient nulle. Or, ce fait conforte partiellement l'hypothèse susmentionnée selon laquelle l'activation des protéines CREB est nécessaire à l'activation des processus moléculaires permettant la mémoire à long terme (Kogan et coll. 1997, 1).

Colombo et ses collaborateurs ont aussi tenté de confirmer cette hypothèse (Brightwell, Smith, Countryman, Neve et Colombo, 2005). Le protocole de l'expérience de ces derniers est très semblable à celui de celle de Kogan et coll. (1997). Cependant, il y a une différence notable entre le protocole de ces deux recherches. En effet, dans l'étude menée par Colombo (2005), les souris observatrices n'ont pas été modifiées génétiquement avant leur développement par le biais d'une manipulation de leur génome (comme c'était le cas pour les souris de l'étude de

Kogan et coll. 1997), mais en injectant *dans leur hippocampe*⁷⁸ un gène codant pour des protéines CREB mutantes, lesquelles ont pour fonction d'entrer en compétition avec les véritables protéines CREB et de les rendre inopérantes. Malgré cette différence dans la manière d'annuler l'action des protéines CREB chez un groupe des souris observatrices, les résultats issus de l'étude de Colombo et de l'étude de Kogan et coll. (1997) sont pratiquement les mêmes. En effet, suivant les résultats de l'expérience de Colombo, comme dans l'étude de Kogan et coll. (1997), les souris modifiées présentent une mémoire à court terme normale, mais une mémoire à long terme défectueuse. Dans l'étude menée par Colombo, de fait, les souris observatrices modifiées, après avoir senti l'haleine de cannelle ou de cacao des souris démonstratrices, ont mangé beaucoup plus de nourriture sentant ce parfum après une minute, mais, après vingt-quatre heures, elles n'ont pas mangé plus de nourriture sentant la cannelle ou le cacao que de nourriture ne présentant pas ces arômes. En d'autres termes, les souris modifiées, à nouveau dans cette autre expérience, n'affichaient pas de préférence alimentaire après vingt-quatre heures. De ce comportement, il faut conclure que ces souris présentaient une mémoire à long terme défectueuse. Vu cela, selon Bickle, cette étude constitue une autre réussite expérimentale respectant la condition « *negative alteration* » susmentionnée (Bickle 2008b, 471). Mais ce n'est pas tout. En effet, cette étude, selon Bickle (ibidem, 471), remplit aussi la condition « *theoretical integration* » (c'est ce qui la rend intéressante, il me semble, aux yeux de ce dernier). Elle remplit cette condition parce que les résultats de cette expérience convergent avec ceux d'autres études montrant que les souris dont l'hippocampe a été retiré (suite à une ablation) présentaient une mémoire à long terme sérieusement défectueuses. De fait, ces chercheurs écrivent :

⁷⁸ Dans l'expérience de Colombo, c'est donc seulement dans l'hippocampe des souris que les protéines CREB deviennent inopérantes (voir Brightwell, Smith, Countryman, Neve et Colombo, 2005). Cela est différent de l'expérience de Kogan et coll. (1997) où il n'y avait plus de protéines CREB *dans l'organisme entier* de la souris.

Our results are consistent with previous lesion research, in which the social transmission of food preference was used as a behavioral paradigm for studying nonspatial hippocampal- dependent memory. Intact short-term memory, but impaired long-term memory, has been reported on several occasions following pretraining lesions to the hippocampus. (Brightwell, Smith, Countryman, Neve et Colombo, 2005, 14)

Enfin, suivant Bickle (2008b, 471), il est intéressant de mentionner brièvement que, dans une autre étude menée par ce même Colombo (Countryman, Orlowski, Brightwell, Oskowitz et Colombo, 2005), la condition « *observation* » a été respectée. Celle-ci a été rencontrée parce que Colombo et ces collaborateurs ont réussi à montrer que des rats observateurs normaux (non modifiés), après l'apprentissage lié à la TSPA, présentaient, dans leur hippocampe ventral, une augmentation de cellules contenant des protéines CREB phosphorylées (on considère qu'il y a une augmentation parce que le nombre de cellules contenant des protéines CREB phosphorylées *avant* l'apprentissage est plus petit que le nombre de cellules contenant des protéines CREB phosphorylées *après* l'apprentissage). La phosphorylation⁷⁹ des CREB est une preuve que ces dernières sont impliquées dans la mémoire à long terme parce qu'une CREB qui est phosphorylée est une CREB qui peut se fixer sur un site CRE d'un gène, laquelle fixation est essentielle à l'amorce de la transcription, de la traduction et, afin de la production de protéines nécessaires aux changements morphologiques des synapses qu'impliquent la formation d'un souvenir.

Comme on peut le constater, les résultats de ces études menées par Colombo rencontrent trois des quatre conditions formulées par Bickle et Silva que j'ai présentées à la fin de la section 3.3. Selon Bickle, parce que la condition dont il ne

⁷⁹ La phosphorylation est l'addition d'un groupe phosphate à une protéine (ici les CREB) ou à une petite molécule comme, par exemple, le glucose ou l'adénine (Wikipédia 2012b). Le phosphate est un sel d'acide phosphorique.

traite pas est sur le point d'être rencontrée grâce à d'autres expériences [le critère *positive alteration*, je suppose] (Bickle 2008b, 470), ce score de 3 sur 4 conforte l'hypothèse selon laquelle la consolidation de la mémoire sociale, chez les souris, peut être expliquée essentiellement par l'activité des protéines CREB opérant dans les neurones de l'hippocampe ventral de ces dernières (ibidem, 474). C'est pour cela que Bickle pense que le *modèle cognitif* de la consolidation de la mémoire sociale chez les souris a été réduit (ou plutôt remplacé) par un *modèle cellulaire et moléculaire* (ibidem, 471 et 474). Soit dit en passant, pour finir, j'aimerais ajouter qu'on pourrait aussi appliquer ce modèle de la mémoire à long terme aux singes et aux humains. De fait, Bickle écrit :

[...] the biochemical and molecular-biological similarities between rodent CREB and primate CREB suggest that these same mechanisms produce social memory consolidation in primates, humans included. (ibidem, 471)

Dans la prochaine section, je montre en quoi consiste, selon Bickle, une explication causale idéale.

3.4 L'explication causale selon Bickle

En utilisant la méthodologie que j'ai présentée aux sections 3.3 et 3.3.1 et dans une perspective qu'il nomme « néocarnapienne » (Bickle 2003, 190), Bickle soutient que les scientifiques et les philosophes d'obédience physicaliste et réductionniste pourront montrer qu'il est possible de construire une théorie scientifique [biomoléculaire] de l'esprit [ou un « langage-objet », pour reprendre l'expression de Carnap,] qui ne fait référence à aucune entité psychologique et qui présente, selon des

moyens plus simples, la même portée explicative que les théories psychologiques (ibidem, 33)⁸⁰.

En conséquence, dans ce cadre théorique biomoléculaire, les *explications causales* concernant l'esprit que l'on retiendra seront strictement de nature cellulaire et biomoléculaire. Par exemple, soutient Bickle (2003, 112-113), on expliquera qu'hier, l'individu imaginaire Kurt a engrangé dans sa mémoire à long terme le numéro de téléphone que quelqu'un lui a dit verbalement par l'énoncé explicatif causal cellulaire et biomoléculaire (2) plutôt que par l'énoncé explicatif causal psychologique (1). Voici le premier énoncé.

⁸⁰ Cette conception de la réduction, pourrait-on ajouter et comme le remarque Alyssa Ney (2008, section 1c), présente un certain air de famille avec celle qu'ont développée John Kemeny et Paul Oppenheim (1956). En 2006, Bickle, suivant une remarque de Robert Richardson, semble lui-même l'admettre :

Robert Richardson noticed that of all the popular accounts of intertheoretic reduction from 20th century philosophy of science, my «intervene molecularly and track behaviorally» account seems closest to John Kemeny and Paul Oppenheim's (1956) account! In particular, Kemeny and Oppenheim stressed that the reducing theory need only explain the empirical data explained by the reduced and that reduction of intermediate theories was not necessary for reduction. Philosophers of science since Schaffner (1967) have dismissed Kemeny and Oppenheim's account as too weak, but perhaps in light of recent neuroscientific practice some of its basic ideas should be re-examined (except, of course, its logical empiricism-inspired account of the nature and role of theories in reduction). (Bickle 2006, 429, note 18)

Je prends le temps de mentionner cela parce que, avec le temps, Bickle semble avoir abandonné l'appareil conceptuel développé par les tenants de l'école structuraliste Wolfgang Balzer, C. Ulises Moulines et Joseph D. Sneed (voir Bickle 2006, 428-429) et avec lequel il avait formulé, dans des publications antérieures, une conception de la réduction (voir Bickle 2002; 2003, 95-102). Depuis 2006, en effet, Bickle semble conceptualiser l'idée de réduction autrement, à savoir essentiellement selon les techniques expérimentales des sciences cognitives cellulaires et moléculaires (voir section 3.3 et 3.3.1). C'est cette dernière conception de la réduction de Bickle que Richardson trouve proche de celle de Kemeny et Oppenheim (1956).

(1) *Énoncé explicatif causal psychologique*

- a. En termes psychologiques fonctionnels et causaux on pourrait dire que le fait que Kurt se *répéta* le numéro quinze fois et qu'il n'y eut pas d'*interférence rétrograde* dans les trente minutes qui suivirent *causa* que ce numéro de téléphone fut *emmagasiné* dans sa mémoire à long terme et lui est accessible aujourd'hui. (Bickle 2003, 113)

On peut représenter cet énoncé de manière plus schématique de la manière suivante :

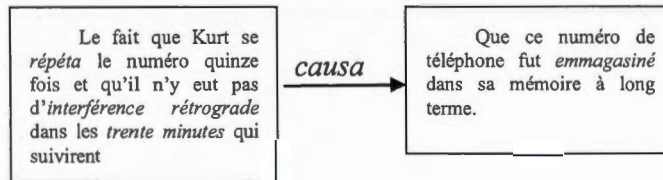


Figure 6 Représentation d'un énoncé explicatif causal psychologique.

Voyons maintenant l'autre énoncé dont le contenu est les événements cellulaires et biomoléculaires qu'implique la potentialisation à long terme tardive (PLT-T). (Je n'aborde pas la potentialisation à long terme précoce [PLT-P] parce que je mets l'accent sur ce qui advient, au niveau moléculaire, quand le numéro de téléphone *se consolide* dans le cerveau de Kurt).

(2) *Énoncé causal explicatif cellulaire et biomoléculaire*⁸¹

- a. Dans le cerveau de Kurt, à la suite du stimulus auditif qui fut provoqué par le son de la voix de cette personne lui disant un numéro de téléphone, l'activité des neurones frontaux conduisit à la libération de dopamine (DA) dans les neurones hippocampiques, lesquels neurones furent aussi activés, au même moment, par la voie anatomique reliée aux récepteurs auditifs.
- b. Cette *activation accrue* des neurones hippocampiques provoqua une entrée massive d'ions de calcium dans le neurone postsynaptique par la voie du récepteur NMDA (N-méthyl-D-aspartate) qui fut débloqué lors de la PLT-P. (On peut identifier le concept « activation accrue » à l'expression psychologique « *Kurt se répéta* »).
- c. Cette entrée massive des ions de calcium dans le neurone postsynaptique provoqua une augmentation de la quantité du complexe nommé « calcium-calmoduline » (Ca^{2+} -CaM)⁸². Ce complexe, conjointement à un récepteur second messenger, lui-même activé par les influx d'inter neurones modulaires, activa la protéine G, laquelle, à son tour, transforma les molécules d'adénosine triphosphate (ATP) en adénosine monophosphate cyclique (AMPc) (Bickel 2002, 130).
- d. L'AMPc se lia aux sous-unités régulatrices des protéines kinases A (PKA). Cette liaison libéra les sous-unités catalytiques de la PKA (ibidem, 130).
- e. Il résulta de cette libération une augmentation (dans les neurones hippocampiques postsynaptiques) de la quantité des sous-unités catalytiques PKA.
- f. Cette augmentation permit à certaines de ces sous-unités catalytiques de la PKA d'éventuellement se translocaliser vers les noyaux des neurones postsynaptiques affectés dans ce processus. À cette étape de

⁸¹ Je tiens à dire que, suivant Bickel (2003, 113), j'ai réécrit de manière plus schématique, dans ce que j'ai appelé « l'énoncé (2) », le passage où ce dernier explicite en quoi consisterait une explication biomoléculaire de la capacité de Kurt de mémoriser un numéro de téléphone qu'il se fait dire. L'énoncé (2) contient aussi quelques détails moléculaires relativement à la biologie de la mémoire que Bickel ne cite pas intentionnellement dans cet exemple mettant en jeu Kurt. On retrouve ces détails et beaucoup d'autres ailleurs dans ce livre de Bickel (2003 chapitre 2, section 4.2). Pour ma part, j'ai pris ces détails additionnels dans Bickel (2002, 130).

⁸² Le complexe nommé « calcium-calmoduline » naît de la liaison d'un ion de calcium avec la calmoduline (une protéine). Ainsi, plus il y a de calcium, plus il y a de calcium-calmoduline.

ce processus biomoléculaire, au moins deux choses importantes eurent lieu :

- i. Premièrement, dans le noyau de ces neurones, les sous unités catalytiques de la PKA phosphorylèrent les protéines CREB-1 (« cAMP-response element binding protein-1 ») lesquelles, ensuite, se lièrent aux sites CRE se situant sur les régions de contrôle de multiples gènes. Cette liaison des protéines CREB-1 aux sites CRE activa la transcription, la traduction et ultimement, la production de protéines. Les protéines produites furent, en premier lieu, des protéines régulatrices dont la fonction fut d'activer les sous-unités catalytiques des PKA pendant une journée. Cette activité prolongée des PKA permit, en deuxième lieu, l'activation de l'expression des gènes transcripteurs à réponse tardive, lesquels transcrivirent et produisirent les protéines requises pour modifier la structure de la membrane synaptique et, ainsi, améliorer durablement son efficacité (voir Bickel 2003, 113).
- ii. Deuxièmement, dans le noyau de ces mêmes neurones, en interagissant avec les MAP Kinase (« mitogen-activated protein kinase »), les sous-unités catalytiques des PKA inhibèrent les CREB-2 (« cAMP-response element binding protein-2 ») (Bickel 2002, 130). Il est nécessaire que cette inhibition des protéines CREB-2 ait lieu puisque l'on croit que les CREB-2 inhibent l'action des CREB-1 en se liant aux CREB-1 et en se fixant sur l'élément de réponse des CRE (ibidem, 130). Bref, l'interaction PKA-MAP kinase bloque les effets répressifs de CREB-2 (ibidem, 130).

Dans le diagramme suivant, on retrouve une *représentation* des étapes (d) à (f) de l'énoncé causal explicatif cellulaire et biomoléculaire précédent (2) :

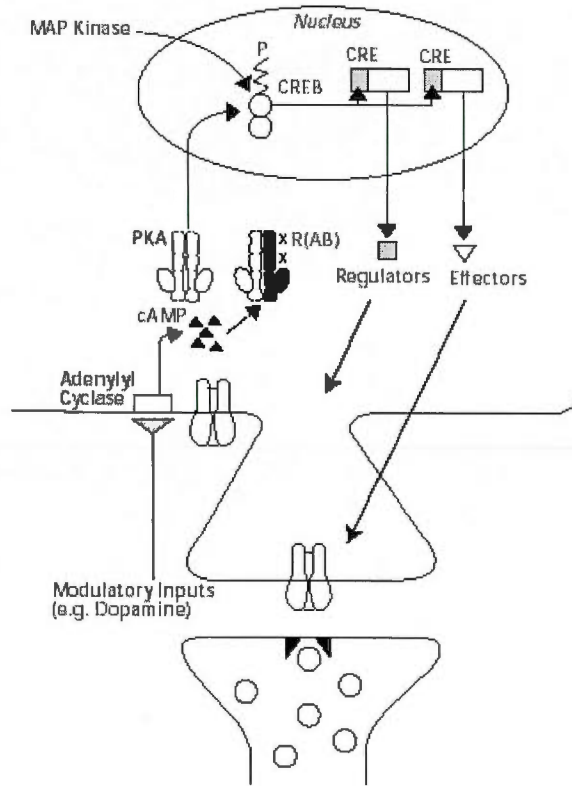


Figure 7 Représentation des étapes (d) à (f) de l'énoncé causal explicatif cellulaire et biomoléculaire précédent (2) (schéma tiré de Bickel 2003, 69).

Selon Bickel, l'énoncé (2) explique causalement beaucoup plus de choses que l'énoncé (1). En fait, en dernière instance, selon lui, l'énoncé (2) discrédite totalement l'énoncé (1) :

Juxtaposed next to each other, one need not possess a detailed philosophical theory of causal explanation to judge how *explanatorily impotent* — how *empty* — the psychological causal story has become. The cellular/molecular neurobiological account explains many key causal processes that the psychological account is either completely blind to or leaves as input-output black boxes. In other words, it *explains* events that the psychological account leaves *unexplained*. When we have neurobiological causal explanations actually in place, psychological causal explanations are rendered *otiose*. No wonder there

is no currently active “psychology of memory consolidation.” There is no longer anything to contribute to the available causal account by working at that level. This conclusion is not a matter of “intuitions” about what constitutes a better explanation. It is rather a report of scientific practice regarding this example over the past thirty years. (Bickle 2003, 112-113)

J’aimerais maintenant rappeler que cette façon de se représenter à l’aide de schémas comme ceux-ci l’explication de la consolidation mnésique s’inscrit dans la *conception sémantique* des théories et, plus spécifiquement, selon moi, dans la formulation qu’en donne Giere (voir section 3.2.1). En effet, nous avons ici à faire à des *modèles représentationnels* et, plus spécifiquement, à des *diagrammes* construits par des agents (ici, Bickle et moi-même) qui tentent, suivant certains objectifs, de représenter selon un certain degré de *similarité* un pan de la réalité (ici, le mécanisme de la consolidation mnésique chez les mammifères) (pour la conception des modèles de Giere voir section 3.2.1). De plus, le diagramme de Bickle représentant les processus cellulaires et biomoléculaires de la mémoire, en plus d’être un *modèle représentationnel*, est un *modèle mécaniste* parce qu’il représente les mécanismes à l’origine d’un comportement ou d’une fonction (ici, c’est la fonction cognitive de la consolidation mnésique) (pour la notion de *modèle mécaniste*, voir section 3.2.2). Enfin, selon moi, suivant ce que j’ai dit à la section 3.2.2, le diagramme représentant les processus cellulaires et biomoléculaires de la mémoire est, plus spécifiquement, un modèle mécaniste *réductionniste* parce que, conformément à la méthodologie réductionniste de Bickle, n’y sont représentés que les mécanismes opérant au niveau cellulaire et moléculaire. Idéalement, en effet, aucun facteur causal psychologique ne doit figurer dans une explication de l’esprit (ou d’une fonction de l’esprit), selon ce dernier.

Mais ce n’est pas tout. En effet, en plus de vouloir se débarrasser, ultimement, des explications psychologiques quand elles n’auront plus d’intérêt heuristique (Bickle 2006, 428), Bickle semble soutenir que les *variables environnementales* ont

beaucoup moins d'importance que les *variables biomoléculaires* dans nos explications des fonctions de l'esprit (voir Bickle 2008b, spécialement la section 5). En conséquence, me semble-t-il, cela signifie que, ultimement, suivant les thèses de Bickle, les explications des troubles mentaux ne devraient pas contenir de variables psychologiques et peu ou pas de variables environnementales.

À mon avis, cette exclusion des variables psychologiques et cette négligence des variables environnementales dans nos explications des maladies mentales devraient être justifiées par un modèle ou une métaphysique de la causalité⁸³. Mais Bickle n'en a pas développé. Plutôt que de construire un tel modèle, comme on vient de le voir, Bickle justifie l'évacuation des *variables causales psychologiques* pour expliquer l'esprit parce qu'il juge, suivant, entre autres, l'exemple de Kurt susmentionné, qu'elles sont vides d'un point de vue explicatif (« *explanatorily impotent* ») si on les compare à la force explicative des *variables causales cellulaires et biomoléculaires* (Bickle 2003, 113).

En ce qui concerne les *variables causales environnementales*, ce dernier semble soutenir qu'elles ont moins d'importance, dans nos explications de l'esprit, que les *variables causales biomoléculaires* parce que les résultats des recherches en neurobiologie moléculaire de la mémoire lui font penser que les *événements biomoléculaires* qui ont lieu dans le cerveau, en plus d'être nécessairement les causes directes de nos pensées et de nos comportements, présentent une efficience causale plus grande que celle que présentent les *événements de l'environnement* (Bickle 2008b, 472-473). Cette thèse de Bickle que j'appelle « thèse neurocausale » est basée sur ce que je nommerais deux constantes expérimentales qui, selon ce dernier,

⁸³ L'idée générale selon laquelle Bickle devrait justifier son épistémologie impitoyablement réductionniste par une théorie de la causalité apparaît dans la thèse de doctorat de Markus Eronen (2010, 30).

démontrent la supériorité de la force causale des événements neuraux sur celle des événements environnementaux. Je formulerais ces deux constantes comme suit : (1) chez un organisme capable de cognition, on peut causer un changement cognitif en intervenant au niveau biomoléculaire dans le système nerveux de ce dernier *sans* intervenir dans son environnement ou *sans* que son environnement change et (2) chez un organisme capable de cognition, *on ne peut pas* provoquer un changement cognitif en intervenant sur l'environnement *sans citer* un changement biomoléculaire au niveau neural. Autrement dit, lors de la production d'un comportement ou d'une pensée, un événement environnemental implique nécessairement un événement biomoléculaire dans le cerveau, mais un événement biomoléculaire n'implique pas nécessairement d'événement environnemental (lequel événement environnemental implique nécessairement un *changement* environnemental).

Dans la prochaine section, je présente plus en détail l'argumentaire sur lequel Bickle fait reposer sa *thèse neurocausale* selon laquelle les événements environnementaux présentent moins d'efficience causale que les événements biomoléculaires.

3.4.1 Les faits empiriques supportant la thèse neurocausale de Bickle

Comme je l'ai dit plus haut (sections 3.3, 3.3.1 et 3.4), chez les limaces de mer, les mouches à fruit, les souris et les humains, le *processus psychologique* qu'est la consolidation mnésique ou, et c'est la même chose d'un point de vue ontologique, le *processus électrochimique et neural* qu'est la potentialisation à long terme tardive (PLT-T) consiste en une chaîne d'événements biomoléculaires menant à des changements de l'expression génétique (à ce sujet, voir section 3.4, points [a] à [f]). Or, nous l'avons vu, ces changements causent la synthèse de protéines lesquelles

entraînent des modifications du cytosquelette du neurone. Entre autres, pour que cette synthèse des protéines et, par voie de conséquence, cette restructuration de la membrane de la synapse se fassent normalement, il faut que les protéines CREB-1 soient opérationnelles (voir section 3.4, points [f]). En effet, ce sont ces dernières qui activent l'expression des gènes responsables de la synthèse des protéines produisant la métamorphose de la membrane de la synapse, laquelle métamorphose permet aux neurones de se raffermir en réseaux, laquelle formation de réseaux est nécessaire au maintien de nos apprentissages à long terme. Or, Bickle pense qu'il est plausible de soutenir que des souris contraintes, lors d'études expérimentales, à vivre dans un environnement isolé (seules et sans stimuli environnementaux) présenteraient des protéines CREB-1⁸⁴ en très petit nombre et, par voie de conséquence, n'auraient pas de mémoire à long terme (Bickle 2008b, 473). En ce cas, remarque Bickle, *l'environnement causerait, par l'entremise d'une cause biomoléculaire directe* (la non-activité des protéines CREB-1 qu'impliquerait l'absence de ces dernières), *une déficience importante de la mémoire à long terme* (ibidem, 473). Mais, explique ce dernier, dans ce même environnement pauvre en stimuli et où les souris sont isolées, il serait possible, lors d'une expérience, d'intervenir directement dans le cerveau de celles-ci en leur injectant dans l'hippocampe ventral des gènes codant pour les protéines CREB-1 afin de réactiver l'activité de leur protéine CREB-1 et, corollairement, de rendre opérationnelle, à nouveau, leur mémoire à long terme (ibidem, 473). En ce cas, remarque-t-il, si nous réussissions cette expérience — laquelle est facilement faisable, selon lui —, un événement neural causerait un changement cognitif *sans que nous ayons à évoquer de causes ou de changements environnementaux*. En effet, l'environnement, observe-t-il, *resterait le même*, c'est-à-dire pauvre en stimuli (ibidem, 473). Or, selon Bickle, le fait que nous puissions provoquer un changement cognitif (ici, le retour de la mémoire à long terme) en

⁸⁴ Bickle ne spécifie pas, dans son exemple, s'il s'agit des protéines CREB-1 ou des protéines CREB-2. Je pense qu'il s'agit des protéines CREB-1. Pour cela, j'emploie, dans ce passage, l'expression « CREB-1 ».

manipulant un événement biomoléculaire dans le cerveau de ce petit mammifère *sans que n'adviennent de changements dans l'environnement* est une preuve empirique démontrant la supériorité de l'efficience causale des événements neuraux sur l'efficience causale des événements environnementaux.

En définitive, considérant ces conceptions épistémique (sections 3.3, 3.3.1, 3.4) et neurocausale (section 3.4.1), il est plausible d'affirmer que Bickle soutiendrait les thèses ontologique (a) et épistémique suivantes (b) :

- a) *Les troubles de santé mentale sont toujours et fondamentalement des maladies du cerveau.* Bien que Bickle affirme que les événements environnementaux font partie du réseau causal d'où émergent les comportements⁸⁵, je pense qu'il soutiendrait, en dernière instance, que les causes véritables des troubles de santé mentale sont des causes biomoléculaires. Je crois cela parce que Bickle accorde, comme nous venons de le voir, une efficience causale supérieure à ces dernières. Conséquemment, je suppose qu'il soutiendrait, comme Kandel, que

[...] les troubles du comportement qui caractérisent les maladies psychiatriques sont des troubles de la fonction cérébrale, même dans les cas où les causes de ces désordres sont clairement d'origine environnementale (Kandel, 2002 [1998], 19).

En d'autres termes, Bickle soutiendrait, à l'instar de Thomas R. Insel et Remi Quirion qu'au final « [...] mental disorders are brain disorders [...] » (Insel et Quirion 2005, 2221).

⁸⁵ De fait, Bickle affirme (avec raison, je pense) : « Events in an organism's environment are often components in the causal nexus that produces behavior » (Bickle 2008b, 473).

- b) *Une explication complète des troubles de santé mentale ne devrait pas citer de variables psychologiques, peu ou pas de variables environnementales, et beaucoup de variables neurobiologiques* (idéalement biomoléculaires, c'est-à-dire des variables faisant référence aux gènes et aux protéines, par exemple).

Avant de clore cette section, j'aimerais signaler qu'à l'heure actuelle, dans le discours social et psychiatrique, on rencontre implicitement et sous d'autres habits la thèse (a) selon laquelle les troubles de santé mentale sont *toujours* et *fondamentalement* des maladies du cerveau. À mon avis, cette conception des maladies mentales suggère implicitement la croyance selon laquelle la pharmacothérapie est la meilleure manière de soigner les troubles de la psyché. Prenons un exemple pour illustrer mon propos. En ce moment, comme le remarque dans le quotidien *Le Devoir* du 15 novembre 2012 le psychologue Jean-Philippe Vaillancourt, « [l]e gouvernement du Québec a récemment acheté du temps publicitaire pour diffuser un message qui soutient que la dépression est une vraie maladie, tout comme le cancer » (Vaillancourt 2012, section A8). À mon avis, l'analogie qui est ici faite entre la dépression majeure et le cancer porte à croire que, fondamentalement, la dépression majeure est *essentiellement* une maladie du cerveau et doit être au moins soignée avec des antidépresseurs. Bien sûr, opérer des changements environnementaux bénéfiques favorisant la guérison et/ou suivre une psychothérapie ne nuira pas à la personne dépressive. Mais, en dernière instance, de la même façon que c'est la chimiothérapie ou la radiothérapie qui guérit le cancer, ce sont les antidépresseurs qui guériront la dépression majeure. Comme le dit Gold, de fait, « [...] if mental illness is a form of brain pathology [(ici la dépression majeure)], then one ought to treat the illness by ministering to the brain » (Gold 2009, 507). Ce raisonnement analogique (à mon avis problématique) véhiculé dans les médias est conforté, entre autres, par le fait que le nombre d'ordonnances d'antidépresseurs est

en constante augmentation et très élevé au Canada comme au Québec. De fait, la prescription de ce type de psychotrope a augmenté de 75 % au Canada de 1999 à 2004 (Dubé 2007, 46). Aussi, en 2004, les médecins canadiens ont rédigé 24, 2 millions d'ordonnances de pilules contre la dépression (Dubé 2007, 46). À ceci s'ajoute, entre autres, le fait qu'

[a]u cours des dernières années, la prévalence de l'usage des antidépresseurs chez les adultes québécois assurés par le régime public d'assurance médicaments a augmenté de façon constante, passant de 13,3 % en 2005 à 14,4 % en 2009. (Gouvernement du Québec 2011, 7)

Est-il exact que la pharmacothérapie est toujours l'approche la plus indiquée pour traiter les troubles de santé mentale? De plus, est-il juste de penser que les maladies mentales sont toujours fondamentalement des maladies du cerveau et qu'elles doivent être expliquées presque essentiellement selon des modèles mécanistes cellulaires et biomoléculaires? Je pense que non et j'explique pourquoi dans la prochaine section.

3.5 Critique de l'épistémologie réductionniste et de la thèse empirique neurocausale de Bickle

Dans les sections précédentes, j'ai présenté le modèle de l'explication mécaniste et réductionniste de Bickle. Ce faisant, j'ai montré pourquoi, suivant son modèle, il fallait expliquer les troubles de santé mentale en termes de mécanismes biomoléculaires et avec peu ou pas de facteurs environnementaux. Dans ce qui suit, j'expose des arguments qui discréditent la thèse épistémologique réductionniste et la thèse neurocausale de Bickle, de même que l'idée véhiculée implicitement par le discours médiatique et selon laquelle les traitements pharmacologiques sont les plus indiqués pour soigner les maladies mentales. Ces arguments antiréductionnistes sont

ceux qu'ont formulés Kendler (2001, 2005, 436-437; 2008a), d'autres philosophes et scientifiques (voir, entre autres, Bunge 2003a; Kirmayer et Gold 2012; McNally 2011, chap. 5 et 6; Murphy 2006) et moi (la première partie de l'argument 2 et l'argument 8).

- (1) Suivant Kendler (2005, 436), contrairement à ce que dit Bickle, les explications des troubles de santé mentale formulées en termes psychologiques ou mentalistes non pas seulement un rôle heuristique passager. Au contraire, elles présentent un intérêt épistémique certain et durable. Pour cette raison, il n'est pas souhaitable de traduire — afin de les éliminer ou de les remplacer — les termes psychologiques et/ou phénoménologiques de la psychopathologie en des termes biomoléculaires. De fait, la psychiatrie et la science en général, il me semble, n'ont rien à gagner de ce projet épistémologique exclusivement réductionniste et, ultimement, éliminativiste que défend Bickle. Il me semble que c'est justement grâce à la diversité de nos concepts scientifiques que nous réussissons à appréhender les multiples facettes de la réalité et, plus spécifiquement, en ce qui nous concerne, les caractéristiques des troubles de santé mentale.

Considérant cela, il s'ensuit qu'il est nécessaire d'étudier intrinsèquement les croyances de l'humain *et* son système biologique afin de le comprendre et d'expliquer les pathologies mentales qu'il peut développer. Comme le dit Kendler, en effet,

to have a complete understanding of humans, two different kinds of explanations are required. Lots of different names have been applied to these two kinds of explanations. The first can be called mental, psychological, or first person. The second can be called material, biological, or third person. (Kendler 2001, 997)

Ainsi, les croyances négatives développées suite à un événement stressant et humiliant peuvent être considérées comme des causes d'une dépression majeure que l'on doit représenter dans nos explications de ce trouble⁸⁶ (voir Kendler 2005, 436). De la même façon, le système des croyances intentionnelles et inconscientes d'une personne présentant un trouble obsessionnel compulsif (TOC) doivent faire partie de nos explications psychiatriques. En effet, on ne comprendrait pas adéquatement ce système de croyances si on tentait de l'analyser de manière essentiellement biomoléculaire, et ce, même si le TOC résulte nécessairement et essentiellement (comme tout ce que nous sommes et faisons) de processus biologiques (au sujet des bases biologiques du TOC, voir Insel 2010, 47-49).

Il est à noter que cette perspective pluraliste implique, comme l'affirme Kendler, que *la biologie ne remplacera pas la psychologie dans notre système explicatif* (Kendler, 2008a, 700). Il s'agira plutôt de raffiner, grâce au progrès des neurosciences, nos construits psychologiques et d'approfondir, du même coup, dans ce processus de recherche, notre compréhension des processus biologiques [dont sont issus les troubles de santé mentale] (ibidem, 700). Ainsi, selon cette approche pluraliste, les perspectives biologique et psychologique coévolueront en s'enrichissant l'une l'autre (ibidem, 700). Comme le dit Kendler,

An iterative relationship between psychology and biology — where initial psychological constructs are better defined and subdivided by initial biological findings, which in turn help clarify the biology — will be needed to reach a more complete understanding. In short, biological and psychological perspectives will coevolve. (ibidem, 700)

⁸⁶ Il est fait mention de cela dans le diagramme de Kendler et coll. que l'on peut trouver à la section 2.6.4 de ce mémoire.

Il est intéressant de noter qu'une coévolution des théories existe aussi dans le domaine de la physique. Par exemple, Sklar (1999, 201-202) explique qu'en mécanique statistique de non-équilibre nous avons besoin d'utiliser la thermodynamique standard pour déterminer la distribution initiale des probabilités des micros états du système physique que l'on veut étudier. En ce sens, la mécanique statistique de non-équilibre (une théorie de bas niveau T_B) *ne peut pas évoluer sans la thermodynamique* (une théorie de haut niveau T_H). En conséquence, cela rend la faisabilité de la réduction de la thermodynamique par la mécanique statistique de non-équilibre peu plausible (et non souhaitable). Comme l'affirme Sklar, « [...] reducibility of thermodynamics to statistical mechanics [is] dubious in a fundamental way » (Sklar 1999, 202).

Donc, selon ce qui précède, à l'inverse de ce que dit Bickle, les explications psychologiques ou mentalistes en psychiatrie, que ce soit en pratique clinique ou en recherche, ont une valeur épistémique et sont pertinentes. Elles ne perdent pas leur intérêt heuristique ou ne deviennent pas « *explanatorily impotent* », pour reprendre l'expression de ce dernier (Bickle 2003, 113), une fois que l'on a déterminé l'histoire biomoléculaire que supposent un trouble de santé mentale et le réseau de croyances qui lui est associé⁸⁷.

Pour finir, j'aimerais ajouter que cette position *épistémologique pluraliste* reste pertinente même si on suppose que l'esprit consiste essentiellement en des processus neuraux et biomoléculaires⁸⁸. De fait, le réductionnisme ontologique n'implique pas le réductionnisme épistémologique (Bunge 2003a, 141). Dans le passage suivant,

⁸⁷ Je bonifie cet argument épistémique pluraliste à la section 3.8.1 quand je montre, suivant Woodward (2008a) et Kendler (2008c), qu'un énoncé explicatif psychologique présente parfois une *valeur explicative* plus grande qu'un énoncé explicatif biologique.

⁸⁸ À l'appendice B, j'aborde brièvement ce qui me semble être la position de Kendler relativement à ce que l'on appelle la relation corps/esprit.

Bunge explique bien cette idée en montrant pourquoi les *concepts* de la psychologie (et ceux de la psychopathologie, ajouterais-je) ne sont pas tous réductibles à ceux des neurosciences :

Assume, for the sake of the argument, that all mental processes are brain processes (ontological reduction). Does this entails that psychology is a branch of biology and, in particular, of neuroscience (epistemological reduction)? Not quite, and this for the following reasons. First, because brain processes are influenced by social stimuli, such as words and encounters with friends or foes. Now, such psychosocial processes are studied by social psychology, which employs sociological categories, such as those of social group and occupation, which are not reducible to neuroscience. A second reason is that psychology employs concepts of its own, such as those of emotion, consciousness, and personality, as well as peculiar techniques, such as interrogation and suggestion, that go beyond biology.

We conclude then that, even though the psychoneural-identity hypothesis is a clear case of ontological reduction, and sensationally fertile to boot, psychology is not reducible to neuroscience even though it has a large overlap with it. (Bunge 2003a, 141)

Dans le même ordre d'idée, je dirais que ce n'est pas parce que la musique se réduit au mouvement des molécules d'airs produites par les vibrations d'un corps (par exemple celles que produit une corde de guitare) qu'il faut expliquer la musique en employant essentiellement le langage de la physique. L'essence de la musique, en effet, réside dans l'effet émotif et/ou intellectuel qu'elle produit chez celui ou celle qui l'écoute et/ou la joue. Ainsi, évacuer le langage psychologique (et narratif et symbolique) des explications psychiatriques serait comme évacuer la musique de la civilisation humaine sous prétexte que le monde physique, intrinsèquement, ne contient pas de musique, mais seulement des atomes en mouvement lesquels, *si personne ne les écoute*, sont des événements physiques qui ne signifient rien.

Bref, de manière générale, on ne peut pas évacuer tout ce qui caractérise psychologiquement l'humain du discours psychiatrique. Pour cela, il nous faut garder les variables psychologiques dans nos explications ou modèles en médecine mentale.

(2) Comme je l'ai dit à la fin de la section 3.4.1, mon avis est que Bickle affirmerait qu'au final «[...] mental disorders are brain disorders [...]» comme le soutiennent Thomas R. Insel et Remi Quirion, (Insel et Quirion 2005, 2221). Je pense que Bickle soutiendrait cette thèse parce qu'il affirme que les événements biomoléculaires du cerveau présentent une efficience causale supérieure à celle que présentent les événements environnementaux. Pour cela, il me semble qu'il est plausible de penser que Bickle accorderait probablement peu d'importance aux *variables environnementales* dans les explications des troubles de santé mentale. Par exemple, il représenterait l'explication d'une dépression majeure comme l'explication mécaniste de cette affliction que j'ai présentée, suivant Insel (2010, 46), à la section 2.6.4 de ce mémoire. Son souhait, aussi, serait idéalement d'établir l'histoire biomoléculaire de cette pathologie.

Pour ma part, comme Bickle, je crois qu'il faut trouver les bases biomoléculaires des troubles de santé mentale. Mais je pense que la thèse neurocausale de Bickle est problématique. En effet, il me semble que ce dernier fait un raisonnement inductif inadéquat lorsqu'il infère (a) la *généralisation ontologique* selon laquelle les événements biomoléculaires présentent *toujours* plus d'efficience causale que les événements environnementaux du (b) *fait particulier* que les souris isolées retrouvent leur mémoire à long terme après avoir subi une manipulation génétique dans leur hippocampe ventral. Autrement dit, on ne peut pas tirer *d'un seul* fait expérimental

une généralisation ontologique relative à la puissance des causes⁸⁹. D'ailleurs, il existe des exemples qui réfutent la thèse de Bickle. Par exemple, si par magie, je me retrouve sur la planète Vénus, il est certain que mon cerveau et toute ma personne fondront. Or, dans ce cas de figure imaginaire, il me semble évident que les événements environnementaux (c'est-à-dire la température qu'on peut identifier à l'énergie cinétique moléculaire moyenne des molécules qui constituent l'atmosphère de Vénus) présentent une efficience causale *plus grande* que les événements biomoléculaires dans mon cerveau.

À mon avis, dans la sphère familiale, sociale, économique et culturelle, il existe des événements environnementaux qui approchent en intensité, si l'on peut dire, la température sur Vénus. Pour cela, ces événements ne sont pas bons pour l'intégrité physique des cerveaux et, en conséquence, peuvent potentiellement causer des troubles de santé mentale. Plusieurs faits empiriques confortent cette hypothèse. Par exemple, selon Richard Wilkinson et Kate Pickett, les personnes vivant au sein de sociétés inégalitaires sont beaucoup plus susceptibles de souffrir de maladies mentales que les personnes vivant dans les sociétés égalitaires (Wilkinson et Pickett 2009, chapitre 5). La figure suivante illustre cette constante⁹⁰.

⁸⁹ Je sais que Bickle dit qu'il ne fait pas d'ontologie (ou de métaphysique). Néanmoins, je trouve que ses considérations sur la puissance causale des événements ont quelque chose qui relève de l'ontologie considérant que l'ontologie est, selon Bunge (2003b, 201, mes *italiques*), « [t]he branch of philosophy that studies the most pervasive features of reality, such as real existence, change, time, *causation*, chance, life, mind, and society ».

⁹⁰ J'ai trouvé ce graphique que l'on retrouve dans le livre de Wilkinson et Pickett (2009, 67) sur le site Internet « The Equality Trust » à l'adresse suivante : <http://www.equalitytrust.org.uk/why/evidence/mental-health/francais>

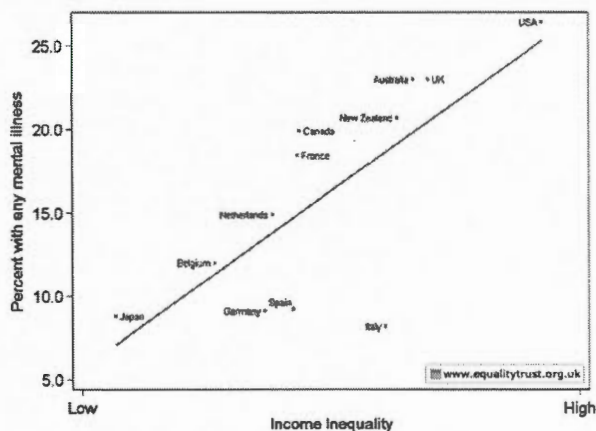


Figure 8 Représentation du fait selon lequel les personnes vivant au sein de sociétés inégalitaires sont beaucoup plus susceptibles de souffrir de maladies mentales que les personnes vivant dans les sociétés égalitaires (graphique tiré de Wilkinson et Pickett 2009, 67).

De ce graphique, on peut, il me semble, tirer la *généralisation invariante*⁹¹ ou l'énoncé explicatif causal suivant : « l'inégalité des revenus cause la maladie mentale ».

Il existe d'autres types de processus environnementaux qui causent le développement des troubles de santé mentale. Par exemple, Kendler affirme que plusieurs études témoignent du fait que les *processus* [(ou *mécanismes*)] *culturels* influent sur les troubles de santé mentale (Kendler 2005, 436, mes *italiques*). Prenons le cas de la boulimie. Richard J. McNally affirme à ce sujet que

[e]pidemiological and historical evidence indicates that bulimia nervosa is a culture-bound syndrome occurring only in contemporary Western societies or in societies strongly influenced by contemporary Western modern media. (McNally 2011, 134-135)

⁹¹ Le concept « généralisation invariante » a été développé par James Woodward (2003). J'aborde ce concept à la section 3.8.1 de ce mémoire.

En effet, plusieurs études, souligne Kendler, montrent que les cas de boulimie ont augmenté de manière significative chez les filles ces dernières années en Occident (Kendler 2005, 436). Aussi, parallèlement à cette réalité, dans les pays orientaux, il semble que la prévalence des cas de boulimie chez les adolescentes augmente dans la mesure où ces dernières sont exposées, par le biais de la télévision et d'autres vecteurs médiatiques, aux idéaux de la beauté féminine de la culture occidentale (ibidem, 436). Donc, ces études suggèrent que la boulimie est causée, au moins partiellement, par les idéaux de beauté corporels que l'on retrouve en Occident.

Cette hypothèse causale est d'autant plus plausible que la boulimie, semble-t-il, n'a pas toujours existé. Comme le dit McNally,

[a] cultural ideal of thinness as beauty is essential for the syndrome [of bulimia] to emerge. Moreover, the syndrome is unlikely to take root in settings where the patient does not have access to abundant food and the privacy to consume it rapidly and then vomit without others noticing. (McNally 2011, 135)

Pour cette raison, il est presque impossible qu'il y ait eu des cas de boulimie quand, il y a longtemps, l'idéal de beauté féminine était, dans plusieurs parties du monde où régnait une culture esclavagiste et où la nourriture était rare, d'être bien en chair, voire obèse (ibidem, 135-136).

Ainsi, comme on peut le constater suivant ces faits empiriques, s'intéresser aux représentations culturelles dans nos recherches en psychiatrie est pertinent parce que ces représentations peuvent avoir une incidence causale réelle sur les comportements et les pensées des gens et sur le développement des troubles de santé mentale. Cela dit, remarque Kendler, bien que la culture existe, en dernière instance, dans le cerveau de chacun des membres d'un groupe social par le biais de systèmes de croyances, il est improbable qu'étudier seulement les corrélats neurobiologiques de ces

représentations culturelles soit la bonne façon de procéder pour *comprendre* et *expliquer* ce qui cause les psychopathologies comme, dans ce cas, la boulimie (Kendler 2005, 436, mes *italiques*). Encore une fois (voir l'argument 1), force est de constater que le réductionnisme ontologique n'implique pas le réductionnisme épistémologique (à ce sujet, voir Bunge 2003a, 141 et Kendler 2001, 997-998). Par conséquent, les variables environnementales doivent figurer dans nos explications — c'est-à-dire dans nos modèles mécanistes — en psychiatrie.

- (3) Les arguments 1 et 2 montrent, comme le remarque Kendler, qu'en plus des facteurs neurobiologiques et génétiques, une explication complète d'au moins certains troubles de santé mentale requiert que l'on tienne compte de facteurs psychologiques ou mentalistes (par exemple, vivre une humiliation) et culturels (par exemple, les idéaux de la beauté féminine) (Kendler 2005, 436). On doit aussi tenir compte de facteurs économiques, comme je l'ai dit (l'inégalité des revenus). Mais il y a plus à considérer. En effet, remarque Kendler, jusqu'à il y a peu de temps, nous avons supposé que ces facteurs biologiques, sociaux et psychologiques ou mentalistes agissaient *de manière indépendante* sur le développement des troubles mentaux (ibidem, 436). Mais la réalité est plus complexe et ce n'est pas le cas. En effet, tous ces facteurs causaux s'auto-influencent (voir postulat 5 de l'annexe B) (ibidem, 437). Ainsi, un individu présentant une prédisposition génétique pour une maladie mentale quelconque verra son risque de la développer modifié à la fois par les conditions environnementales familiales (« rearing environment »), les événements stressants et le poids culturel (ibidem, 437). La boulimie, dont on vient de parler (voir argument 2), est un exemple qui illustre bien ce type d'étiologie où les causes

biologiques, psychologiques et sociales s'interpénètrent⁹². De fait, affirme McNally,

[...] [c]ultural factors are necessary, but not sufficient, for the emergence of bulimia nervosa. Not every young woman bombarded with media depictions of thin models and celebrities develops the disorder. People possessing certain psychological or biological risk factor may be especially vulnerable to fear of becoming fat that motivates the diet-binge purge cycle. (McNally 2011, 136)

Bref, comme l'indique Kendler (2005, 437), les facteurs biologiques favorisant le développement des maladies mentales sont modifiés par des facteurs agissant à un plus haut niveau d'abstraction épistémique (par exemple, les niveaux mental, familial, culturel, économique). Ainsi, en raison de l'impact causal qu'ont ces forces environnementales sur la biologie de l'individu, contrairement à ce que suggèrent les thèses réductionnistes de Bickle, on ne peut les négliger quand l'on représente par un modèle mécaniste une maladie mentale (voir postulat 6 de l'annexe B).

(4) Il faut aussi citer des variables environnementales dans nos explications des maladies mentales parce qu'il n'est pas toujours possible, en utilisant seulement une approche réductionniste et biologique en psychiatrie (comme Bickle), de définir adéquatement une dysfonction (Kendler 2005, 437). En effet, bien que certains symptômes psychiatriques soient considérés comme pathologiques essentiellement sur la base de processus biologiques (par exemple, les hallucinations), plusieurs autres symptômes ne sont jugés dysfonctionnels que dans certains contextes, c'est-à-dire dans certains *environnements* (ibidem, 437). Par exemple, au niveau physiologique, une attaque de panique s'exprime probablement biologiquement de la même manière chez un individu sain passant

⁹² C'est aussi ce que pense Murphy qui prend comme exemple, lui aussi, la boulimie pour faire valoir ce fait (Murphy, 2006, 262-263).

proche de tomber en contexte d'escalade et chez un individu souffrant d'agoraphobie dans un centre commercial bondé de monde (ibidem, 437). Cependant, dans ce cas, la personne qui souffre d'un problème de trouble de panique n'est pas l'escaladeur, mais la personne dans le centre commercial. Comme quoi, note Kendler (ibidem, 437), il faut tenir compte du contexte pour évaluer adéquatement la teneur pathologique d'un symptôme. Ainsi, la seule considération du niveau neurobiologique, dans une explication psychiatrique, est encore une fois insuffisante.

- (5) Pour réfuter Bickle, on pourrait ajouter à l'argument 4 l'idée selon laquelle, dans certains cas, il faut considérer l'interaction environnement/gène lorsqu'on veut déterminer si un gène est normal ou anormal. Si on tient compte de l'environnement, un génotype qui rend un individu vulnérable à la dépression, par exemple, n'est pas nécessairement plus anormal ou pathologique qu'un génotype résilient qui protège l'individu de ce trouble mental. En effet, comme l'affirment Kirmayer et Gold (2012, 324) et McNally (2011, 181), dans un *environnement stressant*, si une personne présentant un profil génétique vulnérable à la dépression risque davantage de développer ce trouble qu'une personne présentant un profil génétique résilient en regard de cette même affliction, il semblerait aussi que, dans un *environnement sain*, quelqu'un présentant un profil génétique vulnérable s'épanouira davantage qu'une personne présentant un profil génétique résilient. En d'autres termes, dans un environnement sain, présenter un génotype vulnérable à la dépression peut être avantageux. Les gènes de vulnérabilité, dans ce contexte, pourraient être conçus, dit McNally, comme des gènes de plasticité (ibidem, 181) et, par voie de conséquence, il me semble, comme des gènes d'adaptabilité. À ce sujet, ce dernier mentionne une analogie florale intéressante formulée à l'origine par les chercheurs Bruce Ellis et Thomas Boyce :

The developmental researchers Bruce Ellis et Thomas Boyce liken children with these genotypes to orchids that decline rapidly under adverse conditions but blossom into flowers of stunning beauty under conditions of nurture and support, whereas children with resilient genotypes are like dandelions that manage to survive no matter what the conditions. (ibidem, 181)

Ensuite, d'un point de vue évolutionniste, McNally formule cette remarque judicieuse au sujet de la présence, dans notre pool génétique, de ces gènes de vulnérabilité ou, selon un autre point de vue, de ces gènes de plasticité :

If genes predisposing a person to psychopathology in adverse environments foster plasticity and flourishing in favorable environments, then it is easy to see why they would remain in the gene pool throughout the course of human evolution. (ibidem, 181)

Bref, vu ces considérations sur (a) la nécessité de considérer l'environnement lorsqu'on détermine certaine dysfonction (argument 4) et sur (b) l'interaction environnement/gène dont il faut tenir compte quand on juge de la normalité ou de l'anormalité d'un gène (argument 5), expliquer un trouble de santé mentale, contrairement à ce que dirait (il me semble) Bickle, commande, au moins pour certains cas, que l'on fasse référence à l'environnement.

(6) Outre le fait qu'il néglige les variables environnementales, le modèle de Bickle ne tient pas compte du fait qu'un ensemble de questions importantes en psychiatrie sont de nature historique et, pour cela, peu sujettes à des explications réductives de nature essentiellement biologique, souligne Kendler (2005, 437). Par exemple, interroge ce dernier, pourquoi les humains sont-ils prédisposés à développer une dépression quand ils sont exposés à l'adversité sociale? Selon Kendler, ce type de question requiert davantage une explication historique et évolutionniste qu'une explication physiologique (ibidem, 437). En regard de cette question, par exemple,

il me semble que l'on trouve une piste de réflexion de type psychoévolutionniste chez Price et coll. (1997) qui ont formulé « l'hypothèse de la dépression due à la compétition sociale » selon laquelle la dépression serait, dans un contexte d'adversité sociale, une stratégie adaptative consistant en une subordination involontaire qui, ultimement, aiderait un individu à accepter sa condition et sa position dans une échelle sociale (Price et coll. 1997 cités dans Faucher 2005, 14). Cette hypothèse suivant laquelle la dépression est due à la compétition sociale est compatible avec l'hypothèse plus générale selon laquelle la dépression serait une fonction adaptative (certes douloureuse) retenue par la sélection naturelle permettant, entre autres, aux individus de mieux réfléchir aux problèmes de leur vie (Andrew et Thomson Jr., 2010). Il est intéressant de noter que, dans cette perspective, la dépression, dans la plupart des cas, ne serait pas pathologique. Plusieurs données scientifiques, notamment en ce qui concerne les gènes, les neurotransmetteurs et leurs récepteurs, la neurophysiologie, la neuroanatomie, la pharmacologie, la cognition, les comportements et l'efficacité des traitements confortent cette hypothèse évolutionniste (ibidem, 2010, 61). Cependant, notons que Paul W. Andrew et J. Anderson Thomson Jr. pensent que la dépression majeure existe. Toutefois, sa prévalence, disent-ils, n'est probablement pas de 30 % (comme l'indiquent certaines études), mais plutôt de 1 % ou de 2 %, comme pour la schizophrénie et les troubles obsessionnels compulsifs (ibidem, 61). Bref, comme on peut le voir, encore une fois, la perspective réductionniste, comme celle qu'adopte Bickle, ne suffit pas pour expliquer les troubles de santé mentale parce qu'elle évacue ces facettes explicatives historiques et évolutionnistes.

- (7) En psychiatrie, comme l'explique Kendler (2005, 437), il est très difficile (et peut-être impossible), comme semble vouloir le faire Bickle, de développer des

modèles explicatifs essentiellement réductionnistes⁹³ dans lesquels seraient représentées des correspondances directes entre des processus biologiques élémentaires et des symptômes de pathologies ou de dysfonctions mentales. En effet, établir des explications réductives comme celle que formule Bickle et selon laquelle l'absence de mémoire à long terme est causée par l'absence du gène codant pour les protéines CREB dans le cerveau est très rare en médecine mentale. Certes, il existe certains cas où ce type de lien presque direct entre une anomalie génétique et les symptômes d'une pathologie est déterminable. Murphy, par exemple, mentionne le cas de la maladie de Huntington (Murphy 2006, 133-135). De fait, affirme-t-il, « the relation between symptoms and cause, as in Huntington's, nearly always holds » (ibidem, 135). Mais ce type d'explication n'est pas la norme en médecine mentale. En effet, dit McNally,

Risk for all (or nearly all) psychiatric disorders is multifactorial. In addition to environmental influences on risk [(voir les arguments 3, 4 et 5)], multiple genes contribute to liability to disorder. Moreover, each gene has only small to modest effect on risk. Most genes have pleiotropic effects, that is, although each gene codes for specific protein, the subsequent function of these proteins is diverse. The upshot is that most genes have multiple downstream phenotypic consequences. (McNally 2011, 172)

Ces considérations empiriques sont vérifiées, par exemple, par le fait que le développement de la schizophrénie n'est pas causé seulement par un allèle anormal (ibidem, 172-173).

⁹³ Kendler (2005, 437) utilise l'expression anglaise « hard reductive models » pour désigner des modèles réductionnistes comme celui de Bickle. Je traduis cette expression anglaise par l'expression « modèles explicatifs essentiellement réductionnistes ».

Bref, pour ces raisons, un modèle explicatif essentiellement réductionniste, [(comme celui de Bickle)], affirme Kendler (2005, 437), ne convient pas à la psychiatrie (voir aussi les postulats 5, 6 et 7 de l'annexe B).

- (8) Pour finir, à la lumière de nos connaissances actuelles en psychiatrie, je me propose de commenter et de critiquer les thèses suivantes de Bickle : (a) nous savons beaucoup de choses relativement au cerveau; (b) nous savons manipuler les événements biomoléculaires ayant lieu dans le cerveau. De fait, Bickle écrit

[...] in current cellular and molecular neuroscience, *not only do we know a lot about how the components of nervous tissue work and interact; we also know how to manipulate these mechanisms, causally and directly, to generate novel behavioral data and significant dissociations in experimental tasks long accepted by behavioral scientists.* (Bickle 2006, 427, mes italiques)

J'aimerais aussi commenter l'idée provenant du discours médiatique et selon laquelle (c) nous avons intérêt à privilégier les traitements pharmacologiques pour soigner les troubles de santé mentale. Comme je l'ai dit, cette idée s'harmonise avec la conception de Bickle selon laquelle les troubles de santé mentale sont *toujours* fondamentalement des maladies du cerveau (voir la fin de la section 3.4.1; voir aussi l'argument 2 de la section 3.5).

Concernant (a), s'il est vrai que nous connaissons beaucoup plus le système nerveux aujourd'hui qu'hier, nous n'en connaissons pas assez sur ce dernier pour prétendre comprendre suffisamment les bases biologiques de la plupart des troubles de santé mentale. Concernant (b), s'il est vrai que les expériences très intéressantes et ingénieuses qui se font dans le domaine des neurosciences biomoléculaires démontrent que nous sommes capables d'intervenir et de manipuler des mécanismes biomoléculaires de toutes sortes, il semble toutefois qu'en psychiatrie notre capacité à

contrôler nos interventions thérapeutiques par voie médicamenteuse ait ses limites. De fait, concernant (a) et (b), Kendler affirme :

We have little more than a glimmer of insight into the etiology of major depression (MD), schizophrenia, or personality disorders. Although our treatments are better studied and probably more effective than those available two or three decades ago, we do not know much about how they work, and they are certainly far from curative. (Kendler 2008a, 1, mes italiques)

Par exemple, concernant (b), il est intéressant de noter qu'en faisant une méta-analyse d'études portant sur l'efficacité des antidépresseurs, Irvin Kirsch et son collègue Guy Sapirstein ont découvert que l'effet thérapeutique des placébos actifs⁹⁴ était à peu près équivalent à celui des antidépresseurs (Kirsch 2009, chapitre 1). Kirsch déduit de cela que l'ingrédient actif de l'antidépresseur ne peut pas être à la source de la diminution des symptômes de la dépression (Kirsch 2009). En fait, pense-t-il, la rémission des symptômes est plutôt due à un effet placebo. L'hypothèse causale de Kirsch qui explique que l'antidépresseur *et* le placebo actif provoquent la diminution des symptômes dépressifs va comme suit : dans une étude clinique quelconque, lorsqu'une personne prend un antidépresseur (sans le savoir) ou un placebo actif (sans le savoir), elle ressent éventuellement des effets secondaires désagréables dus à la prise de l'un ou l'autre de ces médicaments; à ce moment, se forme chez elle la représentation mentale selon laquelle le médicament (qu'elle croit être l'antidépresseur) fonctionne; cette représentation provoque à son tour une représentation mentale d'espoir laquelle cause éventuellement la baisse des symptômes de la dépression.

⁹⁴ Un placebo actif n'est pas une pilule de sucre (dans ce cas, ce placebo serait alors un placebo passif). Un placebo actif est un médicament qui n'est pas le médicament dont on examine l'efficacité. Par exemple, dans le cadre d'une expérience où l'on examine l'efficacité d'un antidépresseur, un placebo actif pourrait être une pilule pour traiter l'acidité gastrique.

L'hypothèse causale de Kirsch selon laquelle les antidépresseurs sont et agissent comme des placébos actifs est d'autant plus plausible qu'il y a de bonnes raisons de penser que la théorie monoaminergique de la dépression n'est pas adéquate (Kirsch 2009, chapitre 4)⁹⁵. De fait, les études portant sur la déplétion des monoamines (sérotonine, dopamine et noradrénaline) montrent que, *chez les personnes saines, cette baisse des monoamines ne provoque pas de baisse de l'humeur* (ibidem, 91, mes *italiques*). Par exemple, lorsqu'on applique la technique de déplétion en tryptophane chez des sujets ne présentant pas de facteurs de vulnérabilité à la dépression, on ne remarque pas chez eux de changements significatifs de l'humeur (Cowen 2008, 435)⁹⁶⁹⁷. Aussi, fait très intéressant (et curieux), la tianeptine, un antidépresseur qui est un *promoteur* sélectif de la recapture de la sérotonine, *est aussi efficace* que les *inhibiteurs* sélectifs de la recapture de la sérotonine (ISRS) dans le traitement de la dépression (Kirsch 2009, 97). Cela signifie paradoxalement qu'un médicament (la tianeptine) qui *abaisse* la quantité de sérotonine dans le cerveau est aussi efficace qu'un médicament qui *augmente* (par exemple, la fluoxétine) la quantité de cette substance⁹⁸.

⁹⁵ Il faut tout de même noter que la quantité de sérotonine serait plus basse chez les gens présentant une dépression majeure, mais cette baisse de la quantité de la sérotonine ne serait pas la cause de la dépression. Elle serait plutôt *causée* par la dépression, c'est-à-dire par d'autres dysfonctionnements cérébraux (Cowen 2008).

⁹⁶ La sérotonine est fabriquée à partir du tryptophane que le sang amène au cerveau. La sérotonine présente dans le cerveau est fabriquée dans le noyau raphé à partir du tryptophane. En effet, la sérotonine que l'on retrouve ailleurs que dans le cerveau ne peut pas aller au cerveau puisque la barrière céphalorachidienne l'en empêche. Or, quand on diminue la quantité de tryptophanes, on diminue, en conséquence, la quantité de sérotonine.

⁹⁷ Toutefois, Cowen affirme aussi que les personnes en rémission d'une dépression présentent des anomalies dans le système cérébral qui traite les émotions, lesquelles vulnérabilités sont exacerbées quand le taux de sérotonines est plus bas (Cowen 2008, 435). Cette hypothèse est confortée pas le fait que, chez les gens ayant déjà été déprimés, l'application de la technique de déplétion en tryptophane entraîne des biais négatifs dans le traitement de l'information émotionnelle (Cowen 2008, 435).

⁹⁸ Même si ce n'est pas un fait qui tend à valider l'hypothèse causale de Kirsch, il est intéressant de noter que les études de ce dernier ont eu une influence importante sur les pratiques en santé mentale en Angleterre. De fait, le *National Health Service* a modifié son approche pour traiter la dépression (CBC NEWS, 19 février 2012) de la manière suivante :

En outre, j'ajouterais à (b) qu'en plus du fait que les médicaments pour l'esprit sont peu efficaces ou inefficaces, ils provoquent à peu près tous des effets néfastes — permanents ou passagers — sur la santé de ceux qui les prennent. En effet, il est loin d'être certain que les médicaments dont on se sert pour traiter les troubles de santé mentale, à savoir, entre autres, les antidépresseurs (par exemple, le chlorhydrate de fluoxétine), les régulateurs de l'humeur (par exemple, les sels de lithium), les antipsychotiques (par exemple, la clozapine), les anxiolytiques (par exemple, le clonazépam) et les stimulants (par exemple, le méthylphénidate), à l'heure actuelle, soient sécuritaires. Par exemple, selon des études récentes, le méthylphénidate et les dérivés de l'amphétamine, médicaments prescrits abondamment par des médecins et des psychiatres à des individus diagnostiqués d'un trouble de déficit de l'attention/hyperactivité (TDAH) pourraient être potentiellement dangereux pour la santé (Higgins, 2009). De fait, ce type de stimulants, à la longue, selon des études faites sur des animaux, pourrait altérer et même léser certaines structures du cerveau à un tel point que l'utilisateur de ces substances, le temps passant, serait susceptible de développer une humeur dépressive, anxieuse et des déficits cognitifs (ibidem, 40). Ces faits empiriques sont particulièrement alarmants si l'on considère qu'un nombre important d'enfants prennent ces médicaments — sans l'avoir choisi — parce qu'ils sont diagnostiqués TDAH.

Donc, toujours concernant (b), contrairement à Bickle qui semble soutenir que nous sommes très habiles pour manipuler les mécanismes biomoléculaires ayant lieu dans le cerveau (Bickle 2006, 427), je pense que nous sommes forcés de constater

-
1. Les antidépresseurs ne sont donnés qu'aux gens sévèrement déprimés (en traitement d'urgence).
 2. Pour les gens faiblement à moyennement déprimés :
 - a. On a mis sur pied un programme de formation massive d'aidants naturels (spécialisés dans l'écoute active).
 - b. On a mis sur pied un traitement de la dépression par l'activité physique.

que, vu les résultats de la pharmacothérapie, nous les maîtrisons somme toute que très approximativement. Si notre connaissance du cerveau était aussi grande que le dit Bickle (voir point [a] de cet argument), en effet, on serait capable de développer des médicaments pour le cerveau non seulement véritablement efficaces, mais aussi n'ayant pas d'effets dangereux pour la santé.

Enfin, concernant (c), à la lumière de ce que j'ai dit sur (a) et (b), il me semble que, à l'heure actuelle, privilégier les traitements pharmaceutiques pour soigner l'esprit serait un mauvais choix. En effet, il faut plutôt tenter, si c'est possible (cela ne l'est pas toujours, en effet), de soigner les gens souffrant de troubles de santé mentale sans l'aide de médicaments.

En somme, dans cette section, nous avons vu (1) (a) que la psychologie n'avait pas seulement un intérêt heuristique et (b) qu'il était épistémiquement pertinent de tenir compte des facteurs causaux psychologiques ou mentalistes dans nos explications et interventions thérapeutiques; (2) que certains troubles de santé mentale étaient causés, au moins partiellement, par des causes environnementales; (3) que l'activation des événements biologiques favorisant le développement de plusieurs troubles de santé mentale est mise en branle par des événements environnementaux; (4) qu'il faut parfois mentionner des facteurs environnementaux lorsque, dans une explication psychiatrique, on fait mention d'une dysfonction; (5) qu'il faut considérer l'interaction environnement/gène lorsqu'on veut déterminer la normalité ou l'anormalité d'un gène; (6) qu'il est pertinent de tenir compte des explications historiques ou évolutionnistes en psychiatrie; (7) qu'il est très difficile (et peut-être impossible), en psychiatrie, de développer des modèles explicatifs essentiellement réductionnistes où seraient formulées des correspondances directes entre des processus biologiques élémentaires et des symptômes observables de pathologies; et (8) (a) que notre connaissance des bases biologiques des troubles de santé mentale

reste très partielle et (b) que notre capacité à manipuler les chaînes d'événements biomoléculaires ayant lieu dans le cerveau est limitée et que, en conséquence (c) privilégier les traitements pharmacologiques pour traiter les maladies mentales est un mauvais choix. Pour toutes ces raisons, je pense que l'on ne peut pas retenir le modèle de l'explication mécaniste et réductionniste de Bickle en psychiatrie. De fait, à la lumière de ce qui précède, il semble que nos explications des troubles de santé mentale doivent présenter, en plus des variables biologiques, des variables psychologiques, environnementales (et même historiques). Dans le même ordre d'idées, nos interventions thérapeutiques doivent elles aussi être diversifiées. Autrement dit, on ne doit pas privilégier l'approche pharmacologique qui doit être employée en dernier recours.

Le modèle de Kendler que je présente dans la prochaine section tient compte de ce pluralisme épistémologique. Il s'accorde aussi avec l'idée selon laquelle l'utilisation de traitements divers (pharmacologique, psychologique, sociale, etc.) est pertinente. C'est pour cela que je soutiens qu'il faut adopter ce modèle de l'explication en psychiatrie.

3.6 Le modèle mécaniste de l'explication pluraliste de Kendler

Les idées dont Kendler s'est servie pour construire son modèle mécaniste de l'explication pluraliste des troubles de santé mentale ont plusieurs origines.

Premièrement, celles-ci proviennent des travaux scientifiques que ce dernier et ses collègues ont menés, à partir de 1986 (Kendler et Prescott 2006, 6), sur l'étiologie de divers troubles de santé mentale. Le livre *Genes, Environment, and*

Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders (2006) rend compte de ces recherches. En effet, on trouve dans ce livre les résultats de l'immense étude intitulée *The Virginia Adult Twin Study of Psychiatric and Substance Use Disorders* (VATSPUD) dans laquelle Kendler était un des chercheurs. Plus précisément, on découvre dans cet ouvrage comment les facteurs génétiques et environnementaux interagissent pour produire plusieurs troubles de santé mentale. Bref, comme l'explique Deborah L. Levy, ce livre

[...] takes the reader on a behind-the scenes excursion into the rationale, methodology, and findings of one of the most ambitious and influential genetic-epidemiological studies in recent history. The Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPUD) is a systematic attempt to explore the roles of genetic and environmental risk factors and their interaction in the etiology of a range of common disorders, including generalized anxiety disorder, major depression, and phobias (grouped as internalizing disorders) and childhood conduct disorder, adult antisocial personality disorder, and substance and illicit drug use (grouped as externalizing disorders). (Levy 2007, 1763)

Deuxièmement, les idées de Kendler proviennent de la littérature en philosophie des sciences. Ce dernier s'inspire, entre autres, des travaux du philosophe Kenneth F. Schaffner. En effet, Kendler (2008a) reprend de ce dernier l'idée selon laquelle les maladies mentales doivent être représentées par « des *modèles causaux interniveaux* qui se chevauchent » et la développe (Schaffner 1994, 281, mes *italiques*)⁹⁹. En effet, s'inspirant des travaux des tenants du néo-mécanisme, il raffine cette idée de

⁹⁹ Ces modèles mettent en jeu ce que Schaffner nomme des « prototypes » ou des « exemplaires » (Schaffner 1994, 283). D'une part, ceux-ci peuvent être des représentations idéalisées d'entités physiques comme, par exemple, des organismes, des cellules et des molécules qui évoluent selon des mécanismes ou des processus physiques s'inscrivant dans le temps. D'autre part, un exemplaire peut être une entité matérielle qui représente une autre entité matérielle. En effet, comme l'explique Schaffner, dans les sciences cliniques comme en psychiatrie, le patient, quand il est malade, est *l'exemplaire clinique* d'une maladie laquelle est un processus physique anormale et pathologique (ibidem, 283). Dans cette perspective, « une théorie de la pathologie — et cela inclut aussi la psychopathologie — peut être interprétée comme une famille de modèles, chacun de ces derniers présentant "quelque chose d'anormale" en regard du processus normal ou sain [(que l'on observe habituellement)] » (ibidem, 283, ma traduction).

Schaffner en la transformant en l'idée selon laquelle les maladies mentales doivent être représentées par des *modèles mécanistes* (voir section 3.2.2). C'est dans ce cadre conceptuel néo-mécaniste que Kendler développe la thèse selon laquelle les maladies mentales résultent de *mécanismes causaux* évoluant sur plusieurs niveaux de la réalité (par exemple, les niveaux biologique, psychologique, social et culturel) (Kendler 2008a). Kendler définit la notion « niveau de la réalité » selon le cadre conceptuel de la méréologie (ibidem, 696). Dans cette perspective, un niveau est identifiable à un mécanisme qui présente une relation partie/tout ou — et c'est la même chose — une relation partie/système (ibidem, 696) (voir la section 3.3.2 à ce sujet). Ainsi, la notion de mécanisme nous permet de représenter le fait que les molécules forment une membrane, que les neurones forment un circuit neural, que les systèmes neuraux forment (d'une manière mystérieuse) ce qu'on appelle un individu et que les individus forment les sociétés (ibidem, 696).

En 2011, dans l'article « What Kinds of Things are Psychiatric Disorders? » qu'il écrit avec le psychologue Peter Zachar et le philosophe Carl Craver, Kendler raffine encore sa conception des troubles de santé mentale. En s'inspirant des travaux du philosophe Richard Boyd sur la notion d'espèce naturelle (Boyd, 1991, 1999), ces trois chercheurs proposent qu'une maladie mentale soit conçue non pas comme une espèce naturelle au sens classique (par exemple, l'or), mais comme une « *mechanistic property cluster (MPC) kinds* » (Kendler, Zachar et Craver 2011, 1143). Une MPC consiste en (1) des réseaux de mécanismes causaux (2) interagissant entre eux (3) se renforçant mutuellement et (4) produisant les symptômes du trouble de santé mentale (ibidem, 1143). Notons aussi que cette conception des troubles de santé mentale absorbe le fait empirique selon lequel les mêmes symptômes d'une maladie mentale peuvent être produits par des mécanismes causaux différents. Autrement dit, en termes philosophiques, les troubles de l'esprit sont multiréalisables. Comme le disent Kendler et ses collègues,

[...] the MPC view allows that the same cluster of symptoms might arise from different etiological, underlying or sustaining mechanisms in different cases. MPC kinds are, in philosophical terms, « multiply realizable » by the mechanisms or sets of mechanisms that produce them. (ibidem, 1148)

Kendler et John Campbell (2009) proposent que l'on conçoive ces mécanismes causaux advenant sur plusieurs niveaux de la réalité dans le cadre du modèle interventionniste (MI) de la causalité développé par Woodward (2003, 2008a)¹⁰⁰. Comme ils le disent,

[...] l'MI peut nous doter d'un cadre conceptuel dans le domaine de la psychiatrie qui pourrait s'appliquer également aux multiples processus causaux — qu'ils aient lieu à l'intérieur ou à l'extérieur de l'esprit/cerveau — à l'œuvre dans les maladies mentales. Le MI nous permet de produire des modèles de l'explication des maladies mentales où figurent des variables causales de types différents. (Kendler et Campbell 2009, 881, ma traduction)

Cette conception des maladies mentales selon laquelle ces dernières consistent en de multiples mécanismes causaux qui s'auto-influencent et se renforcent mutuellement s'inscrit naturellement dans le pluralisme intégratif développé par la philosophe des sciences Sandra D. Mitchell (2003, 2009) et dont une des thèses épistémologiques principales est que différentes théories scientifiques compatibles et complémentaires peuvent nous donner une explication d'un même trouble de santé mentale¹⁰¹. En effet, si on suppose que les symptômes des maladies mentales résultent de mécanismes causaux multiples que l'on représente selon plusieurs niveaux épistémiques, alors on aura besoin d'utiliser plusieurs théories et les méthodes d'investigation que ces dernières supposent pour expliquer chacun de ces

¹⁰⁰ Je présente le MI aux sections 3.8 et 3.8.1 de ce mémoire.

¹⁰¹ Voir Mitchell (2009, 109-110) pour un exemple d'explication pluraliste de la dépression majeure.

niveaux (biologique, psychologique, sociale, etc.). Kendler affirme explicitement son attachement au pluralisme explicatif dans le passage suivant :

I advocate *explanatory pluralism* [...] as the approach best suited to understanding the nature of psychiatric illness. Explanatory pluralism hypothesizes multiple mutually informative perspectives with which to approach natural phenomena. Typically, these perspectives differ in their levels of abstraction, use divergent scientific tools, and provide different and complementary kinds of understanding. Explanatory pluralism is especially appropriate for psychiatry because psychiatric disorders are typically influenced by causal processes operating at several levels of abstraction. (Kendler 2005, 436)

Comme on peut le constater, l'épistémologie pluraliste et intégrative qu'adopte Kendler est très différente de l'approche épistémologique réductionniste qu'adopte Bickle (voir sections 3.3 et 3.4). De fait, selon ce dernier, en dernière instance, seulement les sciences biomoléculaires (et éventuellement microphysiques [Bickle 2006, 431-432]) devraient servir de cadre conceptuel pour donner des explications de l'esprit et, en conséquence, des troubles mentaux.

J'aimerais brièvement souligner que concevoir les maladies mentales en termes de mécanismes causaux multiples qui s'auto-influencent implique que nos interventions thérapeutiques doivent (ou peuvent) viser ces différents mécanismes biologiques, psychologiques, sociaux, économiques et culturels qui pourraient être à l'origine des troubles de santé mentale. Dans cette perspective, entre autres, les thérapies pharmacologiques ou les thérapies par la parole (humanistes) ou les thérapies cognitives et comportementales ou les politiques de santé publique ou un mélange de ces interventions peuvent être indiquées, selon les besoins et les cas, pour soigner les troubles de santé mentale. Cette perspective thérapeutique diffère de celle que suggère le discours social actuel et selon laquelle, idéalement, il est préférable de

soigner les afflictions de l'esprit de manière pharmacologique (voir la fin de la section 3.4.1 et l'argument 8 de la section 3.5).

Enfin, il est intéressant de noter que la conception mécaniste des troubles de santé mentale soutenue par Kendler est très différente de celle que l'on retrouve dans le *Manuel diagnostique et statistique des troubles mentaux. Texte révisé* (DSM-IV-TR) (2003). En effet, contrairement à celle proposée dans ce traité psychiatrique qui ne suppose aucune étiologie mais seulement une symptomatologie des maladies mentales (ou, en termes philosophiques, une phénoménologie antiréaliste), la conception des maladies mentales que formule Kendler repose sur la notion de *cause*. Ainsi, cette conception défendue par Kendler incite à la recherche des *étiologies* des maladies mentales. Selon ce paradigme épistémique que Kendler nomme, suivant Schaffner (1993, 197-199), le « conditionalized realism », nos théories ayant pour objet les troubles de santé mentale devront être révisées en fonction de ce que nous dit *la réalité ou, plus précisément, la structure causale de l'univers*. De fait, Kendler dit :

A conditionalized realism about psychiatric disorders (Schaffner, 1993) is so useful because it forces us to revise our theories when the world tells us that they are wrong. This iterative act of revision proceeds in the most orderly fashion if we take ourselves to be investigating an *objective causal structure* that is independent of our investigative and instrumental practices (Kendler, 2009). (Kendler, Zachar et Craver 2011, 1149, mes *italiques*)

Le « conditionalized réalism » développé par Schaffner (1993), selon Peter Zachar (2012, 192), est un type de position épistémique pragmatique se situant quelque part entre l'instrumentalisme et le réalisme scientifique. Voyons pourquoi.

Les tenants de l'instrumentalisme soutiennent que les théories sont des *instruments* nous permettant d'organiser notre expérience du monde, de faire des prédictions et d'intervenir sur les choses que l'on perçoit. Cependant, selon eux, il n'est pas nécessaire que nos théories nous donnent une représentation conforme à la réalité. Le psychologue B.F. Skinner, comme l'indique Zachar, était un grand représentant de cette tradition (ibidem, 192). Selon moi, les tenants de l'empirisme logique font aussi partie de cette école de pensée (voir les points 1 à 6 de la section 1.1 de ce mémoire).

Contrairement aux défenseurs de l'instrumentalisme, les tenants du réalisme scientifique affirment que le but de l'activité scientifique consiste à construire des théories nous donnant une représentation qui correspond à la réalité. Généralement, par exemple, on considère que Mario Bunge (2006) et Karl Popper (1991 [1979]), chez les philosophes contemporains, font partie des défenseurs de ce mouvement.

Quant à eux, les avocats comme Schaffner ou Kendler de cette forme de pragmatisme qu'est le « conditionalized realism » vont accepter les thèses instrumentalistes, mais ils ne voudront pas abandonner le concept de « réalité » qu'ils jugent très utile (Zachar 2012, 193). En effet, même s'ils concèdent aux défenseurs de l'instrumentalisme que le succès d'une théorie n'est pas garant de sa vérité (plusieurs théories qui ont eu du succès dans le passé se sont avérées fausses, en effet), les tenants du « conditionalized realism » ne peuvent pas s'empêcher de penser que nos théories qui nous permettent d'interagir avec le monde et d'intervenir sur lui *ne se frottent pas minimalement* à ce que l'on pense être la réalité. Comme le dit Zachar,

[t]he pragmatist notion of intervention and interaction highlights the role of both successes and failure. The combinations of being able to do things that were previously unimaginable (like refrigerating food) and also trying and failing (like not discovering the gene of schizophrenia) fits the notion *that we are bumping up*

against the world in some way. As Schaffner says, utility constitute what we think of reality. (Ibidem, 193, mes italiques)

Prenant acte du fait que le succès des théories scientifiques ne peut pas être retenu pour établir que ces dernières représentent la réalité, Schaffner propose, suivant le philosophe Arthur Fine (1986), qu'il existe des normes épistémiques sur lesquelles on peut s'accorder et qui nous permettent de juger si telle théorie ou tel modèle ou tel traitement est valide (Zachar 2012, 193). Par exemple, on peut s'entendre sur le fait que pour vérifier la validité de l'énoncé « ce parapluie est vert » on doit regarder le parapluie à la lumière et demander à quelqu'un dont la vision est normale quelle couleur il perçoit (ibidem, 193). Mais ces normes semblent locales et historiquement situées plutôt que générales et universelles, note Zachar (ibidem, 193). Ainsi, dans cette perspective, *notre connaissance de ce que nous croyons être la réalité ou, plus précisément, la structure fondamentale de l'univers*¹⁰² est déterminée ou *conditionnée* par les meilleures normes épistémiques d'une époque. C'est pour cela que Schaffner, me semble-t-il, nomme sa posture épistémique pragmatique le *conditionalized realism*.

Dans la prochaine section, afin de mieux faire comprendre l'approche néo-mécaniste qu'adopte Kendler, je montre comment celui-ci développe une explication mécaniste et pluraliste du trouble de la dépendance à l'alcool (Kendler 2008a, 696-697). J'en profiterai également pour montrer brièvement en quoi ce trouble est un « *mechanistic property cluster (MPC) kinds* ».

¹⁰² Selon ma compréhension des trois courants épistémiques susmentionnés, je pense que le tenant du réalisme scientifique utiliserait plutôt l'expression « notre connaissance de *ce qui est* la réalité ». Le défenseur de l'instrumentalisme, quant à lui, dirait « notre connaissance *des phénomènes que nous percevons* (la réalité nous échappe absolument; en effet, aucun langage ne peut nous informer sur elle) ».

3.7 La construction d'une explication mécaniste et l'exemple de la dépendance à l'alcool

Selon le DSM-IV-TR,

[1]a caractéristique essentielle de la dépendance à une substance est un ensemble de symptômes cognitifs, comportementaux et physiologiques, indiquant que le sujet continue à utiliser la substance malgré des problèmes significatifs liés à la substance. Il existe un mode d'utilisation répétée qui conduit, en général, à la tolérance, au sevrage et à un comportement de prise compulsive. Un diagnostic de Dépendance à une substance peut s'appliquer à toutes les classes de substances à l'exception de la caféine. (Association américaine de psychiatrie, 2003 [2000], 222)

Dans l'article « Explanatory Models for Psychiatric Illness » (2008a), Kendler propose de construire une explication mécaniste simplifiée du trouble « Dépendance à l'alcool » (ibidem, 696-697). Pour ce faire, explique-t-il, nous devons d'abord faire un croquis du mécanisme (« a mechanistic sketch ») qui représente, à notre avis, cette maladie (ibidem, 696). Faire cette hypothèse mécaniste ou ce croquis consiste à *localiser* (suivant des règles épistémiques précises et dans ce que nous croyions être la réalité) les parties du mécanisme qui, selon nous, interagissent entre elles et sont à l'origine du trouble de la dépendance à l'alcool. De l'avis de Kendler, ce trouble comporte quatre composantes causales (« causal pieces »), à savoir les parties (1) génétique/biologique; (2) psychologique; (3) social et (4) culturel/économique (ibidem, 697). Après avoir localisé ces parties, nous devons tenter de les *décomposer* afin de mieux les comprendre (ibidem, 696). Aussi, si c'est nécessaire, nous devons décomposer les parties de ces parties (et les parties de ces parties et ainsi de suite). En d'autres termes, et pour faire bref, il s'agira de décomposer (1), (2), (3) et (4) le plus possible. Pour le trouble de la dépendance à l'alcool, suivant Kendler (ibidem, 697), on aura, entre autres,

- (1) pour la variable « biologique/génétique », les sous-variables (a) « exposition prénatal à l'alcool » (b) « gènes responsables de la production des structures qui métabolisent l'alcool » (c) « gènes responsables de la fabrication des structures cérébrales qui interagissent directement ou indirectement avec l'éthanol » (par exemple, je suppose, les récepteurs GABA_A);
- (2) pour la variable « psychologique », les sous-variables (a) « névrosisme », « impulsif » et « extraverti » et (b) « tendances psychologiques à rechercher la consommation d'alcool » (« dimensions of alcohol expectancies »);
- (3) pour la variable « sociale », les sous-variables (a) « usage de substances par les pairs » (« peer substance use »), (b) « disponibilité de l'alcool » (« drug availability »), (c) « classe sociale » et (d) « état matrimonial »¹⁰³;
- (4) pour la variable « culturelle/économique », les sous-variables (a) « type d'alcool habituellement consommé » (b) « niveau d'acceptabilité sociale de l'ivresse » (« acceptability of public drunkenness ») (c) « perte des croyances culturelles traditionnelles » (d) « croyances religieuses » (e) « niveau de taxation des breuvages alcoolisés » (f) « grosseur des bouteilles contenant l'alcool » (« sizes of alcoholic beverage containers »).

En opérant cette décomposition (et éventuellement, voir plus bas, une *intégration*), on s'apercevra que, certaines fois, un mécanisme est facilement décomposable et que d'autres fois il l'est difficilement (*ibidem*, 696). Un mécanisme est dit « facilement décomposable » quand les parties qui forment le système sont organisées de manière *agrégative* (*ibidem*, 696). Les composantes de ces systèmes interagissent alors entre elles sans se modifier les unes les autres (*ibidem*, 696). Les *fonctions* de ces composantes dans ce type de système peuvent être comprises sans tenir compte des relations que ces parties entretiennent avec ce système (*ibidem*, 696).

¹⁰³ Dans le paragraphe qu'il consacre aux variables sociales, Kendler ne fait pas mention de la variable « état matrimonial ». Cependant, deux paragraphes plus loin, quand Kendler traite des facteurs causaux de la dépendance à l'alcool, il mentionne la variable « état matrimonial ». C'est la raison pour laquelle je me permets de l'introduire dans cette liste de variables sociales.

Ainsi, ces fonctions peuvent être comprises même dans un système différent (ibidem, 696). On dit de ces parties, dans ce cas, qu'elles sont interchangeables (« intersubstituable ») (ibidem, 696). Ce type de système ne présente pas de *causalité récursive* inter niveau (c'est-à-dire de la récursivité causale entre les parties de niveaux différents) ou intra niveau (c'est-à-dire de la récursivité causale entre les parties d'un même niveau) (ibidem, 696). Par ailleurs, un mécanisme peut être difficilement décomposable. Ces systèmes mécanistes sont composés de mécanismes non agrégatifs. Les composantes de ces systèmes interagissent entre elles et s'auto-influencent. Ce type de système, contrairement aux systèmes facilement décomposables, présente de la causalité récursive intra niveau. Il présente aussi de la causalité récursive inter niveau. À ce sujet, Kendler précise que

[...] [t]he most important kind of interlevel causal loop is where the output of the entire system feeds back to its own constituent components. This is called recursive, or top-down, causality. (Ibidem, 696)

Le trouble de la dépendance à l'alcool que nous venons de décomposer sommairement (suivant Kendler 2008a, 697), comme la plupart des troubles de santé mentale, est un système mécaniste difficilement décomposable. On peut affirmer cela parce que, lorsqu'on procède à l'*intégration* des parties (1), (2), (3) et (4), c'est-à-dire lorsqu'on essaie de déterminer comment celles-ci interagissent, on constate que le trouble de dépendance à l'alcool est composé de parties non agrégatives. En effet, d'une part, ces parties ne sont pas organisées de manière agrégative parce qu'on observe, entre autres, que le degré d'impact des facteurs génétiques sur le développement d'un trouble de dépendance à l'alcool *varie* en fonction des croyances religieuses, de l'état matrimonial et de l'environnement social où évolue l'individu (ibidem, 697). De fait, affirme Kendler,

[...] the impact of genetic risk factors for alcohol dependence fails the additivity and "intersubstitution" assumptions. Their effects are dependent on both biochemical and psychosocial contexts. (Ibidem, 697)

D'autre part, ajoute Kendler, les parties du trouble de dépendance à l'alcool ne sont pas organisées de manière agrégative parce qu'elles interagissent selon des relations *causales récurrentes* (« causal loops ») (ibidem, 697). Kendler formule quatre exemples illustrant cette récursivité (ibidem, 697).

- (1) Si un individu présente des gènes qui le rendent tolérant à l'alcool (facteurs génétiques), alors plus il est probable que celui-ci ressente une sensation de plaisir lorsqu'il boit de cette substance (facteurs psychologiques); plus il est bien et heureux quand il prend de l'alcool (facteurs psychologiques), plus celui-ci aura tendance à fréquenter des lieux où il y a de l'alcool (facteurs environnementaux); plus cet individu fréquente de tels lieux, plus il risque de consommer de l'alcool, et plus il en consomme, plus il risque de fréquenter les endroits où l'on retrouve de l'alcool et ainsi de suite. Bref, comme le résume Kendler,

[...] genes influence subjective ethanol effects, which influence alcohol expectations, which in turn loop out into the environment, influencing consumption patterns, which in turn affect risk of alcohol dependence. (ibidem, 697)

- (2) Si quelqu'un consomme souvent de l'alcool, alors son corps la métabolise plus rapidement et son système nerveux central devient moins sensible à celle-ci (facteur biologique) (ibidem, 697). Sa perception de satiété est donc altérée (facteur psychologique). Vu cela, cette personne aura la sensation, comme on le dit dans le langage populaire, de bien « supporter l'alcool ». Or, plus une

personne a cette sensation, plus il y a de chance qu'elle boive des boissons alcoolisées.

- (3) Les adolescents présentant un caractère impulsif et aimant le risque tendent à se rassembler et à se motiver mutuellement à adopter des comportements antisociaux, à prendre des drogues et à boire de l'alcool (ibidem, 697). Kendler affirme que des facteurs génétiques influencent ce processus [que l'on peut qualifier, je suppose, de psychosocial] (ibidem, 697). Comme il l'écrit,

[...] genetically influenced temperament causes individuals to select themselves into high-risk environments, which feed back on their risk for alcohol dependence by providing easy access to ethanol and encouragement for its excessive use. (ibidem, 697)

- (4) Enfin, les enfants dont les parents (ou un d'entre eux) sont de gros buveurs d'alcool ont tendance à développer une aversion pour l'alcool. En effet, bien que ces enfants présentent un génome qui les dispose à boire de l'éthanol (facteurs génétiques), ils s'abstiennent intentionnellement d'en absorber puisqu'ils connaissent les ravages sur la santé qu'occasionne une trop grande consommation de cette substance (facteurs environnementaux/familiaux).

Conformément à la perspective épistémique du « conditionalized realism » (Schaffner 1993) vue à la section 3.6, et afin de valider ce modèle du trouble de dépendance à l'alcool, nous devons le soumettre à des tests empiriques en respectant des règles épistémiques rigoureuses. En effet, en science, et c'est ce que prône Kendler, on fait une hypothèse causale plausible, on vérifie sa pertinence selon des critères épistémiques précis, puis on la corrige en fonction de ce que nous croyons être la structure causale de l'univers. Ici l'hypothèse prend la forme d'un mécanisme

(pas d'un système de lois comme dans l'empirisme logique). Cette mise à l'épreuve de la réalité nous aidera, éventuellement, à mieux articuler notre modèle pluraliste et mécaniste du trouble de dépendance à l'alcool (voir Kendler 2008a, 696).

Vu ce qui précède, on peut affirmer que le trouble de la dépendance à l'alcool est un « *mechanistic property cluster (MPC) kinds* » (Kendler, Zachar et Craver 2011, 1143). En effet, nous avons effectivement à faire à (1) des réseaux de mécanismes causaux opérant sur les niveaux génétique/biologique, psychologique, social et culturel/économique (2) interagissant les uns avec les autres (3) se renforçant mutuellement et (4) produisant les symptômes du trouble de la dépendance à l'alcool. Ce trouble de santé mentale, aussi, peut être multi réalisé. En effet, des combinaisons différentes de mécanismes causaux peuvent produire les mêmes symptômes. Par exemple, suivant ce que nous avons vu plus haut, alors qu'un individu peut développer un trouble de dépendance à l'alcool parce qu'il porte les gènes qui le prédisposent à bien métaboliser l'alcool, qu'il a accès facilement à cette substance et qu'il fréquente quotidiennement des endroits où on trouve de l'alcool, un autre peut développer ce trouble parce qu'il présente des gènes qui le prédisposent à développer un tempérament impulsif et, corollairement, à fréquenter des gens semblables qui boivent de l'alcool et fréquentent des endroits où on en trouve.

Éventuellement, ce trouble de santé mentale pourrait être représenté selon un *modèle mécaniste et pluraliste* comme l'est la dépression majeure à la section 2.6.4 (suivant Kendler 1993 et coll. cités dans Murphy 2006, 136). De plus, comme je l'ai dit précédemment (voir section 3.2.1), il me semble que ce type de modèle est un *modèle représentationnel* au sens de Giere. Pour cela, à mon avis, la démarche de Kendler s'inscrit, ultimement, dans la conception sémantique des théories scientifiques (voir section 3.2).

Dans les deux prochaines sections, je présente le modèle interventionniste (MI) de la causalité de Woodward. Ce faisant, je montre qu'il est pertinent en psychiatrie et qu'il faut l'inclure dans le modèle mécaniste et pluraliste de Kendler (comme le propose lui-même ce dernier).

3.8 Le modèle interventionniste de la causalité et la psychiatrie

À la section 3.6, suivant Kendler et Campbell (2009), j'ai noté que le modèle interventionniste (MI) de la causalité développé par Woodward est adéquat pour le domaine de la médecine mentale. Pour cela, dans ce qui suit, je présente sommairement le MI. D'abord, je montre qu'il est un cadre théorique qui nous permet de donner une signification¹⁰⁴ aux énoncés du type «*X cause Y*»¹⁰⁵. Ensuite, j'explique brièvement pourquoi le MI n'est pas un modèle anthropocentriste de la causalité. Puis, je montre que le MI nous permet de qualifier les relations causales selon plusieurs sortes de généralisations (invariantes, non invariantes, invariantes avec un grand champ d'application, invariantes avec un petit champ d'application). Après, je développe, suivant les thèses de Woodward et les idées de Kendler (2008c; Kendler et Campbell, 2009), l'idée selon laquelle, contrairement à ce que dit Bickle (sections 3.3 et 3.4), les explications causales citant des variables psychologiques présentent parfois un *contenu explicatif* plus pertinent — et, corollairement, plus utile pour faire des interventions (par exemple, thérapeutiques) — que celui que présente

¹⁰⁴ Comme le souligne justement Alexander Reutlinger (2012, 787-788), l'entreprise de Woodward est une entreprise sémantique. De fait, Woodward affirme : « [...] my project is semantic [...] » (2003, 38). Woodward veut, en effet, mettre en lumière ce que *signifient* les diverses locutions causales. C'est ce que dit Woodward explicitement : « [m]y aim is to give an account of the *content* or *meaning* of various locutions, such as *X causes Y* [...] » (2003, 38, mes *italiques*).

¹⁰⁵ Le MI sert à donner une signification à d'autres types d'énoncés causaux. Mais je ne les aborde pas explicitement dans mon mémoire.

les explications causales contenant des variables biologiques. Ce faisant, j'enrichis l'argument 1 de la section 3.5. Je clos cette section en montrant que le MI est intéressant parce qu'il ne rencontre pas le problème de la pertinence que rencontre le modèle des lois de couverture.

Avant de présenter le MI, toutefois, j'aimerais souligner qu'il existe d'autres modèles de la causalité. Par exemple, Bunge (2001 [1982]; 2006, 91) conçoit la causalité en termes de transfert d'énergie. Salmon (1984, 1994), quant à lui, la conçoit « [...] en termes de processus causaux, caractérisé par des processus physiques capables de transmettre des marques » (Bonnay 2011, 45). Par ailleurs, David Lewis (1986) définit la causalité en termes contrefactuels (Bonnay 2011, 45). Vu ces nombreuses conceptions de la causalité, pourquoi présenter seulement le MI?

J'ai décidé de présenter le modèle de Woodward parce que mon exposition du modèle de Kendler serait incomplète si je ne l'exposais pas. Outre ce souci d'exhaustivité, opter pour le MI m'apparaît pertinent puisqu'il s'harmonise au modèle mécaniste que formule Glennan (2002, S344-S345; 2011, 801-808) et que j'adopte (voir section 3.2.2). Enfin, ce choix me semble adéquat parce que le MI s'applique bien, de manière générale, aux explications causales que l'on retrouve en sciences de la santé dont fait partie la psychiatrie. Comme le dit Christophe Malaterre (2011, 6, note 8, ma traduction), en effet, le MI « [...] semble bien convenir à [la conception de] la causalité [que l'on rencontre] en sciences de la santé [...] ».

Cela dit, je sais que le MI n'est pas parfait et qu'il fait l'objet de critiques sévères (voir, par exemple, Reutlinger 2012). Une des façons de mettre en défaut le modèle de Kendler serait, d'ailleurs, de montrer les failles du MI. Ce n'est pas mon objectif, cependant. En effet, dans cette section, je souhaite prioritairement montrer la richesse conceptuelle du modèle de Woodward.

3.8.1 Présentation du modèle interventionniste

Dans le cadre du MI, un énoncé du type « X cause Y » signifie qu'une intervention I sur X (la variable indépendante) change Y (la variable dépendante). Plus précisément, suivant ce modèle, on dira qu'un énoncé du type « X cause Y » signifie « ... [qu']une intervention I qui change la valeur de X , change aussi de manière régulière et stable [les valeurs de] Y [...], et ce, dans un cadre de circonstances d'arrière-plan quelconque C » (Woodward 2008a, 139, ma traduction). Selon cette logique épistémique et causale, on peut affirmer, par exemple, que la position de l'interrupteur d'une lampe (X) est une cause du fait que la lampe soit allumée (Y) parce que l'on peut changer l'état de la lampe (fermée/ouverte) en manipulant (I) la position de l'interrupteur (fermé/ouvert) (voir Maher 2010). Aussi, dans cette perspective, on peut soutenir que la lecture sur le baromètre (X) n'est pas une cause de la tombée de la pluie (Y) parce que nous ne pouvons pas changer le fait qu'il pleuve ou ne pleuve pas en manipulant (I) ce qui est écrit sur le baromètre (ibidem, 2010). En outre, comme l'explique Craver (2007, 96), suivant Woodward (2003, 98), pour que la relation causale stipulée entre X et Y soit valable, l'intervention I sur X en regard de Y doit respecter quatre critères : (1) I ne change pas directement Y ; (2) ce n'est qu'en passant par X que I change les valeurs des variables causales S_1, S_2, \dots, S_n se situant entre X et Y ¹⁰⁶; (3) I n'est pas corrélée avec une autre variable C qui est une cause de Y et (4) I est la seule cause de X ou (si on formule différemment) aucune autre cause U , à l'exception de I , est la cause de X . Voici un schéma illustrant les quatre critères susmentionnés et que doit respecter une intervention idéale I :

¹⁰⁶ Notons qu'il n'y a pas toujours de variables S_1, S_2, \dots, S_n entre X et Y . En effet, X peut causer directement Y .

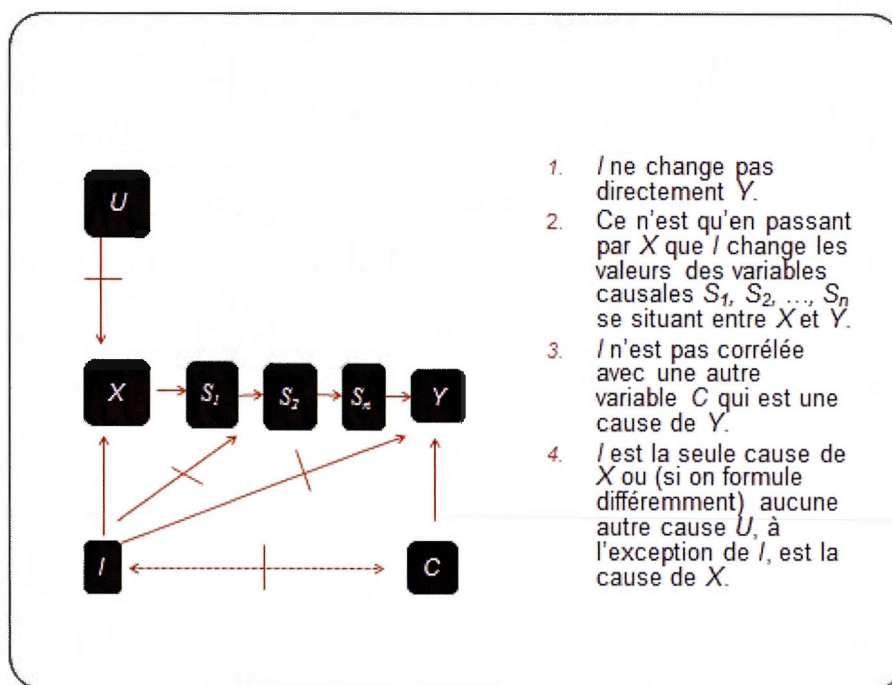


Figure 9 Les quatre critères que doit respecter l'intervention idéale (schéma adapté de Craver 2007, 97).

Les interventions les plus fréquentes dans le contexte de la recherche en biologie et en psychiatrie, explique Woodward (2008a, 139), sont les expériences randomisées, l'invalidation génique et les ablations chirurgicales précises d'aires neurales chez des modèles animaux (ou toutes autres formes d'interventions provoquant une incapacité comme la stimulation magnétique transcrânienne).

Contrairement à ce que ces exemples d'interventions pourraient laisser croire, une intervention, au sens de Woodward, à strictement parler, *ne nécessite pas une activité ou une intentionnalité humaine*¹⁰⁷. En effet, dans la mesure où un processus naturel respecte les quatre conditions susmentionnées, il peut être qualifié

¹⁰⁷ Notons que le modèle mécaniste et réductionniste de Bickle (section 3.3), quant à lui, suppose une action humaine. Suivant cela, entre autres, le MI se distingue du modèle de Bickle.

d'intervention (Woodward 2003, 94). En ce sens, le MI n'est pas anthropocentriste. Comme le dit Woodward,

[...] [t]here is nothing in the interventionist version of a manipulability theory that commits us to the view that all causal claims are in some way dependent for their truth on the existence of human beings or involve a "projection" on to the world of our experience of agency. (Woodward 2008b, section 9)

Plusieurs processus causaux ne nécessitent pas une intervention humaine, en effet. Par exemple, selon le cadre théorique de la relativité générale d'Einstein, la masse du soleil (I) cause une courbure de l'espace-temps (X), laquelle détermine la trajectoire (Y) que suit la Terre autour du soleil. De plus, lorsque cet astre solaire deviendra plus gros et gagnera en masse dans quelques milliards d'années avant de « mourir » (c'est-ce qu'on prévoit), la courbure de l'espace-temps changera et, en conséquence, la trajectoire de la Terre aussi. Ultimement, il y a de bonnes raisons de penser que la Terre finira par être engloutie par le soleil et aucun humain (fort probablement) ne sera là pour constater cet événement.

En formulant des relations causales de ce genre entre une variable X et une variable Y , nous tentons d'identifier ce que Woodward nomme des *généralisations* (2008a, 147-155). Comme nous l'avons vu, Glennan (section 3.2.2), Kendler et Campbell (section 3.6) adoptent explicitement cette manière de concevoir les interactions entre les parties d'un mécanisme. Ces généralisations, explique Woodward (2008a, 147-155), présentent un degré quelconque d'*invariance* ou de *stabilité*. Le degré de stabilité ou d'invariance d'une généralisation est déterminé de la manière suivante : (a) si une relation causale entre X et Y continue de tenir malgré une vaste gamme d'interventions I_1, I_2, \dots, I_n sur X et si cette relation persiste dans plusieurs contextes C_1, C_2, \dots, C_n , alors on dira que le degré d'invariance de cette relation causale est élevé. Inversement, (b) moins une relation causale entre X et Y

continue de tenir suite à un large éventail d'interventions I_1, I_2, \dots, I_n sur X et moins cette relation se maintient dans plusieurs contextes C_1, C_2, \dots, C_n , moins son degré d'invariance ou de stabilité sera jugé élevé. Illustrons maintenant (a) et (b).

- (a) À mon avis, l'énoncé « une diminution de l'activité électrique générale du cerveau (X) cause une sensation de calme (Y) chez la personne présentant ce cerveau » est une *généralisation invariante* (c'est-à-dire une relation causale dont le degré de stabilité est élevé). En effet, une vaste gamme d'interventions I_1, I_2, \dots, I_n visant à baisser l'activité électrique générale du cerveau, dans toutes sortes de contextes C_1, C_2, \dots, C_n , provoqueront un effet de calme chez une personne présentant ce cerveau. Par exemple, si une personne consomme de l'éthanol (c'est-à-dire de l'alcool) ou un anxiolytique, l'activité électrique de son cerveau diminuera et s'en suivra une sensation de calme, et ce, peu importe dans quel contexte cette personne se trouve (cela pourrait être un contexte de guerre, de tension sociopolitique, lors d'une allocution publique, lors d'une entrevue d'embauche, lors d'une fête, etc.). En effet, au niveau biomoléculaire, dans le cerveau de cette personne, lors de l'ingestion (I) de l'une ou l'autre de ces substances, se produira, entre autres, la chaîne causale suivante : les molécules de l'anxiolytique — par exemple celles que l'on appelle « benzodiazépines » — ou les molécules d'éthanol se fixeront, en plus des neurotransmetteurs inhibiteurs GABA, sur différents sites des récepteurs $GABA_A$, un récepteur sensible aux ions de chlore chargés négativement (Cl^-). La combinaison de l'action des benzodiazépines ou de l'éthanol et des neurotransmetteurs GABA sur les récepteurs $GABA_A$ aura pour effet de faire entrer davantage qu'à l'habitude de Cl^- dans les neurones postsynaptiques. Cette augmentation de Cl^- causera une *hyperpolarisation des membranes postsynaptiques* d'une très grande quantité de neurones et rendra ces dernières moins excitables. Puisqu'il y a beaucoup de récepteurs à GABA

dans le cerveau, il s'ensuivra une *diminution de l'activité électrique générale* (X) de ce dernier. Pour cela, au *niveau phénoménologique*, la personne prenant de l'alcool ou un anxiolytique *se sentira plus calme* (Y).¹⁰⁸

- (b) Un exemple de *généralisation instable* serait celui-ci : supposons, dit Woodward (2008a, 146), reprenant un exemple du biologiste Richard Dawkins (1982), un gène Z avec les caractéristiques suivantes : quand quelqu'un présente une forme défectueuse de ce gène, il est incapable d'apprendre à lire et quand il présente la forme normale de ce gène (et que plusieurs conditions génétiques et environnementales sont satisfaites, comme l'exposition à une culture où on retrouve de l'écriture, l'accessibilité à une éducation adéquate, et ainsi de suite) cette personne est capable de lire. Selon Woodward (*ibidem*), cette relation est minimalement causale plutôt que seulement corrélacionnelle. En effet, dit ce dernier, si on fait l'hypothèse que toutes les autres conditions environnementales et génétiques nécessaires à l'apprentissage de la lecture sont présentes, alors une intervention (I) sur ce gène (qui l'activerait ou le désactiverait) causerait qu'une personne développerait ou ne développerait pas la capacité de lire. Cependant, explique Woodward, l'énoncé explicatif causal « le gène Z cause le développement de la capacité de lire » est une *généralisation* plutôt *instable* parce qu'elle ne vaut que dans des conditions très précises. Si on change ces conditions, en effet, cette généralisation ne tient plus. Pour cela, cette généralisation ne peut nous être utile lors d'interventions expérimentales ou thérapeutiques que dans des circonstances très précises (*ibidem*, 149).

¹⁰⁸ À mon avis, ce cas exemplifie adéquatement l'idée selon laquelle les énoncés ou les généralisations empiriques du type « une diminution de l'activité électrique générale du cerveau (X) cause une sensation de calme (Y) chez la personne présentant ce cerveau » réfèrent à une trame précise d'événements ayant lieu dans la réalité matérielle (pour une idée semblable, mais beaucoup plus élaborée, voir Bunge 2006, surtout 295-302).

Notons, par ailleurs, qu'il faut distinguer la notion « stabilité » ou « invariance » et la notion « étendue du domaine d'application de la généralisation » (ibidem, 151). L'étendue d'application d'une généralisation est fonction de sa capacité à *s'appliquer à plusieurs sortes de systèmes* (ibidem, 151, mes *italiques*). Ainsi, une généralisation peut être très stable sans avoir un domaine d'application élevé. Par exemple, la maladie de Huntington est causée par une anomalie génétique bien précise que présente un gène du chromosome 4 (ibidem, 150). Cette anomalie génétique est fatale en toute circonstance. En effet, peu importe les conditions environnementales où évoluerait un individu porteur de cette anomalie génétique et peu importe l'environnement biomoléculaire du gène HD, si cet individu est porteur de ce gène qui présente une répétition de trois acides nucléiques (C, A et G — appelé triplet CAG) supérieure à 40, cet individu développera assurément cette maladie (Wikipédia 2012c). Nous voyons ici que la *généralisation* « posséder un gène HD présentant plus de 40 répétitions de CAG cause le développement de la maladie de Huntington » est *stable* ou *invariante*. Cependant, cette généralisation a un *domaine d'application très limitée* parce qu'elle ne concerne que le rôle d'une anomalie dans un gène spécifique produisant une forme particulière de maladie (Woodward 2008a, 151). Au contraire, la loi newtonienne de la gravitation universelle présente une *très grande étendue* (et une *bonne invariance*). Potentiellement, en effet, elle s'applique, dans l'univers, à tous les systèmes présentant une masse (ibidem, 151).

Mais à quoi nous sert de construire des généralisations invariantes? Eronen (2010, 33) explique que ce type de généralisations nous permet de formuler des questions de type contrefactuelle comme celle-ci : « que ce serait-il passé si les choses avaient été différentes? ». L'intérêt de se poser ce type de questions à la lumière d'une généralisation invariante est que cela nous permet de générer des explications

causales pertinentes et valides et, ultimement, d'opérer des manipulations efficaces sur les choses de la réalité. Par exemple, dit ce dernier,

[...] the ideal gas law can be used to show what the pressure of a gas would have been if the temperature had been different. In this way, the ideal gas law is potentially exploitable for manipulating and controlling the temperature, pressure and volume of a gas. (ibidem, 33)

Il va s'en dire que plus une généralisation est invariante et présente un grand domaine d'application, plus elle présente un potentiel explicatif et un potentiel de manipulation élevé. À l'inverse, toujours selon Eronen (ibidem), une généralisation vraie, mais non invariante ne peut pas nous permettre de générer des explications causales parce qu'elle ne nous permet pas de poser des questions de type contrefactuelle. En conséquence, ce type de généralisation ne peut pas être utilisé à des fins de manipulations expérimentales et de contrôle des phénomènes. L'énoncé suivant est une généralisation vraie, mais non invariante : « toutes les grandes assiettes sur la table de mon frère contenaient du spaghetti le vendredi 7 décembre 2012 ».

Je veux maintenant faire voir que, suivant Woodward (2008a, 159-168), le MI nous permet de montrer que, dans certains cas, un énoncé explicatif mettant en jeu des variables plus spécifiques sera *moins pertinent* qu'un énoncé mettant en jeu des variables moins spécifiques. Ce faisant, je tente de démontrer la fausseté de la thèse de Bickle selon laquelle seulement les explications présentant des variables causales biomoléculaires (et éventuellement microphysiques) ont un intérêt épistémique (voir section 3.4). Aussi, à l'occasion, je bonifie l'argument 1 de la section 3.5 et je montre l'intérêt et la richesse conceptuelle du MI.

Pour illustrer l'idée selon laquelle une explication ne doit pas nécessairement citer les variables les plus spécifiques (« fine grained ») pour être pertinente, Woodward reprend un exemple du philosophe Stephen Yablo (1992). Celui-ci décrit une expérience où un pigeon a été entraîné à picorer une cible seulement quand cette cible est de n'importe quelle teinte de rouge. Supposons maintenant que nous présentions au pigeon une cible pourpre et que ce dernier la picore. Étant donné la situation susmentionnée, se demande Woodward, lequel des deux énoncés suivants serait le plus explicatif et pertinent?

- (3.1) La présentation d'une cible pourpre (X) a causé, chez le pigeon, l'action de picorer la cible (Y).
- (3.2) La présentation d'une cible rouge (X) a causé, chez le pigeon, l'action de picorer la cible (Y). (Woodward 2008a, 161, ma traduction)

L'explication (3.1), selon Woodward (ibidem, 161), est moins appropriée. Selon lui, si on compare (3.1) et (3.2), en tenant compte de l'entraînement susmentionné du pigeon, l'explication (3.1) est *inutilement spécifique* (« inappropriately specific ») (mes *italiques*) (ibidem, 161). De plus, explique Woodward (ibidem), même si nous admettons que l'explication (3.1) est vraisemblable, il semble inadéquat de considérer cette dernière comme étant en compétition avec (3.2) : la vérité de (3.1) (nous avons fait l'hypothèse que [3.1] est vrai), en effet, n'exclut pas la vérité de (3.2).

Mais si ces explications sont toutes deux vraies, qu'est-ce qui pousse Woodward à préférer l'explication (3.2)? Ce qui lui fait préférer l'énoncé (3.2) est la *valeur explicative* qu'il présente. De fait, selon lui, l'explication (3.2) (considérant l'entraînement préalable du pigeon) présente une valeur explicative causale supérieure à l'explication (3.1). Ce qui nous intéresse, en effet, quand nous recherchons la cause qui fit que le pigeon picora est quelque chose, si l'on peut dire, qui rend compte du fait qu'il picora *plutôt* qu'il ne picora pas (ibidem). C'est en

examinant ce *contraste* que nous voyons émerger des informations causales pertinentes à notre explication. De fait, si nous reformulons ces deux énoncés en termes contrastifs, il y a des informations pertinentes que nous donne l'énoncé (3.2) que ne nous donne pas l'énoncé (3.1). Voyons d'abord l'énoncé (3.2) dans une formulation contrastive :

- (3.2) Le contraste entre le fait que la cible soit rouge ou ne soit pas rouge cause le contraste entre le pigeon picorant ou ne picorant pas (ibidem, 161, ma traduction).

Comme on peut le voir, l'explication (3.2) nous dit que si la cible avait été d'une autre teinte de rouge, le pigeon l'aurait picorée quand même. L'énoncé (3.2) nous dit aussi que si la cible n'avait pas été rouge, le pigeon ne l'aurait pas picorée (ibidem). Formulons maintenant l'explication (3.1) en termes contrastifs :

- (3.1) Le contraste entre le fait que la cible soit pourpre ou ne soit pas pourpre cause le contraste entre le pigeon picorant ou ne picorant pas (ibidem, 161, ma traduction).

Si on interprète l'énoncé (3.1) comme supposant que le pigeon ne picorera pas la cible si la cible n'est pas pourpre mais néanmoins rouge, alors l'énoncé (3.1) est faux, explique Woodward (ibidem). En effet, le pigeon picorerait la cible même si cette dernière était dans une teinte de rouge non pourpre. Même si nous trouvons que cette interprétation n'est pas charitable, concède ce dernier, il reste vrai que l'explication (3.1) nous dit moins de choses sur ce que nous voudrions savoir sur les *conditions complètes* [(c.-à-dire sur les valeurs que peut prendre X)] qui provoqueront, chez le pigeon, le comportement de *picorer* ou de *ne pas picorer* [(à savoir les deux valeurs que peut prendre Y)] (ibidem, 161, mes *italiques*). En ce sens, l'explication (3.1) contient des informations causales et explicatives moins pertinentes que l'explication

(3.2), et ce, même si (3.1) cite des informations causales qui sont plus spécifiques que celles que l'on retrouve dans (3.2). Donc, dit Woodward (ibidem, 162), quand il s'agit d'évaluer l'intérêt d'une explication, ce qu'il faut considérer n'est pas les détails et la spécificité, mais les détails *pertinents* et les spécificités *pertinentes*. En effet, suivant le cadre du MI, (1) un facteur (X) est causalement et « explicativement » *inutile* si les variations de sa valeur (quand elles sont produites suite à une intervention I) ne sont pas associées avec des changements des valeurs du phénomène-explanandum (Y) (ibidem, 162). Selon le cadre du MI, de plus, (2) nous voulons citer les facteurs de variations de l'explanans (X) (quand elles sont produites suite à une intervention I) qui sont associées *avec la totalité* de la gamme des variations de l'explanandum (Y) et qui ne suggèrent pas trompeusement que certaines variations [de X (ici, les couleurs pourpre et non pourpre)] seront associées avec des changements dans l'explanandum [Y] lorsque ce n'est pas le cas. [(Dans l'exemple de Woodward, en effet, le pigeon picorerait même si la couleur de la cible n'était pas pourpre, mais d'une autre teinte de rouge)] (ibidem, 162). Comme l'indique Woodward (ibidem, 162), une explication [contenant une variable] complètement non pertinente ou *inutile* (suivant la close [1] susmentionnée) est illustrée par l'exemple classique suivant dû à Salmon (1971, 34) :

Tous les hommes prenant des pilules anticonceptionnelles (X) ne tomberont pas enceintes (Y).

Jean prend des pilules anticonceptionnelles. ($X = x_1$)

Jean est un homme.

Jean n'est pas tombé enceinte ($Y = y_1$).

Comme on peut le constater, les variations de la valeur de X (suite à une intervention I) ne provoqueront pas de changement sur Y . En effet, suite à une I sur X , si $X = x_1$ devient $X = x_0$, Y aura quand même pour valeur y_1 . Donc, X (qui se traduit par le fait de prendre [x_1] ou de ne pas prendre [x_0] de pilule contraceptive) est un facteur

inutile dans ce cas de figure. En effet, puisqu'il est un homme, Jean ne tombera jamais enceinte : Y égalera toujours $Y = y_1$.

Quant à elles, les explications causales du type de (3.1) illustrent ce que nous dit d'éviter le desideratum (2). Il faut, en effet, éviter de formuler des explications causales qui présentent des informations *non pertinentes* (comme la couleur pourpre qui est trop spécifique dans [3.1]) (ibidem, 162). En somme, dit Woodward, une bonne explication devrait citer les variations de X (quand elles sont produites par une intervention) qui sont associées à la *gamme complète* des variations des valeurs du phénomène explanandum Y dont nous essayons de rendre compte et [cette explication pertinente] ne devrait citer que ces facteurs [(de X)] (Woodward 2008a, 162).

Kendler (2008c, 134-135) reprend l'argumentaire de Woodward et l'applique à une situation imaginaire que l'on pourrait rencontrer en psychiatrie. Dans ce qui suit, j'explique cet argumentaire selon la logique woodwardienne précédente¹⁰⁹.

Supposons qu'une personne risque de développer des idéations suicidaires (Y) si elle entretient trop de pensées autodépréciatives (X). Posons que les pensées autodépréciatives (X) surviennent sur plus d'une douzaine d'activités biologiques (par exemple, un taux élevé de cortisol, une dysfonction du système sérotoninergique, une aire 25 défectueuse, une activation du noyau accumbens latéral, etc.). Ainsi, «there are many different ways one's brain can get to the state of having self-derogatory ideation» (Kendler 2008c, 134). Suivant cela, nous pourrions, par exemple, formuler les deux explications causales suivantes :

¹⁰⁹ L'argumentaire de Kendler n'apparaît pas de manière aussi systématique dans le texte de ce dernier. C'est pour cela que je dis que j'« explicite » son argument.

1. La dysfonction du système sérotoninergique (*X*) a causé l'émergence, chez cette personne, d'idéations suicidaires (*Y*).
2. La présence de pensées autodépréciatives (*X*) a causé l'émergence, chez cette personne, d'idéations suicidaires (*Y*).

Selon Kendler qui suit le raisonnement susmentionné de Woodward, même si ces deux explications causales sont potentiellement vraies, la plus *pertinente* des deux est l'explication (2). On pourrait dire, selon la logique de Woodward, que l'explication (1) est inutilement spécifique. Ce qui porte probablement Kendler à préférer l'explication (2) (considérant le fait que les pensées autodépréciatives [*X*] surviennent sur *toutes* les variables biologiques susmentionnées causant [*Y*]) est que cette dernière présente une valeur explicative causale plus grande que l'énoncé explicatif (1). En effet, si nous reformulons ces deux énoncés en termes contrastifs, on s'aperçoit qu'il y a des informations pertinentes que nous donne l'énoncé (2) mais que ne nous donne pas l'énoncé (1). Voyons ces deux énoncés :

1. Le contraste entre le fait que le système sérotoninergique soit défectueux ou non défectueux cause le contraste entre la personne développant des idéations suicidaires ou n'en développant pas.
2. Le contraste entre le fait que des pensées autodépréciatives soient présentes ou ne soient pas présentes cause le contraste entre la personne développant des idéations suicidaires ou n'en développant pas.

Selon ce que j'ai dit plus haut suivant Woodward, si on interprète l'énoncé (1) comme supposant que la personne ne développera pas d'idées suicidaires si son système sérotoninergique n'est pas défectueux, alors l'énoncé (1) est faux. En effet, dans le contexte fixé par Kendler, les pensées suicidaires pourraient être causées par l'action concertée d'autres modules du cerveau que le système sérotoninergique. De fait, l'émergence d'idées suicidaires est liée à l'activité de plusieurs des processus

biologiques susmentionnés sur lesquels surviennent les pensées autodépréciatives. Même si nous trouvons que cette interprétation n'est pas charitable, comme Woodward le disait pour celle concernant les pigeons, il reste vrai que l'explication (1) nous dit moins sur ce que nous voudrions savoir des *conditions complètes* qui provoqueront, chez la personne, le développement ou le non-développement d'idées suicidaires. En ce sens, l'explication (1) contient des informations causales et explicatives moins pertinentes que l'explication (2), et ce, même si (1) cite des informations causales qui sont plus spécifiques que celles que l'on retrouve dans (2). L'énoncé explicatif (1) « la dysfonction du système sérotoninergique (*X*) a causé l'émergence, chez cette personne, d'idéations suicidaires (*Y*) » est effectivement une généralisation causale *moins invariante* que l'énoncé explicatif (2) « la présence de pensées autodépréciatives (*X*) a causé l'émergence, chez cette personne, d'idéation suicidaire (*Y*) ». Il en est ainsi puisque, selon les paramètres qui ont été fixés par Kendler, la stabilité de la généralisation (1) est très sensible aux autres variables du contexte biologique (c'est-à-dire à l'activité des autres modules cérébraux impliqués dans l'émergence des idéations suicidaires par le biais de pensées autodépréciatives). C'est pour cette raison que Kendler dit qu'elle n'est pas, ultimement, généralisable. Comme l'explique ce dernier (2008c, 135), « ...if we intervened on each biological variable ([ici, sur le système sérotoninergique]), the effect would be highly dependent (or nongeneralizable) because of the state of all the other complex biological causes ». Contrairement, toujours selon Kendler, la généralisation (2) est plus généralisable puisqu'elle émerge¹¹⁰ (c'est l'hypothèse que fait Kendler) de l'activité de *tous* les facteurs biologiques causaux susmentionnés. Ainsi, la généralisation (2) présente un bon degré d'invariance et, je dirais même plus, un degré d'invariance plus grand que celui de la généralisation (1). En effet, on peut constater que, dans la généralisation (2) et selon les paramètres fixés par Kendler, les variations de l'explanans (*X*) suite à une intervention (*I*) sont associées *avec la totalité* de la gamme des variations de

¹¹⁰ Dans cet argumentaire, je ne fais pas de différence entre les termes « émerger » et « survenir ».

l'explanandum (Y). Par exemple, suivant Kendler (2008c, 135), en intervenant par une thérapie cognitive-comportementale (I) sur les pensées autodépréciatives (X) la quantité d'idéation suicidaire diminuera (Y) (et, par là même, je suppose, l'activité normale des systèmes biologiques susmentionnés impliqués dans la dépression sera rétablie graduellement). Bref, comme le dit Kendler,

[...] [i]n this artificial case, we can see that if the web of causation is such that a complex and interacting set of lower-order causes flows through a simple set of higher-order causes, intervening at the higher level will likely be more effective, generalizable, and just plain better. (ibidem, 135)¹¹¹

¹¹¹ Je pense que Kendler a en tête l'exemple de physique suivant quand il élabore l'idée selon laquelle, dans ce cas de figure, la généralisation de haut niveau contient plus de *variables explicatives pertinentes* pour effectuer des manipulations causales efficaces que la généralisation contenant des variables spécifiques (c.-à-d. biologiques). Voyons le passage où Kendler et Campbell donnent cet exemple :

Considérons maintenant un exemple simple en physique [...] [comme celui de] l'expansion d'un ballon provoquée quand le gaz que ce dernier contient est chauffé. Le modèle mécanique de la causalité expliquerait ce processus comme la somme des actions de chacune des molécules de gaz qui s'accélèrent et qui s'accélérent font exploser le ballon. Bien que [le mouvement de] chacune des molécules contribue à l'expansion du ballon, faire appel au mouvement individuel de chacune de ces microparticules pour expliquer l'expansion du ballon n'est pas le meilleur niveau [d'analyse] pour comprendre ce processus. *L'information relative à la trajectoire individuelle de n'importe laquelle des molécules n'est pas une information pertinente pour comprendre l'expansion ou la non-expansion du ballon.* [En effet], [l']expansion du ballon est mieux comprise selon un processus se passant à un *niveau supérieur* — l'énergie cinétique moyenne de toutes les molécules de gaz causée par la chaleur. D'intervenir dans ce processus — soit pour réduire ou accélérer l'expansion — consisterait à hausser ou à baisser la chaleur, non à modifier la trajectoire individuelle des molécules. (Kendler et Campbell 2009, 883, je traduis)

J'ai l'impression que, pour Kendler, les trajectoires individuelles des molécules sont à l'expansion ou à la non-expansion du ballon ce que les diverses activités biologiques du cerveau sont à l'émergence des idéations suicidaires. Or, de la même façon que *la trajectoire individuelle de n'importe laquelle des molécules (X) n'est pas une information pertinente pour comprendre l'expansion ou la non-expansion du ballon (Y)*, l'activité individuelle du système sérotoninergique (X) *n'est pas une information pertinente pour comprendre* l'émergence des idéations suicidaires (Y). Enfin, de la même façon que Kendler suppose que l'on est mieux d'intervenir sur la température (X) du gaz pour réduire ou accélérer l'expansion du ballon (Y), il pense que nous devrions intervenir sur les pensées autodépréciatives (X) pour faire réduire la quantité d'idéations suicidaires (Y). C'est pour cela que l'explication causale (2) « la présence de pensées autodépréciatives (X) a causé l'émergence, chez cette personne, d'idéations suicidaires (Y) » présente un potentiel explicatif et un potentiel de manipulation plus élevés que ceux que présente l'explication causale (1) « la dysfonction du système

On pourrait rétorquer, suivant Insel (2010), que c'est seulement l'aire 25 défectueuse (X) qui est à l'origine des dysfonctions du circuit mettant en jeu les modules cérébraux qui produisent les idéations suicidaires (Y) (voir section 2.6.4 de ce mémoire). Suivant cela, on pourrait affirmer qu'en intervenant (I) sur cette aire on rétablirait le fonctionnement normal de ce circuit et qu'en conséquence les idéations suicidaires (Y) diminueraient.

Il me semble que cette hypothèse causale est problématique puisque les bases biologiques exactes de la dépression majeure nous échappent encore et qu'il est improbable qu'une seule structure cérébrale (comme l'aire 25) soit à elle seule à l'origine de ce trouble. En effet, ce genre d'explication réductionniste ne convient pas à l'explication biologique de la dépression majeure et des troubles mentaux en général. Comme le dit le chercheur en neurosciences John Suckling du département en psychiatrie de l'université de Cambridge,

[...] *it looks unlikely that a reductionist approach will lead to a sufficient or useful description of MDD [(major depressive disorder)]. Indeed, this is the case more generally across the inventory of mental health disorders. Compartmental models in which areas are imbued with specific functions and sensitivities do not acknowledge the distributed and integrative nature of the brain function. As the techniques to measure the brain mature, an evolution is also needed in our conceptualization of its organization, function and dysfunction. (Suckling 2012, 6, mes italiques)*

sérotoninergique (X) a causé l'émergence, chez cette personne, d'idéations suicidaires (Y) ». Enfin, comme la généralisation invariante (2) produit des *explications causales psychologiques*, l'intervention causale sur (X) sera de nature psychologique. D'où l'idée de Kendler d'intervenir avec une thérapie cognitivo comportementale (I) sur les pensées autodépréciatives (X) afin de faire baisser la quantité d'idéations suicidaires (Y).

Cela dit, je ne voudrais pas que l'on pense que je m'oppose à la recherche des bases neurobiologiques de la dépression majeure ou de celles de tous les autres troubles de santé mentale. Au contraire, je pense qu'il faut tenter de comprendre la dynamique qu'entretiennent ces modules et processus biologiques desquels émergent les symptômes des maladies mentales. Ces systèmes sont peut-être constitués de mécanismes non agrégatifs et, pour cela, difficilement décomposables (voir section 3.7), mais il est possible d'élucider leur fonctionnement à force de beaucoup de travail et d'intelligence. Comme le dit Kendler,

[...] although this nondecomposability greatly complicates our search to understand explanatory mechanisms in psychiatry, cynicism and pessimism are as premature and unwarranted as is zealous oversimplification [...]. Causal loops are not irrevocable barriers to detailed scientific understanding [...] (Kendler 2008a, 698)

Une des façons de faire, explique Kendler, pour comprendre ces systèmes non agrégatifs, consiste à trouver en leur sein des zones locales décomposables et de les étudier en isolation du système (ibidem, 698). Dans cette perspective, par exemple, que nous ayons découvert que l'aire 25 est une des parties maîtresse du mécanisme biologique de la dépression majeure est un pas dans la bonne direction (Insel 2010). Le fait que nous ayons découvert que le système sérotoninergique était défectueux chez les gens présentant une dépression ou en ayant déjà présenté une est aussi une trouvaille intéressante (Cowen 2008). Cependant, si je suis les propos que tient Kendler dans son expérience de pensée, *inférer* un traitement pharmacologique *d'un seul mécanisme* comme celui de la dysfonction du système sérotoninergique ne sera pas très efficace. À ce propos, le seul fait que les inhibiteurs sélectifs de la sérotonine (ISRS) ou que les autres types d'antidépresseurs soient très peu ou pas efficaces dans le traitement de la dépression est un fait qui tend à conforter cette idée (sur l'efficacité

des antidépresseurs, voir Kirsch 2009; voir aussi l'argument 8 de la section 3.5 de ce mémoire)¹¹².

En somme, ce qui précède nous démontre que le MI de la causalité développé par Woodward, contrairement à celui qu'a développé Bickle¹¹³ (sections 3.3 et 3.4), n'est pas caractérisé par un fondationnalisme épistémique causal (Woodward 2008a, 146). Le MI de Woodward, en fait, est a priori neutre relativement à l'efficacité des causes d'un point de vue épistémique. Si une généralisation psychologique ou environnementale présente un degré d'invariance plus important qu'une généralisation biologique « alors soit », dirait Woodward! Et si l'inverse est le cas, « qu'il en soit ainsi », dirait-il pareillement! Ainsi, le MI de Woodward s'accommode davantage que le modèle de l'explication de Bickle avec la diversité des étiologies des troubles de santé mentale (voir McNally 2011, chap. 5 et 6 pour une revue vulgarisée de cette littérature épidémiologique).

Avant de terminer cette section, j'aimerais montrer brièvement qu'il est pertinent d'adopter le MI parce qu'il ne rencontre pas le problème de la pertinence causale

¹¹² Comme je l'ai dit à l'argument 8 de la section 3.5, de manière générale, si nous n'avons pas de traitements pharmacologiques efficaces qui ne provoquent pas d'effets indésirables (voire néfastes et dangereux pour la santé) pour soigner les troubles de santé mentale, c'est parce que notre connaissance du cerveau est insuffisante. En effet, et là je formule mon argument dans le cadre du MI, nous connaissons très peu de *généralisations invariantes* qui régissent le système (mécaniste) cérébral produisant les pensées, les émotions et les comportements. Si nous ne connaissons pas ces généralisations, c'est parce qu'au niveau cellulaire nous ne savons pas comment *tous* les neurones du cerveau sont connectés entre eux. Autrement dit, pour reprendre une expression du physicien et neuroscientifique Sebastian Seung, nous ne connaissons pas le « connectome » de chacun de nos cerveaux (Seung 2012). Si nous réussissions à déterminer et à comprendre le connectome du cerveau de chacune des personnes présentant un trouble de santé mentale, alors nous aurions plus de chance de développer des outils pharmacologiques précis. (Je suppose que cette précision serait d'autant plus grande si notre connaissance de ces connexions s'étendait au niveau biomoléculaire et, pourquoi pas, au niveau microphysique.)

¹¹³ Woodward ne fait pas référence à Bickle. C'est moi qui fais la comparaison.

(section 2.6.2) que pose la théorie nomologique de la causalité développée par Hempel (section 1.2). Reprenons l'exemple de Craver vu à la section 2.6.2 :

Tous les neurones bénis stimulés par un courant de 10 nanoampères pendant une seconde produisent des potentiels d'action.

Les neurones ont été bénis (cela signifie qu'ils ont été aspergés d'eau bénite isotonique).

Les neurones ont été stimulés par un courant de 10 nanoampères pendant une seconde.

Les neurones ont produit un potentiel d'action.

Comme nous l'avons vu, même si cet argument contient une information non pertinente (les neurones bénis), le modèle des lois de couverture admet cette explication causale. À l'inverse, le MI ne l'admet pas. Dans ce cadre interventionniste, en effet, on sait que le fait que les neurones présentent la caractéristique d'être bénis X ne peut pas être une cause de la production des potentiels d'action Y parce que même si nous empêchions les neurones d'être bénis (c'est l'intervention I), — disons en faisant en sorte que le prêtre ne réussisse pas à jeter l'eau bénite isotonique — les neurones, si on les stimulait électriquement, produiraient néanmoins des potentiels d'action. On pourrait aussi affirmer que cette relation n'est pas causale parce qu'elle ne respecte pas la clause selon laquelle l'intervention I ne doit pas être corrélée avec une autre variable C qui est une cause de Y . De fait, dans cet exemple de Craver, I (l'action de bénir les neurones) est corrélée avec une autre variable C (la stimulation de 10 nanoampères pendant une seconde) qui est une cause de Y (la production du potentiel d'action).

En définitive, comme on peut le constater (1) le MI nous permet de concevoir le contenu ou la signification des énoncés causaux du type « X cause Y » qui nous servent à décrire des mécanismes; (2) le MI n'est pas un modèle de la causalité

anthropocentriste; (3) le MI nous permet de caractériser les généralisations desquelles on peut tirer des explications causales et des possibilités de manipulations expérimentales (ou thérapeutiques); (4) le MI nous permet de montrer que, parfois, une généralisation mettant en jeu des variables *plus* spécifiques est moins pertinente, *en regard de sa valeur explicative*, qu'une généralisation mettant en jeux des variables *moins* spécifiques; (5) on peut déduire de cela que le MI ne présente pas de fondationnalisme causale d'un point de vue épistémique (à l'inverse du modèle de Bickle). J'ajouterais, au passage, que l'étendue d'application du MI est plus grande que celle du modèle mécaniste et réductionniste de Bickle. Il en est ainsi parce que le modèle de ce dernier fut explicitement développé selon les techniques et méthodes de recherche qui sont propres à la neurobiologie cellulaire et moléculaire alors que le MI est issu d'une théorie générale de l'explication causale dont le domaine de référence est, en principe, toutes les sphères de la réalité matérielle (à l'exception, peut-être, de celle décrite par la microphysique). (6) Enfin, le MI ne présente pas le problème de la pertinence causale que pose la théorie nomologique de la causalité. Pour ces raisons, je soutiens que le MI convient à la médecine mentale et qu'il est pertinent, comme le fait Kendler, de l'intégrer au modèle de l'explication mécaniste et pluraliste de ce dernier.

3.9 Conclusion du troisième chapitre

Dans ce troisième chapitre, d'abord, j'ai présenté la conception sémantique des théories scientifiques (section 3.2). Puisque le concept de modèle est central dans cette conception des théories, j'ai présenté cette notion (section 3.2.1). Plus précisément, j'ai présenté trois versions de cette notion, à savoir celle de Suppes, de Van Fraassen et de Giere. De ces trois notions, j'ai retenu celle de Giere puisqu'elle m'a semblé la plus pertinente pour rendre compte de l'activité théorique en

psychiatrie. J'ai par la suite présenté les notions de mécanisme et de modèle mécaniste. J'ai adopté celles qu'a développées Glennan (section 3.2.2). À cette occasion, j'ai affirmé que les psychiatres et les chercheurs s'expliquaient les maladies mentales selon deux types de modèle mécaniste, l'un réductionniste, l'autre pluraliste. J'ai présenté le modèle de l'explication réductionniste et mécaniste formulé par John Bickle (sections 3.3 et 3.4). J'ai ensuite fait voir que le modèle de l'explication de ce dernier n'était pas approprié pour la psychiatrie (section 3.5). Ensuite, en raison des insuffisances du modèle de Bickle, j'ai présenté le modèle de l'explication pluraliste et mécaniste de Kendler (sections 3.6 et 3.7), un modèle adéquat pour formuler des explications en psychiatrie, selon moi. Comme nous l'avons vu, ce modèle peut être bonifié du MI développé par Woodward (sections 3.8 et 3.8.1).

CONCLUSION GÉNÉRALE DU MÉMOIRE

Dans ce mémoire, j'ai évalué l'intérêt de trois modèles de l'explication pour la psychiatrie, à savoir (1) les modèles DN et IS (c.-à-d., les modèles des lois de couverture) que j'ai interprétés de manière réductive (pour une présentation classique de ces modèles, voir Hempel 1965a [1942], 1965b; Hempel et Oppenheim 1948); (2) le modèle de l'explication réductionniste et mécaniste développé par Bickle (2003, 2006, 2007, 2008a, 2008b) et (3) le modèle de l'explication pluraliste et mécaniste conçu par Kendler (2001, 2005, 2008a, 2012; Kendler et Campbell 2009; Kendler, Zachar et Craver, C. 2011).

J'ai rejeté l'interprétation réductionniste du modèle des lois de couverture principalement en raison de problèmes formels. En effet, (1) j'ai montré qu'il est impossible, au sein du cadre syntaxique des théories scientifiques, de réduire des lois de haut niveau par des lois de bas niveau (voir Fodor 1974). Il en est ainsi parce que les espèces naturelles que l'on retrouve dans les théories de haut niveau sont réalisées de façon multiple au niveau moléculaire. Pour cette raison, j'ai conclu que nous ne sommes pas en droit d'utiliser seulement des lois biomoléculaires dans nos explications DN et IS des *phénomènes naturels*. Par ailleurs, (2) j'ai ajouté à ce constat que même si nous réussissions à réduire ces lois psychiatriques aux lois biomoléculaires, on ne pourrait toujours pas utiliser les modèles de lois de couverture en psychiatrie. En effet, ces derniers rencontrent des problèmes importants lorsqu'il

s'agit d'expliquer le fonctionnement du cerveau. D'abord, suivant Craver (2007), j'ai montré qu'il est possible de citer des variables non pertinentes dans les explications neurales DN et IS sans que cela affecte la force de ces explications. Ensuite, j'ai souligné, toujours suivant Craver (ibidem, 39-40), que l'exigence de prévisibilité nomique qui est au cœur des modèles de lois de couverture nous empêchait de dire que la production d'un potentiel d'action cause (et explique) la libération de neurotransmetteurs. Enfin, (3) j'ai fait voir que même s'il était possible de construire une explication empirique IS (ou DN) ne présentant pas de difficultés formelles, il resterait un problème. En effet, à mon avis, une explication — qu'elle soit réductionniste ou pluraliste — formulée en termes d'entités linguistiques évoluant dans un réseau nomique est une mauvaise façon de se représenter les pathologies en médecine mentale.

Comme les trois problèmes que je viens de nommer découlent du fait que la réduction interthéorique et les modèles des lois de couverture s'inscrivent dans la conception syntaxique des théories scientifiques, j'ai proposé de concevoir l'explication en psychiatrie selon une autre conception des théories, à savoir la conception sémantique des théories scientifiques. Cette conception repose sur l'idée selon laquelle les théories scientifiques sont des familles de *modèles*. J'ai adopté la conception des modèles développée par Giere (1988, 1999a, 1999b, 2006 et 2010) et j'ai proposé de concevoir les maladies mentales, suivant Glennan (2002, 2005), selon des *modèles mécanistes* qui peuvent être réductionnistes (Bickle) ou pluralistes (Kendler).

Dans le cadre du modèle de l'explication réductionniste et mécaniste de Bickle, les troubles de santé mentale sont toujours et fondamentalement des maladies du cerveau. Pour cette raison, une explication complète des troubles de santé mentale ne devrait pas citer de variables psychologiques, peu ou pas de variables

environnementales, et beaucoup de variables neurobiologiques. J'ai rejeté cette thèse épistémique pour diverses raisons empiriques (section 3.5).

Enfin, j'ai présenté et défendu le modèle de l'explication pluraliste et mécaniste développé par Kendler. Ce modèle ne rencontre pas les problèmes auxquels fait face le modèle de Bickle. Dans le cadre du modèle de Kendler, les maladies mentales sont conçues comme (1) des réseaux de mécanismes causaux (2) interagissant entre eux (3) se renforçant mutuellement et (4) produisant les symptômes du trouble de santé mentale (Kendler, Zachar et Craver 2010, 1143). Le modèle de Kendler explique les troubles de santé mentale en *représentant* les mécanismes multiniveaux d'où émergent ces troubles. Enfin, le modèle de Kendler s'harmonise bien au MI de la causalité (Woodward (2003, 2008a). Comme nous l'avons vu, le MI ne présente pas le problème de la pertinence (comme celui de Hempel) et n'est pas caractérisé par un fondationnalisme causal (comme celui de Bickle).

Comme on a pu le constater à la lecture de ce mémoire, les outils de la philosophie des sciences peuvent nous aider, entre autres, à expliciter les fondements d'une discipline. Pour ma part, j'ai essayé d'éclairer ce qu'est une explication adéquate d'un trouble de santé mentale. Je souhaiterais que tous les psychiatres fassent ce type d'exercice cognitif. La pratique clinique de ces derniers ne pourrait que bénéficier de cette réflexion. Vu tout cela, comme l'affirme Kendler (2008b, 4, ma traduction), on peut penser que « [...] les philosophes peuvent être de bons alliés et de bons collaborateurs en psychiatrie au 21^e siècle ».

APPENDICE A

LES CINQ PRINCIPES DE KANDEL

Eric Kandel (2002 [1998]), dans son article « Un nouveau cadre conceptuel de travail pour la psychiatrie », formule cinq principes réductionnistes qui, selon lui, devraient fonder la psychiatrie moderne. Ces principes s'harmonisent très bien avec l'épistémologie réductionniste et la thèse empirique « neurocausale » de Bickle. Dans ce qui suit, je cite ces cinq principes puis je les commente.

1. Tous les processus mentaux, y compris les processus psychologiques les plus complexes, dérivent d'opérations dans le cerveau. L'axe central de cette perspective est que ce que nous appelons communément l'esprit est un ensemble de fonctions portées à l'extérieur par le cerveau. Les actions du cerveau ne concernent pas seulement des comportements moteurs relativement simples, comme marcher ou manger, mais l'ensemble des actions cognitives complexes, conscientes et inconscientes, que nous associons avec le comportement spécifiquement humain, comme penser, parler, créer une œuvre littéraire, musicale et artistique. En corollaire, *les troubles du comportement qui caractérisent les maladies psychiatriques sont des troubles de la fonction cérébrale, même dans les cas où les causes de ces désordres sont clairement d'origine environnementale*¹¹⁴.
2. Les gènes et leurs productions protéiniques sont des déterminants importants de l'organisation des interconnexions (sic) entre neurones dans

¹¹⁴ Mes italiques.

le cerveau et du détail de leur fonctionnement. Les gènes, et en particulier les combinaisons de gènes, exercent ainsi un contrôle significatif sur le comportement. Comme corollaire, une des composantes qui contribuent au développement des principales maladies mentales est génétique.

3. L'altération des gènes n'explique pas, en elle-même, l'ensemble de la variabilité des troubles mentaux majeurs. Des facteurs sociaux ou développementaux y contribuent également de façon très importante. De la même façon que des combinaisons de gènes contribuent au comportement, y compris le comportement social, de même le comportement et les facteurs sociaux exercent en retour des actions sur le cerveau qui modifient l'expression des gènes et ainsi la fonction des cellules nerveuses. L'apprentissage, y compris celui qui résulte d'un comportement anormal, produit des modifications de l'expression génétique. Ainsi tout ce qui concerne la « culture » est-il finalement exprimé comme de la « nature ».
4. Les modifications de l'expression génétique produites par l'apprentissage induisent des changements dans les formes des connections (sic) neuronales. Ces changements contribuent non seulement à la base biologique de l'individualité, mais sont probablement responsables de l'initiation et du maintien d'anomalies du comportement qui sont induites par les contingences sociales.
5. Puisque la psychothérapie ou le conseil a un effet réel et produit des changements à long terme dans le comportement, elle le fait probablement à travers l'apprentissage, en produisant des transformations dans l'expression génétique qui modifient la force des connections (sic) neuronales, et induisent des modifications structurelles de la forme anatomique des interconnexions des cellules nerveuses dans le cerveau. Comme la résolution de l'imagerie cérébrale s'améliore, elle pourrait éventuellement rendre possible une évaluation quantitative des effets de la psychothérapie.

Le *premier postulat* stipule que tous les processus cognitifs résultent de processus neuraux. Autrement dit, s'il n'y a pas de cerveau, il n'y a pas de pensée ou de conscience ou d'émotions. Kandel défend ici la *thèse ontologique réductionniste* selon laquelle tous les processus cognitifs *sont* des processus neuraux. Kandel

formule cette thèse de la manière suivante : « l'esprit est un ensemble de fonctions portées à l'extérieur par le cerveau » (ibidem, 19). Cette thèse est aussi soutenue par le généticien, biologiste et prix Nobel de médecine Francis Crick qui, selon moi, la formule plus clairement :

you, your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules (*mes italiques*). (Crick (1994) cité dans Noë, 2009, 5)

Le *deuxième postulat* m'apparaît être un raffinement réductionniste du premier puisqu'il nous apprend que le développement des aires et régions cérébrales et, par conséquent, des fonctions du cerveau, est ultimement dû aux gènes. La recherche concernant ces derniers, de fait, nous renseigne sur la constitution et les fonctions du système cérébral. En effet, dit Kandel,

Collectively the methods of molecular genetics have provided more new information and a deeper understanding of almost all facets of neural structure and function than at any period in the past half century or more. (Cowan et Kandel, 2001, 596)

Ce fait porte Kandel à affirmer que la composante principale qui contribue au développement des troubles de santé mentale doit être génétique. Ce postulat montre que Kandel, comme Bickle, accorde une efficience causale primordiale aux gènes — lesquels sont des entités biomoléculaires — dans la détermination des comportements et des fonctions (cognitives, émotives, etc.) et, en conséquence, dans le développement *des principaux troubles de santé mentale*. Pour cette raison, la recherche en psychiatrie doit prioriser un *réductionnisme méthodologique*. De fait, trouver les gènes à l'origine de ces maladies, comprendre ce que ces gènes produisent

et dans quelles conditions biologiques ceux-ci s'expriment est — bien que très laborieuse — une quête primordiale. De fait, Kandel et Cowan affirment :

The identification of the gene involved in a disorder is an essential first step toward its elucidation; but it is only the first step, and in many cases it has proved to be the easiest step. Discovering the product of the gene, when and where it is expressed, how its production is regulated, and what it actually does are usually much more difficult. Further progress toward these ends will likely depend heavily on developments in the newly emerging fields of functional genomics and proteomics. (Cowan et Kandel, 2001, 596-597)

On pourrait considérer que le *troisième postulat* n'est pas réductionniste, mais il l'est. De fait, comme l'explique Murphy (Murphy, 2006, 116), bien que Kandel postule que les facteurs sociaux et psychologiques engendrent une altération de l'expression génétique, ce dernier dit en même temps que tout ce qui concerne la « "culture" est [...] finalement exprimé comme de la "nature" » (Kandel, 2002 (1998), 19) et que « les gènes déterminent le phénotype » (ibidem, 22). Essentiellement, cela signifie que nous pouvons expliquer entièrement le phénotype avec le génotype. Cela veut dire, en d'autres termes, que nous pouvons expliquer totalement la culture et la psyché humaine en étudiant les gènes de l'homo sapiens. Dans cette perspective, remarque Murphy (2006, 116), les théories de haut niveau (psychologiques, sociologiques, anthropologique, etc.) ne présentent aucune autonomie explicative en regard des troubles de santé mentale. De fait, ils ne font que nous aider dans la découverte des processus que la biologie moléculaire finira par expliquer entièrement (Murphy, 2006, 116; voir aussi Schaffner K.F cité dans Faucher, Poirier et Lachapelle, 2006, 158). Dans cette perspective, en regard des troubles de santé mentale, il est plausible de penser que la force causale des événements biomoléculaires éclipse celle des événements environnementaux. Comme Murphy, on pourrait dire, pour cela, que Kandel — et, selon moi, Bickle — adopte une position métaphysique donnant une

priorité causale au niveau moléculaire sur les autres niveaux de réalité (Murphy 2010a, section 2.3).

Le *quatrième postulat* est un corollaire des trois premiers. De fait, si les processus cognitifs et affectifs (l'esprit) sont des processus neuraux ayant lieu dans le cerveau (postulat 1), si le développement de ce cerveau est déterminé par les gènes (postulat 2) et si les gènes déterminent et expliquent toujours causalement le phénotype (postulat 3), alors on peut déduire qu'un individu qui a développé un trouble de santé mentale à la suite d'une expérience traumatisante (qui est un apprentissage [postulat 4]) présentait nécessairement des *anomalies génétiques* ou une *fragilité génétique* qui le rendaient plus vulnérable lors de cette expérience. On peut illustrer cette idée en considérant ce que pense Kandel de la maladie du choc post-traumatique. De fait, ce dernier affirme que, bien qu'elle soit acquise par expérience, cette maladie est susceptible d'apparaître davantage chez ceux qui étaient *génétiquement prédisposés à la développer* [mes italiques] (Kandel, 2002 [1998], 29). Il me semble que cette compréhension du trouble sous-tend une posture réductionniste parce qu'elle nie, comme nous l'avons vu dans le troisième postulat, l'autonomie explicative des théories de haut niveau comme, par exemple, les théories psychologiques et sociologiques. De fait, dans le cas d'un soldat qui revient de guerre, ce n'est pas d'avoir vu et vécu des horreurs qui explique causalement, en dernière instance, l'apparition de la maladie du trouble du choc post-traumatique, mais plutôt la « fragilité » génétique du soldat. Comme nous l'avons vu dans le troisième postulat, Kandel, à l'instar de Bickle, semble accorder une efficience causale primordiale aux événements biomoléculaires. Ceci concorde, je pense, avec ce que soutient ce Nobel de médecine dans la première clause, à savoir que « les troubles du comportement qui caractérisent les maladies psychiatriques sont des troubles de la fonction cérébrale, même dans les cas où les causes de ces désordres sont clairement d'origine environnementale » (Kandel, 2002 [1998], 19).

Enfin, le *cinquième postulat* obéit à la même logique que le quatrième. De fait, puisqu'une psychothérapie est une expérience constructive et bénéfique (en théorie), alors elle implique une transformation « favorable » de l'expression génétique chez l'individu et, corollairement, elle engendre une transformation physique du cerveau. Cette transformation du système nerveux consiste en une réorganisation cérébrale qui, selon ma compréhension des choses, « renforce » certaines structures susceptibles de « protéger » les structures cérébrales plus vulnérables de l'individu. Par exemple, une personne très anxieuse pourra développer des stratégies cognitives protectrices qui « renforceront » son lobe préfrontal. Ce faisant, cet individu aura plus de contrôle sur son amygdale qui est, pour des raisons génétiques, généralement hyperactive.

En somme, on constate que les idées de Kandel concernant les maladies mentales s'harmonisent très bien avec l'épistémologie réductionniste et la théorie empirique neurocausale de Bickle. Dans ce cadre, en dernière instance, une explication complète citera essentiellement les processus biomoléculaires provoquant les dysfonctions cognitives et émotives caractérisant les troubles de santé mentale. Pour l'instant, cependant, puisque nous connaissons peu les mécanismes biomoléculaires à l'origine de la plupart des troubles de santé mentale, cette explication sera accompagnée très souvent de schémas représentant les systèmes d'aires cérébrales jugées dysfonctionnelles (voir la section 2.6.4 pour ce type d'explication).

APPENDICE B

LES HUIT PRINCIPES DE KENDLER

Kendler a formulé huit principes qui rendent bien compte de ses idées pluralistes cardinales relatives aux troubles de santé mentale (Kendler 2005, 433, je traduis) :

1. La psychiatrie est irrévocablement enracinée dans le mental, c'est-à-dire dans l'expérience à la première personne.
2. Le dualisme cartésien des substances est faux.
3. L'épiphénoménalisme est faux.
4. La causalité cerveau→esprit¹¹⁵ et esprit→cerveau est réelle.
5. Les troubles psychiatriques sont étiologiquement complexes, et on ne s'attend plus à des découvertes [...] qui permettraient d'expliquer l'origine de ces troubles en termes simples.
6. Les *approches explicatives* pluralistes sont préférables aux approches monistes explicatives, et en particulier au réductionnisme biologique.
7. La psychiatrie [...] [doit] aborde[r] la complexité de manière empiriquement rigoureuse avec des *modèles explicatifs* pluralistes.
8. Enfin, nous devons accepter le « réductionnisme incomplet » [patchy reductionism], c'est-à-dire que nous devons tenter d'intégrer et d'expliquer graduellement les étiologies complexes qui caractérisent les maladies psychiatriques.

¹¹⁵ Il faut considérer la notion d'esprit en termes épistémiques. Kendler ne soutient pas que l'esprit est une substance autre que celle du corps (voir Kendler 2001).

On peut résumer la substance de ces postulats de la manière suivante : *postulat 1* : bien que les neurosciences et la biologie moléculaire soient importantes pour comprendre les troubles de santé mentale, la vie de l'esprit du patient l'est tout autant (Kendler 2005, 433-434); *postulat 2* : si nous rejetons le dualisme, et adoptons le monisme biologique — c'est ce que fait Kendler —, il s'ensuit que toutes les maladies mentales résultent¹¹⁶ essentiellement de processus biologiques (comme tout ce que nous faisons, d'ailleurs) (ibidem, 434); *postulat 3* : cette proposition veut simplement dire que les croyances des individus — lesquelles croyances *sont* des processus biologiques — ont un pouvoir causal. En effet, Kendler, me semble-t-il, soutient la théorie de l'identité psychoneurale. Notons que, dans ce cadre, les états mentaux existent. En effet, comme le dit Rachel Cooper, les tenants de la théorie de l'identité psychoneurale (contrairement aux tenants du matérialisme éliminativiste) « [...] accept mental states so long as they are material » (Cooper 2007, 114; voir aussi Kendler 2001); *postulat 4* : dans le cadre de la théorie de l'identité psychoneurale, ce postulat signifie que (a) certaines activités biologiques du cerveau ont une influence sur (b) certaines activités biologiques du cerveau sur lesquelles surviennent l'esprit et vice versa¹¹⁷; *postulat 5* : ce principe signifie que la grande majorité des troubles de santé mentale sont étiologiquement complexes. En effet, l'étiologie de la plupart des maladies mentales *ne ressemble pas* à celle, par exemple, de la maladie de Huntington qui est causée par une seule anomalie génétique (voir section 3.8.1); *postulat 6* : à mon sens, il s'agit d'un corollaire épistémologique du cinquième postulat. De fait, si les troubles de santé mentale *sont causés* par plusieurs facteurs que l'on retrouve sur différents niveaux de la réalité (c'est la complexité

¹¹⁶ Selon ma compréhension des choses, le verbe « résulte » ne signifie pas « est causé ». Murphy affirme à peu près la même chose, il me semble, quand il écrit : « nous préférons dire que la neuropathologie fait advenir [realizes] la maladie, plutôt qu'elle ne la cause » (Murphy, 2009, 113, ma traduction).

¹¹⁷ Selon moi, cette conception de Kendler de la relation qu'entretiennent le corps et l'esprit est très proche de celle de Bunge. En effet, dit Bunge, « [...] the so called psychophysical relations are interactions between different subsystems of the CNS [(central nervous system)], or between them and other components of the organism » (Bunge 2001 [1977], 86).

étiologique), nous devons utiliser diverses théories (biologiques, psychologiques, sociales, etc.) pour appréhender ces étiologies; *postulats 7* : il nous faut un *modèle de l'explication* capable de rendre compte de cette diversité et de cette complexité causale. Ce modèle est le modèle de l'explication pluraliste et mécaniste de Kendler bonifié du MI de Woodward (voir sections 3.6, 3.7 et 3.8) (voir la section 2.6.4 pour un exemple de modèle de la dépression majeure chez les femmes; voir aussi Kendler 2008b, 12-13 pour un autre exemple); *postulat 8* : selon ma lecture de Kendler, comme nous sommes loin du moment où nous serons en possession d'un modèle complet de chacune des maladies mentales (Kendler 2005, 138), ce postulat signifie que nous devons travailler à développer et intégrer graduellement les différentes explications (biomoléculaire, psychologique, sociale, familiale, etc.) que nous avons d'une même affection de la psyché. Suivant Schaffner, Kendler nomme cette démarche ou cette approche le « patchy reductionism » (ibidem, 138).

RÉFÉRENCES

- Akil, et coll. 2010. « The Future of Psychiatric Research: Genomes and Neural Circuits ». *Science*, vol. 327, n° 5973, p. 1580-1581.
- Allan V. Horwitz et J. C. Wakefield. 2007. *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. New York : Oxford University Press.
- Andrew, P.W et Anderson T. Jr. 2010. « Depression's Evolutionary Roots ». *Scientific American Mind*, vol. 20, n° 7 (janvier/février), p. 56-61.
- Association américaine de psychiatrie. 2003. *Manuel diagnostique et statistique des troubles mentaux. Texte révisé (DSM-IV-TR), 2000*, Paris, Masson.
- Bais, S. 2012. *Une relativité bien particulière précédé de Les équations fondamentales de la physique*. Paris : Gallimard (Folio essais).
- Barberousse, A, Kistler, M. et P. Ludwig 2000. *La philosophie des sciences au 20^e siècle*. Paris : Flammarion.
- Barondes, S. 1999. *Molecules and Mental Illness* (couverture souple). New York : Scientific American Press.
- Batterman, R. 2006. « Reduction ». In *Encyclopedia of Philosophy* (deuxième édition), D. Borchert (dir.). Detroit : Macmillan Reference USA.

- Batterman, R. 2012. « Intertheory Relations in Physics ». In *The Stanford Encyclopedia of Philosophy* (Automne 2012), E. N. Zalta (dir.), URL=<http://plato.stanford.edu/archives/fall2012/entries/physics-interrelate/>>. Consulté le 8 juillet 2012.
- Bear, M. F., Connors, B. W. et M. A. Paradiso. 2002. *Neurosciences à la découverte du cerveau* (deuxième édition). Paris : Éditions Pradel.
- Bechtel, W. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London : Routledge.
- Bechtel, W. et A. Hamilton. 2007. « Reductionism, Integration, and the Unity of the Sciences » in *Philosophy of Science: Focal Issues*, T. Kuipers (dir.), p. 377-430. The Netherlands : Elsevier.
- Bechtel, W. et R. C. Richardson. 2010 [1993]. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research* (deuxième édition). Cambridge, MA : MIT Press/Bradford Books.
- Bechtel, W. et R. C. Richardson. 2010. « Introduction : Discovering Complexity — Further Perspectives ». Introduction in *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research* (deuxième édition), p.xvii-liv. Cambridge, MA : MIT Press/Bradford Books.
- Bickle, J. 1998. *Psychoneural Reduction: The New Wave*. Cambridge, MA : MIT Press.
- Bickle, J. 2002. « Concepts Structured Through Reduction: A Structuralist Resource Illuminates the Consolidation — Long-Term Potentiation (Ltp) Link ». *Synthese*, vol. 130, n° 1, p. 123-133.
- Bickle, J. 2003. *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Norwell, MA: Kluwer Academic Press.
- Bickle, J. 2006. « Reducing Mind to Molecular Pathways: Explicating the Reductionism Implicit in Current Cellular and Molecular Neuroscience ». *Synthese*, vol. 15, p. 411-434.

- Bickle, J. 2007. « Ruthless Reductionism and Social Cognition ». *Journal of Physiology* (Paris), vol. 101, p. 230-235.
- Bickle, J. 2008a. « Vous avez dit réalisation multiple? Je réponds neurosciences moléculaires ». In *Des neurones à la philosophie : Neurophilosophie et philosophie des neurosciences*. L. Faucher et P. Poirier (dir.), p.181-204. Paris : Syllepse.
- Bickle, J. 2008b. « The Molecules of Social Recognition Memory: Implications for Social Cognition, Extended Mind, and Neuroethics ». *Consciousness and Cognition*, vol. 17, n° 2, p. 468-474.
- Bird, A. et E. Tobin. 2008. « Natural Kinds ». In *The Stanford Encyclopedia of Philosophy* (Édition été 2010), E. N. Zalta (dir.). URL = <http://plato.stanford.edu/archives/sum2010/entries/natural-kinds/>.
Consulté le 7 octobre 2012
- Bolton, D. 2008. *What is Mental Disorder? : an Essay in Philosophy, Science, and Values*. Oxford, New York : Oxford University Press.
- Bonnay, D. 2011. « L'explication scientifique ». In *Précis de philosophie des sciences*, A. Barberousse, D. Bonnay et M. Cozic (dir.), p. 13-61. Paris : Vuibert.
- Boyd R. 1991. « Realism, Antifoundationalism and the Enthusiasm for Natural Kinds ». *Philosophical Studies*, vol. 61, p. 127-148.
- Boyd R. 1999. « Homeostasis, Species, and Higher Taxa ». In *Species : New Interdisciplinary Essays* (dir. R. A. Wilson), p. 141-185. A Bradford Book/MIT Press : Cambridge, MA.
- Brigandt, Ingo et A. Love. 2012. « Reductionism in Biology ». In *The Stanford Encyclopedia of Philosophy* (Édition été 2012), E. N. Zalta (dir.), URL = <http://plato.stanford.edu/archives/sum2012/entries/reduction-biology/>.
Consulté le 8 juillet 2012.

- Brightwell, J. J., Smith, C. A., Countryman, R. A., Neve, R. L., et P. J Colombo. 2005. « Hippocampal Overexpression of Mutant CREB Blocks Long-Term but not Short-Term Memory for a Socially Transmitted Food Preference ». *Learning and Memory*, vol.12, p. 12-17.
- Bromberger, S. 1966. « Why-Questions ». In *Readings in the Philosophy of Science*, Baruch A. Brody (dir.), p. 66-84. Englewood Cliffs: Prentice Hall, Inc.
- Broome, M. et L. Bortolotti (dir.). 2009. *Psychiatry as Cognitive Neuroscience: philosophical perspectives*. Oxford University Press.
- Bunge, M. 1998 [1967]. *Philosophy of Science. From Explanation to Justification*. (Édition révisée). New Brunswick (É-U) et Londre (UK): Transaction Publishers.
- Bunge, M. 2001 [1977]. « Emergence and the Mind ». In *Selected Essays of Mario Bunge. Scientific realism*, M. Mahner (dir.), p. 75-91, New York : Prometheus Books.
- Bunge, M. 2001 [1982]. « The Revival of Causality ». In *Selected Essays of Mario Bunge. Scientific realism*, M. Mahner (dir.), p.57-74, New York : Prometheus Books.
- Bunge, M. 2003a. *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge*. Toronto : University of Toronto Press.
- Bunge, Mario. 2003b. *Philosophical Dictionary* (édition augmentée). New York : Prometheus Books.
- Bunge, M. 2004. *Matérialisme et humanisme. Pour surmonter la crise de la pensée*. Édition Liber : Montréal.
- Bunge, M. 2006. *Chasing Reality. Strife for Realism*. Toronto : University of Toronto Press.

Carnap, R. 1950. « Empirism, Semantics, and Ontology » (transcrit en hypertexte par Andrew Chrucky le 19 septembre 1997).

URL : http://philosophy.ru/library/carnap/02_eng.html

Consulté le 8 juillet 2012.

Carnap, R. 1995 [1966]. *An Introduction to the Philosophy of Science*. New York : Dover Publications, Inc.

Caspi, A. et coll. 2003. « Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene ». *Science*, vol. 301, p. 386 -389.

Caspi A. et Terrie E. Moffit. 2006. « Gene-Environment Interactions in Psychiatry: Joining Forces with Neuroscience ». *Nature Reviews Neuroscience*, vol.7 (juillet), p. 583-590.

CBC NEWS. 19 février 2012. « Treating Depression: Is there a Placebo Effect? ».

URL : <http://www.cbsnews.com/video/watch/?id=7399362n>

Consulté le 7 octobre 2012.

Chalmers F., A. 1987. *Qu'est-ce que la science? Popper, Kuhn, Lakatos, Feyerabend*. Paris : La découverte.

Churchland, P. 1979. *Scientific Realism and the Plasticity of Mind*. London, New York, Melbourne : Cambridge U. P.

Cooper, R. 2005. *Classifying Madness: A Philosophical Examination of Diagnostic and Statistical Manual of Mental Disorders*. Dordrecht, Berlin, Heidelberg, New York : Springer.

Cooper, R. 2007. *Psychiatry and Philosophy of Science*. Montréal et Kingston : McGill-Queen's University Press.

Countryman, R. A., Orlowski, J. D., Brightwell, J. J., Oskowitz, A. Z., et P. J. Colombo. 2005. « CREB Phosphorylation and c-Fos Expression in the Hippocampus of Rats During Acquisition and Recall of a Socially Transmitted Food Preference ». *Hippocampus*, vol. 15, p. 56-67.

- Cowan, W. M., et E. Kandel. 2001. « Prospects for Neurology and Psychiatry ». *Journal American Medical Association*, vol.285, p. 594-600.
- Cowen, P. J. 2008. « Serotonin and Depression: Pathophysiological Mechanism or Marketing Myth? ». *Trends in Pharmacological Sciences*, vol. 29, n° 9, p. 433-436.
- Craver, C. F. 2007. *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience*. Oxford : Oxford University Press.
- Dawkins, R. 1982. *The Extended Phenotype*. New York : W.H. Freeman.
- Dizadji-Bahmani, F., R. Frigg et S. Hartmann. 2010. « Who's Afraid of Nagelian Reduction? ». *Erkenntnis*, vol. 73, n° 3, p. 393-412.
- Dubé, C. 2007. « Médicaments. La vie en Rose, bleu, blanc ». *Québec Sciences*, vol.45, n° 6, p. 42-46.
- Endicott, R. 1998. « Collapse of the New Wave ». *Journal of Philosophy*, vol. 95, p. 53-72.
- Endicott, R. 2001. « Post-Structuralist Angst-Critical Notice: John Bickle, Psychoneural Reduction: The New Wave ». *Philosophy of Science*, vol. 68, p. 377-393.
- Eronen, M. I. 2010. *Reduction in Philosophy of Mind: A Pluralistic Account*. Thèse de doctorat, University of Osnabrück.
- Faucher, L. 2005. « Evolutionary Psychiatry and Nosology ». *Les Cahiers du LANCI*, vol. 4 (novembre 2005), p. 1-19.
- Faucher, L. (dir.). 2006a. « Philosophie des psychopathologies ». Numéro spécial de *Philosophiques*.
- Faucher, L. 2006b. « What's Behind a Smile? The Return of Mechanism: Reply to Schaffner ». *Synthese*, vol. 151 : p. 403-409.

- Faucher, L., P. Poirier et J. Lachapelle. 2006. « La théorie des systèmes développementaux et la construction sociale des maladies mentales ». *Philosophiques*, vol. 33, n° 1, Printemps 2006, p. 147-182.
- Feldman R.S., S. Giroux et F. Cauchy. 1994. Introduction à la psychologie. Approches contemporaines. Montréal : McGraw-Hill.
- Feyerabend, P. K. 1962. « Explanation, Reduction, and Empiricism ». In *Minnesota Studies in the Philosophy of Science*, vol. 3, G. Maxwell and H. Feigl (dir.), p. 28-97. D. Reidel Publishing Company.
- Feyerabend, P. K. 1965. « Problems of Empiricism ». In *Beyond the Edge of Certainty*, R.G Colodny (dirs.), p. 145-260. Englewood Cliffs, N.J.: Prentice-Hall, Inc.
- Feyerabend, P. K. 1979 [1975]. *Contre la méthode. Esquisse d'une théorie anarchiste de la connaissance*. Paris : Éditions du Seuil.
- Fine, A. 1986. *The Shaky Game: Einstein, Realism and The Quantum Theory*. Chicago : University of Chicago Press.
- Fodor, J. 1974. « Special Sciences: Or the Disunity of Science as a Working Hypothesis ». *Synthese*, vol. 28, p. 97-115
- Fuchs T. 2009. « Embodied Cognitive Neuroscience and its Consequences for Psychiatry ». *Journal of Pollesis and Praxis*, vol. 6, p. 219-233.
- Giere, R. N. 1988. *Explaining Science. A Cognitive Approach*. Chicago. Chicago : The University of Chicago Press.
- Giere, R. N. 1999a. *Science without Laws*. Chicago et Londres : The University of Chicago Press.
- Giere, R. N. 1999b. « Using Models to Represent Reality ». In *Model-Based Reasoning in Scientific Discovery*, L. Magnani, N. J. Nersessian, et P. Thagard (dir.), p. 41-57. New York : Kluwer/Plenum.
- Giere, R. N. 2004. « How Models are Used to Represent Reality ». *Philosophy of Science*, vol. 71, n° 5, p. 742-752.

- Giere, R. N. 2006. *Scientific Perspectivism*. Chicago : The University of Chicago Press.
- Giere, R. N. 2010. « An Agent-Based Conception of Models and Scientific Representation ». *Synthese*, vol. 172, p. 269-281.
- Glennan, S. 1996. « Mechanisms and the Nature of Causation ». *Erkenntnis*, vol. 44, p. 49-71.
- Glennan, S. 2002. « Rethinking Mechanistic Explanation ». *Philosophy of Science*, vol. 69 (Suppl.), S342-S353.
- Glennan, S. 2005. « Modeling Mechanisms ». *Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 36, p. 443-464.
- Glennan, S. 2008. «Mechanisms». In *The Routledge Companion to the Philosophy of Science*, S. Psillos et M. Curd (dir.), p. 375-384. London; New York°: Routledge.
- Glennan, S. 2011. « Singular and General Causal Relations: A Mechanist Perspective». In *Causality in the Sciences*, P. McKay Illari, F. Russo et J. Williamson (dir.), p. 789-817. New York : Oxford University Press.
- Gold I. 2009. « Reduction in Psychiatry ». *Canadian Journal of Psychiatry*, vol. 54 n° 8, p. 503-512.
- Gold, I. 2011. « Re : Reduction in Psychiatry ». *Canadian Journal of Psychiatry*, vol. 56, n° 2, p. 125-126.
- Gouvernement du Québec. 2011. *Portrait de l'usage des antidépresseurs chez les adultes assurés par le régime public d'assurance médicaments du Québec* (janvier 2011), Québec, Institut national d'excellence en santé et en services sociaux.
- Graham, G. et Stephens Lynn. 2006. « Reconcevoir le délire ». *Philosophiques*, vol. 33, n° 1 (printemps), p. 183-195.

- Haines, E. 2006. *Knives Don't Have Your Back*. Toronto, Ontario : Last Gang Records. 1 disque sonore (45 min, 45 s) : numérique; 12 cm + 1 livret.
- Hempel, C. et P. Oppenheim. 1948. « Studies in the Logic of Explanation ». *Philosophy of Science*, n° 15. p. 135-175.
- Hempel, C. 1965a [1942]. « The Function of General Laws ». In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, p. 231-244. New York : Free Press.
- Hempel, C., 1965b. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press.
- Hempel, C., 1965c. « Aspects of Scientific Explanation ». In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, p. 331-496. New York : Free Press.
- Hempel, C. 1972 [1966]. *Éléments d'épistémologie*. Paris : Armand Colin.
- Higgins, E S. 2009. « Do ADHD Drugs Take a Toll on the Brain? ». *Scientific American Mind*, vol.20, n° 4 (juillet/août), p. 38-43
- Hooker, C. A. 1981. « Towards a General Theory of Reduction ». *Dialogue*, vol. 20, p. 35-59, 201-236 et 496-529.
- Insel T. R. 2010. « Faulty Circuits ». *Scientific American*, vol. 302 (avril 2010), n° 4, p. 44-51.
- Insel T. R. et R. Quirion. 2005. « Psychiatry as a Clinical Neuroscience Discipline ». *Journal of the American Medical Association (JAMA)*, vol. 294, n° 17, p. 2221-2224.
- Kandel, E. 2002 [1998]. « Un nouveau cadre conceptuel de travail pour la psychiatrie. (A new intellectual framework for psychiatry [1998]) ». *L'évolution psychiatrique*, vol. 67 p. 12-40.
- Katz, B. et R. Miledi. 1967. « The Timing of Calcium Action during Neuromuscular Transmission ». *Journal of Physiology*, vol. 189, p. 535-544.

- Kemeny, J. et P. Oppenheim. 1956. « On Reduction ». *Philosophical Studies*, vol. 7, p. 6-19.
- Kendler, K. S. 2001. « A Psychiatric Dialogue on the Mind-Body Problem ». *American Journal of Psychiatry*, vol. 158, p. 989-1000.
- Kendler, K. S. 2005. « Toward a Philosophical Structure for Psychiatry ». *American Journal of Psychiatry*, vol. 162, p. 433-440.
- Kendler, K. S. 2008a. « Explanatory Models for Psychiatric Illness ». *American Journal of Psychiatry*, vol. 165, p. 695-702.
- Kendler, K. S. 2008b. « Introduction. Why does Psychiatry Need Philosophy? ». In *Philosophical issues in psychiatry: Natural kinds, mental taxonomy and causation*, K. Kendler et J. Parnas (dir.), p. 1-16. Baltimore : Johns Hopkins University Press.
- Kendler, K. S. 2008c. « Introduction ». In *Philosophical issues in psychiatry: Natural kinds, mental taxonomy and causation*, K. Kendler et J. Parnas (dir.), p. 132-136. Baltimore : Johns Hopkins University Press.
- Kendler, K. S. 2012. « Levels of Explanation in Psychiatric and Substance Use Disorders: Implications for the Development of an Etiologically Based Nosology ». *Molecular Psychiatry*, vol. 17, p. 11-21.
- Kendler, K. S., R.C. Kessler, M.C. Neale, A.C. Heath et L. Eaves. 1993. « The Prediction of Major Depression in Women: Towards and Integrated Etiologic Model ». *American Journal of Psychiatry*, vol. 150, p. 1139-1148.
- Kendler, K. S. et C.A. Prescott. 2006. *Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric And Substance Use Disorders*. New York : Guilford Press.
- Kendler, K. S. et J. Parnas (dir.). 2008. *Philosophical issues in Psychiatry: Explanation, Phenomenology, and Nosology*. Baltimore, MD : Johns Hopkins University Press.

- Kendler, K.S. et J. Campbell. 2009. « Interventionist Causal Models in Psychiatry: Repositioning the Mind-Body Problem ». *Psychological Medicine*, vol. 39, p. 881-887.
- Kendler, K.S., Zachar, P. et C. Craver. 2011. « What Kinds of Things are Psychiatric Disorders? ». *Psychological Medicine*, vol. 41, p. 1143-1150.
- Kim, J. 2000. « L'émergence, les modèles de réduction et le mental ». *Philosophiques*, vol. 27, n° 1, printemps 2000, p. 11-26.
- Kim J. 2006 [1998]. *L'esprit dans un monde physique. Essai sur le problème corps-esprit et la causalité mentale*. Paris : Éditions Syllepse (collection Matériologiques).
- Kirsch, I. 2009. *The Emperor's New Drugs: Exploding the Antidepressant Myth*. Londre: Bodley Head.
- Kirmayer, L. et I. Gold. 2012. « Re-Socializing Psychiatry: Critical Neuroscience and the Limits of Reductionism ». In *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*, S. Choudhury et J. Slaby (dir.), p. 307-331. Oxford : Blackwell.
- Kistler, M. 2007. « La réduction, l'émergence, l'unité de la science et les niveaux de réalité ». *Matière Première* 2, p. 67-97.
- Kistler, M. 2009. « Explication mécaniste et causalité descendante ». In *L'unité des sciences. Nouvelles perspectives*, T. Martin (dir.), p. 69-89. Paris : Vuibert.
- Kistler, M. 2011. « La causalité ». In *Précis de philosophie des sciences*, A. Barberousse, A. Bonnay, D. et M. Cozic (dir.), p. 100-140, Paris : Vuibert.
- Kitcher, P. 1984. « 1953 and All That: A Tale of Two Sciences ». *Philosophical Review*, vol. 93, p. 335-373.
- Klein, C. 2009. « Reduction Without Reductionism: A Defence of Nagel on Connectability ». *Philosophical Quarterly*, vol. 59, p. 39-53.

- Koch, C. 2012. « Movies in the Cortical Theater ». *Scientific American Mind*, vol. 22, n° 6, p. 20-21.
- Kogan, J. H., Frankland, P. W., Blendy, J. A., Coblenz, J., Marowitz, Z., Schütz, G., et coll. (1997). « Spaced Training Induces Normal Long-Term Memory in CREB Mutant Mice ». *Current Biology*, vol. 7, p. 1-11.
- Kyburg, H. 1965, « Comment. ». *Philosophy of Science*, vol. 32, p. 147-51.
- Le cerveau à tous les niveaux. 2012a. « Les cellules gliales »
URL : http://lecerveau.mcgill.ca/flash/d/d_01/d_01_cl/d_01_cl_ana/d_01_cl_ana.html
Consulté le 8 juillet 2012.
- Le cerveau à tous les niveaux. 2012b. « Les neurotransmetteurs de l'anxiété »
URL : http://lecerveau.mcgill.ca/flash/d/d_04/d_04_m/d_04_m_peu/d_04_m_peu.html
Consulté le 6 septembre 2012.
- Levy, D. L. 2007. « Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders ». *The American Journal of Psychiatry*, vol. 164, n° 11, p. 1763-1764.
- Lewis, D. 1986. « Causal Explanation », in *Philosophical Papers*, vol. II, Oxford : Oxford University Press, p. 214-241.
- Lilienfeld, S.O. 2007. « The Assault on Scientific Mental Health. Ethical Implications », *Science and Ethics. Can Science Help Us Make Wise Moral Judgments?*, P. Kurtz et D. Koepsell (dir.), p. 208-224. Amherst, New York, Prometheus Books.
- Machamer, P., Darden, L., et C. Craver (2000). « Thinking about Mechanisms ». *Philosophy of Science*, vol. 67, p. 1-25.
- Maher, P. 2010. « Lecture 29 : Woodward on Manipulation and Causation » (notes de cours).
URL : <http://patrick.maher1.net/318/lectures/wood1.pdf>
Consulté le 18 avril 2012.

- Malaterre, C. 2011. « Making Sense of Downward Causation in Manipulationism. With Illustrations from Cancer Research ». [manuscrit]
 URL : <http://philsci-archive.pitt.edu/8856/>
 Consulté le 27 décembre 2012
- Mayes, R. G. 2005. « Theories of Explanation ». In *Internet Encyclopedia of Philosophy* (éditions 2010).
 URL : <http://www.iep.utm.edu/explanat/>
 Consulté le 23 juin 2011.
- McNally, R. J. 2011. *What is Mental Illness?* Cambridge, Massachusetts, London et England : The Belnap Press of Harvard Press.
- Mitchell, S. D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge : Cambridge University Press.
- Mitchell, S. D. 2008a. « Explaining Complex Behavior ». In *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*, K. Kendler et J. Parnas (dir.), p. 19-38. Baltimore : Johns Hopkins University Press.
- Mitchell, S. D. 2008b. « Comment : Taming Causal Complexity ». In *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*, K. Kendler et J. Parnas (dir.), p. 125-131. Baltimore : Johns Hopkins University Press.
- Mitchell, S. D. 2009. *Unsimple Truths: Science, Complexity and Policy*. Chicago et Londres : University of Chicago Press.
- Moulines U. C. 2006. *La Philosophie des sciences. L'invention d'une discipline (fin XIXe - début XXIe siècle)*. Paris : Éditions Rue d'Ulm.
- Murphy, D. 2006. *Psychiatry in the Scientific Image*. Cambridge : MIT Press.
- Murphy, D. 2008. « Levels of Explanation in Psychiatry ». In *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*, K. Kendler et J. Parnas (dir.), p. 99-125. Baltimore : Johns Hopkins University Press.

- Murphy, D. 2009. « Psychiatry and the Concept of Disease as Pathology ». In *Psychiatry as Cognitive Neuroscience : Philosophical Perspectives*, M. Broome et L. Bortolotti (dir.), p. 103-117. Oxford University Press.
- Murphy, D. 2010a. « Philosophy of Psychiatry ». In *The Stanford Encyclopedia of Philosophy* (Édition automne 2010), Edward N. Zalta (dir.).
URL : <http://plato.stanford.edu/archives/fall2010/entries/psychiatry/>.
Consulté le 20 septembre 2010.
- Murphy, D. 2010b. « Explanation in Psychiatry ». *Philosophy Compass*, p. 602-610.
- Nagel, E. 1949. « The Meaning of Reduction in the Natural Sciences ». In *Science and Civilization* R.C. Stouffer (dir.), p. 99-135. Madison : University of Wisconsin Press.
- Nagel, E. 1961. *The Structure of Science. Problems in the Logic of Explanation*. New York : Harcourt, Brace & World, Inc.
- Nagel, E. 1970. « Issues in the Logic of Reductive Explanations ». In *Mind, Science, and History*, H.E. Kiefer and M.K. Munitz (dir.), p. 117-137. Albany : State University of New York Press.
- Ney, A. 2008. « Reductionism ». In *Internet Encyclopedia of Philosophy* (éditions 2010).
URL : <http://www.iep.utm.edu/red-ism/>
Consulté le 19 novembre 2011.
- Nichols, S., 2002. « How Psychopaths Threaten Moral Rationalism, or is it Irrational to Be Amoral? ». *The Monist*, vol. 85, p. 285-304.
- Noë, A. 2009. *Out of Our Heads. Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. New York : Hill and Wang.
- Opie, J. 2010. « Levels and Explanations ». In *ASCS09 : Proceedings of the 9th Conference of the Australasian Society for Cognitive Science*, W. Christensen, E. Schier et J. Sutton (dir.), p. 270-276. Sydney : Macquarie Centre for Cognitive Science.

- Oulis, P. 2010. « Nature and Main Kinds of Psychopathological Mechanisms ». *Dialogues in Philosophy, Mental and Neuro Sciences*, vol. 3, n° 2, p. 27-34.
- Price, J., Sloman, L., Gardner Jr, R., Gilbert, P. et P. Rohde. 1997. « The Social Competition Hypothesis of Depression ». In *The Maladapted Mind : Classic Readings in Evolutionary Psychopathology*, S. Baron Cohen (dir.), p. 241-254. Hove: Psychology Press.
- Poirier, P. 2000. « L'empire contre-attaque : le retour de la réduction psychophysique ». *Philosophiques*, vol. 27, n° 1 (printemps), p. 39-62.
- Poland, J. 2006. « Étapes vers un monde sans schizophrénie ». *Philosophiques*, vol. 33, n° 1 (printemps), p. 99-124.
- Popper, K. 1991 [1979]. *La connaissance objective*. Paris : Champs/Flammarion.
- Psillos, S. 2002. *Causation and Explanation*. Montréal et Kingston : McGill-Queen's University Press.
- Radden, J. (dir). 2004. *The Philosophy of Psychiatry: a Companion*. Oxford. Oxford University Press.
- Reutlinger, A. 2012. « Getting Rid of Interventions ». *Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 43, p. 787-795.
- Riel, R. 2011. « Nagelian Reduction Beyond the Nagel Model ». *Philosophy of Science*, vol. 78, n° 3, p. 353-375.
- Ruphy, S. 2005. « Why Metaphysical Abstinence Should Prevail in the Debate on Reductionism ». *International Studies in the Philosophy of Science*, vol. 19, p. 105-121.
- Salmon, W. C. 1971. *Statistical Explanation and Statistical Relevance*. London : University of Pittsburgh Press.
- Salmon, W. C. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton : Princeton University Press.

- Salmon, W. C. 1989. « Four Decades of Scientific Explanation ». In *Minnesota Studies in the Philosophy of Science: Scientific Explanation* (volume 13), W. Salmon et P. Kitcher (dir.), p. 3-219. Minneapolis : University of Minnesota Press.
- Salmon, W. C. 1994. « Causality Without Counterfactuals ». *Philosophy of Science*, vol.61, n° 2, p. 297-312.
- Salmon, W. C. 1999. « Scientific Explanation ». In *Introduction to The Philosophy of Science*, M. H. Salmon et coll. (dir.), p. 7-41. Hackett Publishing Company.
- Sarkar, S. 1992. « Models of reduction and Categories of Reductionism ». *Synthese* vol. 91, n° 3, p. 167-194.
- Sarkar, S. 1998. *Genetics and Reductionism*. Cambridge : Cambridge University Press.
- Sarkar, S. 2008. « Reduction ». In *The Routledge Companion to Philosophy of Science (Routledge Philosophy Companions)*, S. Psillos et M. Curd (dir.), p. 425-434. New York : Routledge.
- Schaffner, K. F. 1967. « Approaches to Reduction ». *Philosophy of Science*, vol. 34, n° 2, p. 137-147.
- Schaffner, K. F. 1969a. « Correspondence Rules ». *Philosophy of Science*, vol. 36, n° 3, p. 280-290.
- Schaffner, K. F. 1969b. « The Watson-Crick Model and Reductionism ». *The British Journal for the Philosophy of Science*, vol. 20, n° 4, p. 325-348.
- Schaffner, K. F. 1974. « Reductionism in Biology: Prospects and Problems ». In *Proceedings of the 1974 Meeting of the Philosophy of Science Association*, R.S. Cohen et A. Michalos (dir.), p. 613-632. Dordrecht: D. Reidel Publishing Company.
- Schaffner, K. F. 1993. *Discovery and Explanation in Biology and Medicine*. Chicago : University of Chicago Press.

- Schaffner, K. F. 1994. «Reductionistic Approaches to Schizophrenia». In *Philosophical Perspectives on Psychiatric Diagnostic Classification*, J. Z. Sadler, O. P. Wiggins et M. A. Schwartz (dir.), p. 279-294. Baltimore : Johns Hopkins University Press.
- Schaffner, K. F. 2008a «Etiological Models in Psychiatry: Reductive and Nonreductive». In *Philosophical Issues in Psychiatry: Natural kinds, mental taxonomy and causation*, K. Kendler et J. Parnas (dir.), p. 48-90. Baltimore : Johns Hopkins University Press.
- Schaffner, K. F. 2008b. «Theories, Models, and Equations in Biology: The Heuristic Search for Emergent Simplifications in Neurobiology». *Philosophy of Science*, vol.75, n° 5, p. 1008-1021.
- Schaffner, K. F. 2010. «The Person and Philosophy of Science and Medicine». *International Journal of Integrated Care*, vol. 10, p. 44-47.
- Seung, S. 2012. *Connectome: How the Brain's Wiring Makes Us Who We Are*. New York: Houghton Mifflin Harcourt.
- Sklar, L. 1967. «Types of Inter-Theoretic Reduction». *British Journal for the Philosophy of Science*, vol. 18, n° 2, p. 109-124.
- Sklar, L. 1999. «The Reduction(?) of Thermodynamics to Statistical Mechanics». *Philosophical Studies*, vol. 95, p. 187-202.
- Solms, M. 2006. «Freud Returns». *Scientific American Mind*, vol. 17, no 2, p. 28-34.
- Stahl, S. M. 2008. «Personalized Medicine, Pharmacogenomics, and the Practice of Psychiatry: On the Threshold of Predictive Therapeutics». *Psychopharmacology*, vol. 13, n° 2, p.115-118.
- Suckling, J. 2012. «Evidencia sobre la depresión con técnicas de imagen: ¿ hay biología en la bibliografía? » (Imaging Evidence for Depression: Is there Biology in the Bibliography?). *Rev Psiquiatr Salud Ment (Barc.)*, vol. 5, n° 1, p. 5-7.

- Suppe, F. 1974. *The Structure of Scientific Theories*. Chicago : University of Illinois Press.
- Suppe, F. 1989. *The Semantic View of Theories and Scientific Realism*. Champaign IL: University of Illinois Press.
- Suppes, P. 1957. *Introduction to Logic*. Princeton : Van Nostrand.
- Suppes, P. 1960. « A Comparison of the Meaning and Use of Models in Mathematics and the Empirical Sciences ». *Synthese*, vol. 12, p. 287-301.
- Suppes, P. 1962. « Models of Data ». In *Logic, Methodology and Philosophy of Science : Proceedings of the 1960 International Congress*, E. Nagel, P. Suppes et A. Tarski (dir.), p. 252-261. Palo Alto CA : Stanford University Press.
- Suppes, P. 1967. « What is a Scientific Theory? » In *Philosophy of Science Today*, Morgenbesser, S. (dir.), p. 55-67. New York : Basic Books.
- Suppes, P. 2002. *Representation and Invariance of Scientific Structures*. Stanford : CSLI Publications.
- Szasz, T. 2007. « Secular Humanism and «Scientific Psychiatry» ». In *Science and Ethics. Can Science Help Us Make Wise Moral Judgments?*, P. Kurtz et D. Koepsell (dir.), p. 195-204. Amherst, New York, Prometheus Books.
- Thagard, P. 1998. « Explaining Disease: Correlations, Causes, and Mechanism ». *Mind and Machines*, vol 8, p. 61-78.
- Thagard, P. 2006. *Hot Thought: Mechanisms and Applications of Emotional Cognition*. Cambridge, Massachusetts : MIT Press.
- Thagard, P. 2008. « Mental Illness From the Perspective of Theoretical Neuroscience ». *Perspectives in Biology and Medicine*, vol. 51, p. 335-352.

- Thagard, P. et S. Findlay. 2012. « Conceptual Change in Medicine : Explanations of Mental Illness from Demons to Epigenetics ». In *The Cognitive Science of Science. Explanation, Discovery, and Conceptual Change*, p. 261-280. Cambridge, Massachusetts : The MIT Press.
- Vaillancourt, J-P. 2012. « La dépression, une “vraie” maladie? », *Le Devoir*, 15 novembre 2012, A8.
- Van Fraassen, B. C. 1980. *The Scientific Image*. Oxford : Oxford University Press.
- Van Fraassen, B. C. 1987. « The Semantic Approach to Scientific Theories ». In *The Process of Science*, Nersessian, N. J. (dir.), p. 105-124. Martinus Nijhoff Publishers.
- Vorms, M. 2011. *Qu'est-ce qu'une théorie scientifique?* Paris : Vuibert.
- Wakefield, J. C. 1992. « Disorder as Harmful Dysfunction: A Conceptual Critique of DSM-III-R's Definition of Mental Disorder ». *Psychological Review*, vol. 99, p. 232-247.
- Wakefield, J. C. 2006. « Fait et valeur dans le concept de trouble mental : le trouble en tant que dysfonction préjudiciable ». *Philosophique*, vol. 33, n° 1 (Printemps), p. 37-63.
- Whitehead, A. et Russel, B. 1910-1913. *Principia Mathematica*. (3 volumes). Cambridge : Cambridge University Press.
- Wikipédia. 2011. « Gaz parfait ».
 URL : http://fr.wikipedia.org/wiki/Gaz_parfait
 Consulté le 22 août 2011.
- Wikipédia. 2012a. « Dépolarisation ».
 URL : <http://fr.wikipedia.org/wiki/D%C3%A9polarisation>
 Consulté le 8 juillet 2012.

- Wikipédia. 2012b. « Phosphorylation ». URL : <http://fr.wikipedia.org/wiki/Phosphorylation>
Consulté le 7 septembre 2012.
- Wikipédia. 2012c. « Maladie de Huntington ». URL: http://fr.wikipedia.org/wiki/Maladie_de_Huntington
Consulté le 6 décembre 2012.
- Wilkinson R. G. et K. E. Pickett. 2009. *The Spirit Level*. London : Penguin.
- Winther, R. G. 2009. « Schaffner's Model of Theory Reduction: Critique and Reconstruction ». *Philosophy of Science*, vol. 76, p.119-142.
- Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford : Oxford University Press.
- Woodward, J. 2008a. « Cause and Explanation in Psychiatry: an Interventionist Perspective ». In *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*, Kenneth Kendler et Josef Parnas (dir.), p. 132-184. Baltimore : Johns Hopkins University Press.
- Woodward, J. 2008b. « Causation and Manipulability ». In *The Stanford Encyclopedia of Philosophy* (Édition hiver 2008), Edward N. Zalta (dir.). URL°: <http://plato.stanford.edu/archives/win2008/entries/causation-mani/>.
Consulté le 6 décembre 2012.
- Woodward, J. 2009. « Scientific Explanation ». In *The Stanford Encyclopedia of Philosophy* (Édition automne 2010), Edward N. Zalta (dir.). URL : <http://plato.stanford.edu/archives/spr2010/entries/scientific-explanation/>.
Consulté le 20 juin 2011.
- Yablo, S. 1992. « Mental Causation ». *Philosophical Review*, vol.101, p. 245-280.
- Zachar, P. 2012. « Comments : Validity, Utility, and Reality : Explicating Schaffner's Pragmatism ». In *Philosophical Issues in Psychiatry II: Nosology*, Kenneth Kendler et Josef Parnas (dir.), p. 190-196. Oxford : Oxford University Press.