

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ÉNACTIVISME ET CONSCIENCE, TROP DE BRUIT POUR RIEN ?

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN PHILOSOPHIE

PAR

ALEXANDRE B. ROMANO

JUIN 2014

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.»

AVANT-PROPOS ET REMERCIEMENTS

Ce travail est le fruit de nombreux détours. Étant indécis, à la fin de mes études de premier cycle, entre l'histoire de la phénoménologie et la philosophie de la psychologie, ce mémoire devait être un prétexte pour satisfaire mes divers intérêts. Il était ainsi initialement prévu d'évaluer la faisabilité du programme de recherche visant à la « naturalisation de la phénoménologie », autant du point de vue historique : *comment concilier la phénoménologie naturalisée à l'antinaturalisme notoire de Husserl ?* que du point de vue des sciences cognitives : *est-ce que la tradition phénoménologique peut faire aujourd'hui une contribution philosophique substantielle quant au problème de la conscience ?* Finalement, aucune de ces deux questions ne seront abordées dans ces pages.

Le lecteur trouvera ici un travail moins ambitieux que ne le prévoyait le projet initial. Dans ce mémoire, je vise à mettre en contexte et à évaluer les arguments en faveur de la théorie éactive de l'expérience perceptuelle. Cette réorientation n'est pas seulement due à des contraintes de temps et d'espace ; elle marque aussi l'évolution de mon intérêt pour un ensemble de problèmes philosophiques touchant la perception et la notion de représentation mentale. Maintenant que ce travail est achevé, il ne me reste qu'à espérer que ce changement de cap n'affectera pas trop la qualité ou la cohérence générale de l'ensemble. En effet, l'écriture de ce mémoire ayant été fait en moins de trois mois, plusieurs arguments présentés ici sont restées à l'état embryonnaire et auraient pu être développées davantage.

Bien que cet ouvrage soit de modeste envergure, la rédaction de celui-ci n'eut été possible sans le support et les encouragements de plusieurs personnes, à qui il me faut maintenant exprimer toute ma gratitude.

Je désire d'abord et avant tout remercier Marina. Tu as choisi d'être ma compagne pendant ces moments difficiles. Sans ta patience et tes encouragements constants, de même que tes commentaires sur plusieurs extraits de ce texte, il est douteux que j'eusse aujourd'hui terminé cet ouvrage. Pour ces raisons, et bien d'autres encore, je te dédie ce mémoire. Maintenant que cet épisode (pénible !) est derrière nous, j'espère avoir la chance de poursuivre mon chemin à tes côtés pendant bien longtemps encore.

Mes remerciements et mon admiration vont ensuite à mes deux directeurs de recherche : Pierre Poirier et Denis Fisette, dont les nombreuses discussions et les avis éclairés m'auront permis de développer les aptitudes nécessaires à ma rédaction. Ici, je dois exprimer une gratitude toute particulière à Pierre, dont l'aide et la compréhension m'auront permis de consacrer autant de temps que nécessaire à la recherche. C'est lui, aussi, qui aura eu à subir les inconvénients de mon manque d'assiduité, alors que certains problèmes d'ordre personnel m'éloignaient des enjeux de nature plus académique. Ainsi, que ce soit dans les fastidieuses démarches visant à l'obtention de financement, que dans ses commentaires sur des versions précédentes de ce mémoire, sa présence constante aura marqué, et marque encore, l'essentiel de mon parcours académique. De nombreux amis méritent aussi de recevoir ma gratitude. Je pense ici notamment à Louis-Philippe Auger, Alexandre Duval et Simon Brien, dont les discussions auront donné lieu à d'intéressantes réflexions, qu'elles se retrouvent, ou non, dans ce mémoire.

Enfin, une partie importante du travail de recherche ayant donné lieu à ce mémoire n'a été possible que grâce au financement du CRSH (2011–2012) et de l'UQAM (2012–2013), pour lesquels j'aimerais exprimer toute ma gratitude. Avec les assauts constants du capitalisme néo-libéral sur nos institutions universitaires, je ne peux que me considérer chanceux d'avoir eu droit à ses ressources, de même qu'à la confiance des professeurs qui ont choisi de bien vouloir me les accorder.

TABLES DES MATIÈRES

AVANT-PROPOS ET REMERCIEMENTS	II
RESUME	VI
INTRODUCTION	1
CHAPITRE I	11
1.1 L'HETEROGENEITE DU TERME « CONSCIENCE »	12
1.1.1 Conscience de créature et conscience d'état	12
1.1.2 Conscience phénoménale et conscience d'accès	13
1.2 LE PROBLEME DE LA CONSCIENCE	17
CHAPITRE II	22
2.1 LE MODELE TRADITIONNEL ET LA NOTION DE REPRESENTATION MENTALE	23
2.2 BREF APERÇUS DES THEORIES CONTEMPORAINES DE LA CONSCIENCE	29
2.3 QUELQUES CRITIQUES ADRESSEES AU MODELE TRADITIONNEL	32
CHAPITRE III	35
3.1 INTRODUCTION ET CONTEXTE	36
3.2 DE L'EXTERNALISME PSYCHOLOGIQUE À L'EXTERNALISME ÉNACTIF	41
3.2.1 L'EXTERNALISME PSYCHOLOGIQUE	45
3.2.2 L'EXTERNALISME ACTIF	49
3.2.3 EXTERNALISME ÉNACTIF	59
3.3 L'EXTERNALISME ET L'EXPLICATION EN SCIENCES COGNITIVES ET EN NEUROSCIENCES	61
3.3.1 NIVEAU PERSONNEL ET NIVEAU SUBPERSONNEL	63
3.3.2 L'EXPLICATION MÉCANISTE	68
3.3.3 EXPLICATIONS MECANISTE, REDUCTIONNISME ET NIVEAU D'EXPLICATION	74
3.4. CONCLUSION PARTIELLE : ENACTION ET EXPLICATION MECANISTE	77

CHAPITRE IV	79
4.1. LES APPROCHES CLASSIQUES DANS LES SCIENCES DE LA VISION ET DE LA CONSCIENCE	86
4.1.1 L'INTERNALISME ET LA SCIENCE DE LA VISION	87
4.1.2 INTERNALISME ET CONSCIENCE : LA RECHERCHE DES CORRELATS NEURONAUX	92
4.2 L'APPROCHE ENACTIVE, UN TOUR D'HORIZON	97
4.2.1 LES CONTINGENCES SENSORIMOTRICES	101
4.2.2 LES CONNAISSANCES OU HABILITÉS SENSORIMOTRICES	104
4.2.3 DEPENDANCE ET CONSTITUTION	105
4.2.4 L'ARGUMENT DÉVELOPPEMENTAL	107
4.2.5. DEUX TYPES DE CÉCITÉS	108
4.3 L'APPROCHE ÉNACTIVE ET LE CARACTÈRE PHÉNOMÉNAL DE LA PERCEPTION	112
4.3.1 LA CONCEPTION BIDIMENSIONNELLE DE LA PERCEPTION	112
4.3.2 LA PRÉSENCE DANS L'ABSENCE	115
4.3.3. LE RÔLE DES ATTENTES DANS L'EXPÉRIENCE PERCEPTUELLE	117
4.4 ANTIREPRESENTATIONNALISME : REJET DU <i>BRIDGE LOCUS</i>	119
4.5 LES MODALITÉS SENSORIELLES ET L'ARGUMENT DE LA PLASTICITÉ	122
4.6 L'APPROCHE ENACTIVE ET LA CONSCIENCE	129
4.7 CONCLUSION	132
CONCLUSION	134
BIBLIOGRAPHIE	138

RÉSUMÉ

Ce mémoire porte sur l'approche éactive (O'Regan & Noë 2001a, Noë 2004), qui est une théorie de l'expérience perceptuelle qui a connu une attention grandissante en sciences cognitives au cours des dernières années. L'idée principale de cette approche est que la conscience n'est pas produite par l'*occurrence* de quelque chose dans le cerveau ou dans l'esprit; autrement dit, elle n'est ni le fait d'une certaine activité neuronale, ni le fait de la relation entre certains états mentaux. Pour ces raisons, l'approche éactive se trouve en opposition frontale avec la majorité des approches en neurosciences, en sciences cognitives et en philosophie, qui souscrivent majoritairement à une forme d'*internalisme*. D'après la théorie éactive, la conscience est *constituée* par l'*activité* de l'organisme dans son environnement de sorte qu'elle ne peut être expliquée adéquatement en s'en tenant seulement à ce qui se passe dans la boîte crânienne de l'agent.

Les deux thèmes centraux de cette théorie sont l'*externalisme* du contenu perceptuel ainsi qu'une forme d'*antireprésentationnalisme*. Par *externalisme*, on entend ici l'idée que les processus cognitifs donnant lieu au contenu perceptuel sont situés partiellement à l'extérieur de l'organisme; alors que par *antireprésentationnalisme* on veut dire que la perception, notamment la perception visuelle, ne nécessite pas l'élaboration d'une représentation détaillée de ce dont nous faisons l'expérience (autrement dit, il n'y a pas de représentation interne qui soit *suffisante* pour donner lieu à notre expérience visuelle). Les détails de notre expérience ne sont donc pas dans notre cerveau, ils sont plutôt dans notre environnement. Les éactivistes justifient leur externalisme par leur antireprésentationnalisme : c'est parce que notre expérience n'est pas explicable en termes de représentations internes qu'il faut « se tourner vers l'extérieur ». Les deux derniers chapitres de ce mémoire portent sur chacun de ces deux aspects.

Bien que l'approche éactive soit une théorie originale qui mérite d'être discutée davantage, celle-ci fait face à de nombreuses limitations. Bien que l'externalisme qu'elle adopte ne soit pas incompatible avec la pratique scientifique, il n'est pas certain que cette posture théorique permette d'expliquer la conscience *en tant que telle*, comme le prétendent O'Regan et Noë (2001a). De plus, si on concède que l'approche éactive permet d'expliquer l'expérience *visuelle*, il n'est pas certain qu'elle puisse expliquer l'expérience d'autres modalités sensorielles, et il est encore plus douteux qu'elle puisse rendre compte de la douleur, un cas paradigmatique d'expérience consciente.

Mots clés : conscience, philosophie de l'esprit, sciences cognitives, théorie éactive, perception, expérience visuelle.

INTRODUCTION

Longtemps un sujet proscrit en science et en philosophie, la conscience est aujourd'hui un sujet à la mode. En effet, l'intérêt pour cet aspect essentiel de notre vie mentale a atteint un point tel, qu'il serait probablement approprié d'affirmer qu'il y eut davantage de pages publiées sur la conscience ces dernières années que sur n'importe quel autre sujet en philosophie de l'esprit. Qui plus est, cette effervescence ne se limite plus aux philosophes : on observe, effectivement, un intérêt grandissant de la part de la communauté scientifique pour « le problème de la conscience ». Cet intérêt est motivé non seulement par les progrès récents en neuroimagerie dans la recherche des corrélats neuronaux de la conscience, mais aussi par la place centrale que semble devoir occuper la conscience dans l'architecture de l'esprit¹ et par la curiosité suscitée par certains phénomènes comme la déconnexion interhémisphérique [*split brain*] (Gazzaniga 1988) et la vision aveugle [*blindsight*] (Weiskrantz 1986). Cette attitude tranche donc avec le dédain manifesté par les scientifiques jusqu'à tout récemment à l'encontre de la conscience, phénomène qui fût longtemps perçu comme le « dernier

¹ À titre d'exemple, on peut penser à la *Global Workspace Theory* (GWT) défendue initialement par Bernard Baars (1988), et qui est aujourd'hui un des modèles de la conscience les plus populaires en sciences cognitives. Ce modèle part de la prémisse, généralement acceptée, voulant que plusieurs processus cognitifs fonctionnent en parallèle de manière inconsciente. Des processus mentaux (perceptions, souvenirs, pensées, etc.) deviennent conscients lorsqu'ils sont recrutés dans la *Global Workspace* par un processus attentionnel. La *Global Workspace* est une mémoire de travail qui rend disponibles l'information contenue dans ces processus mentaux aux autres systèmes psychologiques (ceux impliqués dans la cognition ou le comportement, par exemple). Dehaene et Naccache (2001) identifient trois classes de computations qui ne semblent pas pouvoir être effectués inconsciemment : « la mise à jour [*maintenance*] d'information de manière durable et explicite, de nouvelles combinaisons d'opérations et le comportement intentionnel. » (p. 9). Certes, la GWT est un modèle *psychologique* de la conscience, mais elle a aussi été *importée* au sein des neurosciences (e.g., Dehaene *et al.* 2006). Pour une recension comparative des différents modèles de la conscience en sciences cognitives voir l'excellent compte rendu fait par McGovern et Baars 2007.

bastion des propriétés occultes, des épiphénomènes et des états subjectifs non mesurables » (Dennett 1978).

Cela étant dit, la conscience n'est pas le seul sujet à faire couler beaucoup d'encre en philosophie de l'esprit et en sciences cognitives. Les théories souscrivant à une conception incarnée de l'esprit [*embodied mind*] de même que les approches d'inspiration phénoménologique sont aussi des sujets à la mode. Leur popularité est d'autant plus forte, que ceux qui s'identifient à ces approches sont habitués à faire des déclarations à l'emporte-pièce à propos de la conscience et de la cognition, de même que sur la nature de l'explication en sciences cognitives (dans ce mémoire, j'utiliserai « cognition » et « esprit » de manière interchangeable). De sorte que l'intérêt pour ces approches est, entre autres, dû à leur prétention à être radicalement différentes des approches dites « traditionnelles », « classiques » ou « cartésiennes », actuellement dominantes en sciences cognitives. Dans cette perspective, ces approches ont été particulièrement critiques de l'hypothèse, très influente en sciences cognitives, selon laquelle la meilleure manière de comprendre la cognition est en termes de « structures représentationnelles dans l'esprit et de procédures computationnelles qui opèrent sur ces structures » (Thagard 2012, c.f. le deuxième chapitre).

Ces positions sont aujourd'hui reconnaissables sous la forme d'une multitude de programmes de recherche distincts, parmi lesquels on retrouve l'*approche éactive*, qui trouve son origine dans le projet philosophique de Francisco Varela. Biologiste de formation devenu philosophe, les principales idées caractérisant sa philosophie trouvent leur première formulation cohérente en sciences cognitives dans son livre publié en collaboration avec Evan Thompson et Eleanor Rosch en 1991 : *The Embodied Mind*². La conception de la cognition et de l'explication en sciences

² À vrai dire, plusieurs des idées de Varela étaient déjà formulées dans *Autopoiesis and Cognition* (Maturana & Varela, 1972 en espagnol, 1980 en anglais), mais cet ouvrage est aujourd'hui rarement mentionné dans la littérature en sciences cognitives.

cognitives que l'on retrouve dans cet ouvrage proviennent non seulement de la tradition philosophique husserlienne, mais aussi d'une multitude de traditions différentes, entre autres : la théorie de l'autopoïèse en biologie, les théories dynamiques de la cognition, la cybernétique et la psychologie écologique. Ces auteurs concevaient aussi *The Embodied Mind* comme s'inscrivant dans la « continuation » du travail initié par Merleau-Ponty, leur idée principale étant que les sciences de l'esprit et la tradition phénoménologique peuvent s'enrichir réciproquement (p. xv). Partant du fait qu'une « théorie satisfaisante de la cognition doit tenir compte de la phénoménalité », et donc, du point de vue à la première personne, Varela, Thompson et Rosch aspiraient à développer une nouvelle manière de concevoir et d'étudier la conscience : l'énactivisme³.

Steward, Gapenne et Paolo (2010, p. xiii) soulignent dans leur récent ouvrage que c'est justement l'idée selon laquelle les sciences cognitives doivent tenir compte de l'expérience telle qu'elle est vécue à la première personne qui distingue l'approche énative d'autres approches similaires, comme la psychologie écologique de Gibson (1979). Francisco Varela (1996) concevait d'ailleurs sa « neurophénoménologie »

³ Le verbe anglais « to enact » veut dire à la fois *jouer ou performer quelque chose sur scène*, comme dans le cas d'une pièce de théâtre, ou encore *décréter une loi*. Varela a proposé le nom « approche énative » lorsqu'il commença à rédiger *The Embodied Mind* en 1986 avec Evan Thompson pour faire référence à une multitude d'idées, dont sa conception incarnée, anti-cartésienne et antireprésentationnelle de l'esprit. Par là, il espérait mettre l'accent non seulement sur l'action (d'où l'approche *enactive*), mais aussi mettre l'accent sur le fait que les organismes sont créateurs de significations ou de valeurs qui ne leur sont pas préexistantes. On se souviendra qu'avant d'utiliser cette expression, il utilisait plutôt l'appellation « d'approche herméneutique », qui avait aussi été utilisée par d'autres auteurs en faveur de l'approche incarnée (c.f. Thompson, Lutz, et Cosmelli, p. 42 note 6). Au cours des dernières années, l'étiquette « approche énative » a été principalement revendiquée par Evan Thompson (2007) et ses collaborateurs en lien avec la conception autopoïétique de la cognition; la position d'Alva Noë et de Susan Hurley étant davantage reconnaissable sous l'étiquette « approche sensorimotrice ». Mentionnons cependant, qu'étrangement, il n'est pas question d'autopoïèse dans *The Embodied Mind*. Enfin, une troisième approche a récemment été défendue par Daniel Hutto et Erik Myin dans *Radicalizing Enactivism* (2012, voir aussi Menary 2006b pour une présentation de cette approche ainsi que certaines critiques). Sans souscrire à l'énactivisme-autopoïétique de Thompson, Hutto et Myin vont plus loin que Noë et O'Regan en rejetant toute notion de contenu représentationnel.

comme un « remède méthodologique » au problème de la conscience qu'il jugeait indissociable de l'approche énative.

Cependant, déterminer comment cette « exigence phénoménologique » doit être interprétée ou mise en oeuvre, n'est pas chose aisée. D'une part, certains auteurs ont adhéré au programme de recherche visant à *naturaliser la phénoménologie*, c'est-à-dire d'intégrer aux sciences cognitives certaines *thèses* ou *méthodes* propres à la tradition husserlienne (e.g. Petitot *et al.* 1999). Ces chercheurs doivent cependant justifier ce « tournant naturaliste » de la phénoménologie, qui semble aller à l'encontre de l'antinaturalisme notoire de Husserl (1913/1989). D'autre part, plusieurs chercheurs, ont tenté d'adopter une approche phénoménologique, à savoir, une approche à la première personne, dans le cadre du problème de la conscience, sans pour autant « naturaliser la phénoménologie ».

Ce débat tourne principalement autour de la place et du statut épistémique que nous devrions accorder aux rapports à la première personne dans l'étude de la conscience et de la cognition en général (e.g. Gallagher 2002, 2003, Gallagher & Overgaard 2005, Jack et Roepstorff 2002, 2004, Lutz 2002, Lutz *et al.* 2002, Velmans 2007a, b). L'enjeu de ce débat est cependant parfois difficile à saisir, dans la mesure où personne, sinon quelqu'un adoptant une forme dogmatique de béhaviorisme, ne serait prêt à nier l'utilisation de rapports à la première personne dans l'étude de la conscience. En effet, ceux-ci sont essentiels, que ça soit pour décrire adéquatement l'*explanandum* d'une théorie de la conscience ou encore dans le cadre des protocoles expérimentaux mis en place par les scientifiques lorsqu'ils doivent déterminer si un sujet est conscient d'un stimulus. Toutes ces questions ont donné lieu à une longue controverse entre les défenseurs de la neurophénoménologie et Daniel Dennett (c.f. le numéro spécial de *Phenomenology and Cognitive Science*, 2007, vol. 6 sur l'hétérophénoménologie édité sur le sujet).

L'approche énaactive n'est cependant pas seulement une critique de la conception traditionnelle de la cognition et des sciences cognitives, c'est aussi, et surtout, un modèle original de ce qu'est la cognition et de ce que devraient être les sciences cognitives. À cet effet, un aspect important de ce programme de recherche est de faire comprendre en quoi des processus cognitifs comme la perception visuelle ou la conscience ne sont pas des événements ou des états *internes* localisés dans le cerveau des organismes. Ceux-ci sont plutôt des propriétés des organismes en tant que tels, et reposent sur l'engagement de l'organisme avec son environnement, engagement qui prend la forme d'interactions sensorimotrices en temps réel. Les énaactivistes ont ainsi rejeté deux distinctions traditionnelles qui caractérisent notre compréhension de ce qu'est l'esprit. D'une part, les énaactivistes rejettent la séparation arbitraire entre perception, cognition et action ; d'autre part, les énaactivistes rejettent aussi la distinction entre l'organisme et son environnement. Les processus cognitifs sont ainsi conçus comme étant le fait d'un système étendu et complexe qui intègre ces différents aspects, de sorte que, selon les énaactivistes, nous ne pouvons comprendre les phénomènes cognitifs réels indépendamment les uns des autres.

Cela étant dit, on retrouve aujourd'hui dans la littérature deux conceptions de l'énaactivisme⁴, la première ayant une portée plus large que la seconde :

L'énaactivisme–autopoïétique : Cette conception large de l'approche énaactive porte sur la nature de la *cognition en général* et tente de faire voir la profonde continuité entre la vie et l'esprit. L'idée centrale de l'énaactivisme–autopoïétique est que « la cognition est fondé [*grounded*] dans l'activité de création de sens [*sense-making*] d'agents autonomes — des êtres qui se génèrent et se maintiennent eux-mêmes, et par le fait même, *énaactent* ou mettent de l'avant, leur propre domaine de signification et de valeur » (Thompson & Stapleton 2008, p. 23). Cette conception de l'énaactivisme

⁴ Pour une distinction similaire entre ces deux conceptions de l'approche énaactive, voir Torrance 2005.

s'inscrit directement dans la lignée des travaux de Varela et son principal promoteur aujourd'hui est Evan Thompson (2007).

L'énactivisme–sensorimoteur : L'énactivisme au sens étroit est une théorie portant sur le statut métaphysique du *contenu perceptuel*. La notion de *contenu* est un construit théorique central en philosophie de l'esprit de même qu'en sciences cognitives qui sera introduit plus en détails dans le deuxième chapitre. Dans le cadre de la discussion qui nous intéresse, la notion de contenu permet d'expliquer les deux propriétés essentielles des phénomènes mentaux, à savoir l'intentionnalité et leur caractère phénoménal. En effet, comme le dit si bien David Rosenthal :

Tous les états mentaux, quelle qu'en soit la sorte, manifestent des propriétés qui appartiennent à l'une des deux catégories suivantes : soit les propriétés *intentionnelles*, soit les propriétés *phénoménales* ou sensibles. [...] Quoi qu'on puisse encore affirmer de vrai des états mentaux, il n'en demeure pas moins qu'il est évident qu'on ne considérerait pas du tout un état comme mental s'il n'avait pas une propriété intentionnelle ou phénoménale. (Rosenthal 2002, p. 168).

Nous dirons ainsi d'un état mental, comme une croyance, un désir ou une perception qu'il est doté d'intentionnalité ou d'un caractère phénoménal, c'est-à-dire :

1. l'intentionnalité : un état mental, comme une perception, est intentionnel ou représentationnel dans la mesure où il est *à propos de quelque chose*. À cette fin, on utilisera ce terme pour désigner à la fois *l'information qu'il nous transmet à propos du monde* ainsi que la *manière* dont cette information nous est présentée.

2. le caractère phénoménal : par caractère phénoménal, on désigne *l'aspect sensoriel* de l'expérience, ou encore « ce que cela fait » pour un organisme, que d'encourir cette expérience.

C'est deux notions seront traitées plus en détails dans les deux prochains chapitres. Retournons, pour l'instant, à l'énactivisme–sensorimoteur, aussi connu sous le nom d'*approche sensorimotrice* de la perception, et qui a été défendu principalement par Susan Hurley, Alva Noë, Kevin O'Regan et Eryk Myin. Cette interprétation restreinte de l'énactivisme–sensorimoteur est sujette à controverses, d'autant plus que ces auteurs ne sont pas toujours clairs sur les ambitions et la portée de leur théorie. Les propos de Susan Hurley (1998) encouragent à concevoir l'énactivisme–sensorimoteur comme une théorie de la cognition en général, alors qu'Alva Noë et Kevin O'Regan la présentent à la fois comme une théorie de la conscience en général, une théorie de la perception, une théorie de la vision et une théorie de l'expérience visuelle.

La **thèse centrale** de l'énactivisme–sensorimoteur est la suivante : le contenu perceptuel, qui détermine le caractère phénoménal de notre *expérience perceptuelle*, est *constitué* par l'exercice des connaissances des contingences sensorimotrices.

Les conséquences de cette thèse, si elle s'avère, sont nombreuses pour le statut de l'expérience perceptuelle et pour le développement futur de la recherche en sciences cognitives et en neurosciences. D'abord, parce que, tel que je viens de le mentionner, selon l'approche énaactive, la conscience est une propriété d'un système complexe, composé de l'organisme en tant que tel et de son activité dans son environnement. La conscience n'est donc ni une propriété attribuée à certains états mentaux, ni due à l'occurrence d'une activité neuronale particulière, tel que le présuppose le programme de recherche visant à identifier les corrélats neuronaux de la conscience. À en croire les énaactivistes, ce programme de recherche est d'ailleurs voué à l'échec.

Mais l'approche énaactive nous invite aussi à repenser notre manière d'aborder plusieurs phénomènes étudiés par la science de la vision. En effet, dans leur article phare publié dans *Behavioral and Brain Sciences* en 2001, Kevin O'Regan et Alva

Noë prétendent que l'énactivisme est une « approche unifiée » qui nous permet de mieux comprendre une multitude de phénomènes perceptuels qui paraissent, à première vue, indépendants les uns des autres : la complétion visuelle, la cécité au changement, la stabilité perceptuelle, la perception prosthétique, la perception des couleurs, l'adaptation aux lentilles inversées, la restauration de la vision après une opération chirurgicale et l'agnosie visuelle (p. 971).

Malgré leur différences, ces deux conceptions de l'énactivisme partagent au moins un objectif commun : trouver une explication satisfaisante au problème de la conscience (Thompson 2004), et ce, même si ces dernières années l'approche énaïve a été étendue à d'autres phénomènes, comme la cognition sociale, l'empathie et l'intersubjectivité (voir par exemple les textes dans Steward, Gapenne et Paolo 2010). Cependant, le remède énaïf au problème de la conscience est partiellement conceptuel : pour les énaïvistes, nous ne résoudrons pas ce problème en en apprenant davantage sur le fonctionnement du cerveau, mais plutôt en concevant la conscience autrement, à savoir comme étant une activité de l'organisme en tant que tel.

Trame narrative de ce mémoire

Ce mémoire constitue un ensemble de réflexions sur ce que je considère être les deux aspects centraux de l'énactivisme : l'externalisme du contenu perceptuel et l'antireprésentationnalisme. Dans cette perspective, mon analyse sera consacrée principalement à l'énactivisme-sensorimoteur⁵ défendu par Alva Noë et Kevin O'Regan, même si, à bien des égards, cette analyse s'applique aussi à la théorie d'Evan Thompson, qui adopte, lui aussi, un externalisme et un

⁵ Pour alléger le texte, je dirai, désormais « énaïvisme » pour désigner l'énactivisme-sensorimoteur.

antireprésentationnalisme similaires (c.f. Noë & Thompson 2004, Thompson 2007 p. 253 et suiv.).

Ce choix découle de ce que je pense être à la fois l'aspect original et l'aspect le plus important de l'approche énactive : une thèse métaphysique sur la nature du contenu perceptuel, notamment sur son caractère phénoménal. On peut considérer, en effet, que la conception sensorimotrice de la vision avait déjà été mise de l'avant par certains chercheurs en intelligence artificielle et en robotique au début des années quatre-vingt-dix (e.g. Ballard 1991, Brooks 1991a, b). Daniel Dennett avait aussi proposé une idée similaire quant aux détails de notre expérience perceptuelle (e.g. 1991, p. 354 et suiv.). En ce sens, je montre que l'approche énactive se démarque non pas par sa conception incarnée, active ou sensorimotrice de la vision ou de la cognition, mais par sa position externaliste et antireprésentationnaliste dans le cadre du problème de la conscience : l'énactivisme nous invite, en effet, à repenser totalement notre manière de concevoir ce qu'est la conscience et comment nous pouvons espérer l'expliquer. Leur idée étant, je le répète, que lorsque nous cessons de « regarder en dedans » et tournons notre regard « vers l'extérieur », le mystère de la conscience s'en trouve soudainement dissolu.

À cette fin, ce mémoire est constitué de quatre chapitres. Les deux premiers chapitres visent à introduire les notions essentielles pour comprendre les enjeux dont il sera question dans ce mémoire, alors que les deux derniers, plus longs, sont plus argumentatifs et concernent les deux traits principaux de l'approche énactive : son externalisme et son antireprésentationnalisme.

Le **premier chapitre** porte sur la conscience. Le terme de conscience est hétérogène et renvoie à divers concepts. Il en va de même de ce que l'on entend par « problème de la conscience », puisque celle-ci donne lieu à une multitude de problèmes empiriques et conceptuels de nature différente.

Le **deuxième chapitre**, quant à lui, consistera en un aperçu rapide de ce que l'on appelle le modèle « traditionnel » de la cognition. La notion de représentation mentale sera aussi introduite, puisqu'elle viendra jouer un rôle important au sein de mon exposé.

Le **troisième chapitre** porte sur ce que j'appelle l'externalisme actif, thèse celle selon laquelle l'étude de certains processus cognitifs (ici, la conscience) demande que l'on étudie un système complexe constitué de l'agent et de son environnement. Plus précisément, je me concentre sur les conséquences de l'approche énative pour l'explication scientifique, autant en sciences cognitives qu'en neurosciences. Dans ce chapitre, je présente d'abord les différents types d'externalismes, avant d'évaluer si l'externalisme des théories incarnées est compatible avec l'explication mécaniste, qui est le modèle d'explication dominant en biologie, en neurosciences et en sciences cognitives.

Le **quatrième chapitre**, quant à lui, porte plus spécifiquement sur l'antireprésentationnalisme de même que sur la conception sensorimotrice de la perception. Les énativistes défendent cette position en faisant appel à plusieurs expériences en psychologie, notamment aux cas de cécité au changement ou à la plasticité sensorielle. Ces expériences seront présentées et j'évaluerai dans quelle mesure elles viennent effectivement soutenir la thèse énative. Une fois cela fait, je préciserai dans quelle mesure l'approche énative peut être qualifiée d'antireprésentationnaliste.

La **thèse** que je soutiens dans ce mémoire double : d'une part, contrairement à ce que suggère Prinz (2012, voir aussi Jack & Prinz 2004), l'approche énative n'est pas incompatible avec les sciences cognitives et les neurosciences ; d'autre part, contrairement à ce que soutiennent O'Regan et Noë (2001a) l'approche énative ne

peut pas être considérée, pour l'instant, comme une théorie de la conscience. En effet, alors que l'approche éactive contribue à comprendre pourquoi nous sommes conscient d'une *tache rouge* plutôt que d'une *tache verte*, ou conscient *visuellement* plutôt qu'*auditivement*, elle ne nous permet pas de comprendre pourquoi certains états mentaux sont *dotés d'un caractère phénoménal* plutôt que le contraire. En ce sens, j'espère dans les pages qui suivent non seulement remettre en contexte la théorie éactive de la perception, mais aussi tracer les limites de son application.

CHAPITRE I

« LE PROBLÈME » DE LA CONSCIENCE

Dans leur introduction à *The Embodied Mind*, Varela, Thompson et Rosch déploraient le fait que les sciences cognitives n'aient pratiquement rien à dire à propos de notre expérience subjective, c'est-à-dire de la conscience (1991, p. xv). À bien des égards, cette critique formulée au début des années quatre-vingt-dix n'est plus valide aujourd'hui. En témoigne, en effet, le nombre d'ouvrages parus sur ce sujet au fil des trente dernières années⁶.

Cependant, malgré un nombre grandissant de publications traitant de la conscience, le terme « conscience » reste l'objet de plusieurs confusions conceptuelles. Une pratique, malheureusement trop courante, consiste à introduire une notion complexe, comme celle de conscience, par une notion tout aussi complexe, e.g. celle de subjectivité, de qualia, de phénoménalité, ou encore, par l'expression « ce que cela fait » [*what is it like*]. Afin d'éviter les confusions qui caractérisent souvent les discussions sur la conscience, je présenterai dans la prochaine section un bref aperçu des différents concepts qui sont désignés par le terme « conscience » dans la littérature. Dans la seconde section, je présenterai certains des problèmes empiriques, épistémologiques et conceptuels qui touchent ce phénomène.

⁶ Voir entre autres : Baars 1988, Dennett 1991, Edelman 1992, Searle 1992, Dretske 1995, Tye 1995, Crick 1995, Lycan 1987, 1996, Chalmers 1996, Siewert 1998, Carruthers 2001, Rosenthal 2005.

1.1 L'hétérogénéité du terme « conscience »

Le terme de conscience est **hétérogène**, c'est-à-dire que le terme conscience recouvre un ensemble de concepts distincts et possiblement indépendants. Cette hétérogénéité est d'ailleurs amplifiée par la diversité des disciplines qui s'y intéressent, chacune d'entre elles adoptant parfois sa propre terminologie. La phrase bien connue de Thomas Nagel (1974) « ce que cela fait » [*what is it like*] est aujourd'hui largement utilisée dans la littérature pour faire référence à la conscience, et ce, non sans causer quelques maux de tête. En effet, par cette phrase, Nagel affirme que « fondamentalement un organisme a des états mentaux conscients si cela lui fait un certain effet *d'être* cet organisme — un certain effet *pour* l'organisme. » (1974/1987 p. 392), mais il est cependant douteux que cette formulation clarifie, réfère ou explique quoi que ce soit. Pour reprendre les mots de Lycan, cette expression serait maintenant « inutile », « pernicieuse », voire même « dangereuse » dans la mesure où elle est désormais utilisée uniquement pour renvoyer à tout ce que l'on considère comme étant *mystérieux* ou *inexplicable* dans la conscience (1996, p. 77).

Une étape incontournable pour un philosophe qui s'intéresse à la conscience consiste à faire quelques distinctions dans la manière d'utiliser le terme « conscience ». Ces clarifications conceptuelles sont d'autant plus nécessaires que le terme « conscience » ne renvoie pas à une notion théorique aux limites bien définies, mais est tiré du langage ordinaire. Ces clarifications seront l'objet de cette section. Je vais d'abord introduire la distinction entre conscience de créature et conscience d'état, avant de me tourner vers la distinction entre conscience phénoménale et conscience d'accès.

1.1.1 Conscience de créature et conscience d'état

Lorsqu'on dit d'une créature qu'elle est consciente, on peut vouloir dire plusieurs choses différentes. Elle peut d'abord être dotée de **sentience**, c'est-à-dire avoir la

disposition à avoir des sensations et à réagir au monde; elle peut ensuite être **éveillée**, c'est-à-dire être en train d'exercer actuellement cette disposition : la créature⁷ est donc alerte, par opposition à endormie ou dans le coma. Une créature peut aussi être dite consciente si elle a une forme, plus ou moins élaborée, de **conscience de soi**. Pris dans un sens fort, les animaux et les enfants en bas âge ne sont peut-être pas capables d'une conscience de soi. Communément, on considère aussi une créature comme consciente lorsqu'elle est dotée d'une certaine **forme de subjectivité** : le monde a une apparence particulière du point de vue de la créature. C'est ce que Thomas Nagel entendait par *what is it like*. Enfin, une créature peut être dite consciente lorsqu'elle a des **états mentaux conscients** (Van Gulick 2011).

Malgré le fait que les enjeux philosophiques et scientifiques touchant la conscience de créature soient dignes d'intérêt en eux-mêmes, la majorité des philosophes se sont plutôt intéressés aux propriétés phénoménales caractérisant certains de nos états mentaux⁸. Cela n'est pas surprenant, dans la mesure où, en philosophie, l'intérêt pour la conscience provient essentiellement du problème corps-esprit, c'est-à-dire la relation entre événements mentaux et événements physiques.

1.1.2 Conscience phénoménale et conscience d'accès

Lorsqu'on se tourne vers la conscience propre aux états mentaux, on retrouve une seconde distinction très importante dans la littérature : celle entre *conscience*

⁷ Il me faut ici mentionner le fait que la conscience de créature ne doit pas nécessairement être comprise comme étant une propriété *binnaire* : celle-ci viendrait plutôt en *degrés*. Les travaux de Petra Stoerig (c.f. *inter alia* 2007) en neurosciences, par exemple, suggèrent que les organismes sont sujets à plusieurs niveaux de conscience différents et qu'ils sont donc rarement complètement inconscients.

⁸ Rosenthal, par exemple, reconnaissant que les questions concernant la conscience de créature et la conscience d'état ne sont peut-être pas indépendantes, affirme cependant que : « Knowing that a creature's being conscious means that it is awake and sentient does not help us understand what it is for a mental state to be conscious, or place useful constraints on what such state consciousness consists in. » (Rosenthal 1997, p. 730).

phénoménale et *conscience d'accès*. Introduite d'abord par Ned Block (1995), celle-ci a ensuite été discuté par de nombreux auteurs :

Par **conscience phénoménale (conscience-p)**, on fait généralement référence au *caractère qualitatif* distinctif de nos expériences mentales conscientes, dont les cas paradigmatiques sont la douleur et les autres sensations (e.g. les couleurs). Cette catégorie peut aussi être étendue, comme le fait Block, pour y inclure non seulement les expériences perceptuelles et les sensations corporelles, mais aussi les émotions et les affects ou sentiments. Dans cet article, d'ailleurs, Block présente cette notion de la façon suivante :

[...] we have P-conscious states when we see, hear, smell, taste, and have pains. P-conscious properties include the experiential properties of sensations, feelings, and perceptions, but I would also include thoughts, desires, and emotions. A feature of P-consciousness that is often missed is that differences in intentional content often make a P-conscious difference. What it is like to hear a sound as coming from the left differs from what it is like to hear a sound as coming from the right. P-consciousness is often representational [...] So far, I don't take myself to have said anything terribly controversial. The controversial part is that I take P-conscious properties to be distinct from any cognitive, intentional, or functional property. (Block 1995, p. 230).

Pour référer au caractère qualitatif, la notion de **qualia** a aussi souvent été employée, non sans controverses. Ainsi plusieurs « vertus » ont été attribuées aux qualia, dont le fait d'être des propriétés intrinsèques, ineffables, privées et monadiques de notre expérience. Comme on peut le voir dans cette citation, Block (1995, 1997) a défendu que notre expérience est dotée de propriétés intrinsèques qui ne peuvent être réduites à des propriétés représentationnelles, puisqu'il est concevable que certains états mentaux aient le même contenu représentationnel tout en ayant un caractère phénoménal différent. En ce sens, les qualia ou les propriétés phénoménales sont les « composantes » de notre expérience (la rougeur d'une expérience visuelle, etc.) et

s'apparentent aux qualités secondes de Locke ou aux sense-data des empiristes⁹ : « les qualia et les sense-data jouent un rôle similaire : les sense-data sont les objets de [notre] expérience et les qualia [sont] leurs propriétés. » (Crane 2000, p. 171).

D'un autre côté, Harman (1990), Dennett (1991) et Tye (1992) ont argué que les qualia n'existaient tout simplement pas. Il est difficile de comprendre comment ces auteurs peuvent avoir des positions aussi diamétralement opposées sur la question de l'existence des qualia, si par « qualia » on ne fait que faire référence au fait que certains états mentaux ont un caractère qualitatif. Pour cette raison, j'utiliserai ce terme technique de la manière la plus neutre possible afin de faire référence aux qualités de notre expérience, par exemple l'expérience de l'amertume lorsque je bois un café.

Enfin, déterminer quels types d'états mentaux sont dotés de qualia (seulement nos états perceptuels, ou bien aussi nos états cognitifs comme nos croyances ?), ainsi que la nature de la relation entre les qualia et le monde extérieur sont aussi des questions qui suscitent d'intenses discussions que je ne pourrai aborder ici.

Certains auteurs ont cependant aussi fait valoir que la conscience phénoménale ne se borne pas à l'aspect *qualitatif* de notre expérience. Uriah Kriegel (2009), par exemple, a proposé la distinction entre le **caractère subjectif** et le **caractère qualitatif**, qui seraient tous deux des composantes de la conscience phénoménale : « [n]ous saisissons ce qu'est le caractère subjectif en fixant ce qui est commun à tous les états phénoménalement conscients, et saisissons ce qu'est le caractère qualitatif en fixant ce qui varie entre eux. » (p. 1, c.f. aussi Thompson 2007 chap. 7, Zahavi 2002, 2003). Kriegel justifie l'introduction de la notion de caractère subjectif en faisant valoir que

⁹ Pour une discussion détaillée à propos de la relation historique entre la notion de *sense-data* et celle de *qualia*, c.f. Crane (2000).

nos expériences conscientes ne sont pas simplement *en nous*, mais sont aussi *pour nous*, ce qu'il désigne d'ailleurs par l'expression de *mienneté* [*for-me-ness*].

Dans une perspective similaire, Van Gulick a fait remarquer que « la structure phénoménale de la conscience recouvre une grande partie de l'organisation spatiale, temporelle et conceptuelle de notre expérience du monde, et de nous-mêmes comme étant des agents s'y trouvant », ce qui dépasse le cadre strict de la notion de qualia telle qu'employée dans la littérature (2011). Pour cette raison, il distingue la conscience qualitative, qui est purement sensorielle et la conscience phénoménale, qui inclut la *structure intentionnelle* caractéristique de notre expérience consciente. Dorénavant, lorsque je parlerai de « caractère phénoménal », je ferai référence à cette conscience phénoménale entendue au sens plus large.

La **conscience d'accès (conscience-a)** quant à elle est la *conscience d'intelligence*, autrement dit, un état mental est conscient-a lorsqu'il est accessible à nos pensées (i.e. à interagir avec nos autres états mentaux) et accessible à l'expression verbale. Par exemple, je suis conscient-a que le département de philosophie est au cinquième étage si je suis capable de rapporter ma croyance lorsque la situation s'y prête, où d'utiliser cette connaissance pour me rendre à mes séminaires. En ce sens, la conscience-a n'est pas une propriété des états mentaux en tant que tel, mais dépend plutôt de la relation entre l'état mental en question et sa *disponibilités* aux mécanismes responsables de guider le comportement intelligent de l'organisme. Ainsi, alors que la conscience-p est conçue comme étant une propriété occurrente et intrinsèque des états mentaux, la conscience-a est une propriété dispositionnelle et relationnelle. Enfin, mentionnons que Ned Block défend la thèse selon laquelle nous pouvons être conscients-a sans être conscients-p, et inversement, mais plusieurs philosophes refusent cette thèse (Cohen & Dennett 2011, Prinz 2012 sont des exemples récents).

1.2 Le problème de la conscience

Le terme « conscience »¹⁰ n'est pas le seul à être hétérogène : ce que l'on entend par « problème de la conscience » l'est tout autant.

À cet effet, Van Gulick (2011) identifie trois catégories de questions :

1. les questions descriptives, par exemple : *qu'est-ce que la conscience et quelles sont ses caractéristiques principales ? Quelles sont les méthodes adéquates pour l'étudier, l'expliquer et la modéliser ? si l'étude de la conscience nécessite une approche à la première personne, quel est son statut épistémologique ?*
2. les questions explicatives, par exemple : *est-ce que la conscience est un aspect primitif de la réalité, comme les entités postulées par la physique, et si ce n'est pas le cas, comment est-ce que le cerveau occasionne la conscience ?*
3. les questions fonctionnelles, par exemple : *quel est le rôle de la conscience dans l'économie du mental ? a-t-elle un impact causal ?*

En un certain sens, les problèmes suscités par la conscience sont une partie importante du problème corps-esprit qui a traversé l'histoire de la philosophie depuis l'avènement de la modernité. En effet, dans les deux cas, la question en arrière-plan reste celle de la place des phénomènes mentaux dans le monde physique. Le fait que la conscience soulève autant de questions différentes est aussi un signe que les

¹⁰ Pour une bonne introduction au problème de la conscience, c.f. Carruthers 2000 chap. 3, dont cette section est partiellement inspirée.

philosophes ne s'entendent pas entre eux sur la nature du problème. Pour certains, le problème est *causal* et porte sur l'interaction entre le mental et le physique, alors que pour d'autre le problème est *explicatif* et vise à déterminer quel type d'explication permettra de rendre compte du mental d'une manière qui soit compatible avec les sciences contemporaines (Crane & Patterson 2000, p. 1). Or, bien que ces questions soient interreliées, je soutiens ici que les questions explicatives constituent le coeur du « mystère de la conscience ». Enfin, mentionnons qu'alors que les philosophes se sont principalement intéressés à donner une explication naturaliste de la *conscience phénoménale*, les chercheurs en sciences cognitives et en neurosciences ont concentré la majeure partie de leurs efforts à expliquer la *conscience d'accès*. Dans ce mémoire, lorsqu'il sera question du « problème de la conscience », c'est de la *conscience phénoménale* dont il sera question.

Plusieurs philosophes ont d'ailleurs souligné les enjeux explicatifs relatifs à la conscience. Dans le sillage de Thomas Nagel, Colin McGinn (1989, 1993 chap. 1–2), a soutenu que bien que *nous devons croire que le physicalisme est vrai*, en un certain sens nous ne savons pas *comment* il peut être vrai. Au cours des prochains paragraphes, je vais me concentrer plus particulièrement sur les propos de David Chalmers concernant le problème difficile de la conscience et de Joseph Levine concernant le déficit d'explication, propos qui ont été particulièrement influents.

Dans son célèbre article intitulé *Facing Up to the Problem of Consciousness* (1995), David Chalmers distingue les problèmes faciles **du problème difficile de la conscience**. Les problèmes faciles concernent les divers processus cognitifs souvent associés à la conscience et à son rôle dans l'économie cognitive de l'organisme vivant, à savoir : l'état d'éveil, l'introspection, la disposition à rapporter le contenu de nos états mentaux, la conscience de soi, l'attention, le contrôle volontaire de l'action, la connaissance (« je suis conscient du problème ») (Chalmers 1996, p. 24–5). Ces

concepts, comme ceux d'autres sciences (en chimie, en biologie, en psychologie etc.) sont *fonctionnels*, c'est-à-dire qu'ils sont caractérisés par leur *rôle causal distinctif* plutôt que par le substrat physique où ils sont réalisés ou implémentés et dont nous faisons abstraction lors de l'explication. C'est d'ailleurs souvent ainsi que l'on conçoit les états mentaux, telle que les croyances ou les douleurs : une croyance, par exemple, est une chose qui est causée par une perception ou une inférence, et qui peut causer à son tour des actions après avoir interagi avec des désirs et d'autres croyances. Cette conception fonctionnaliste (Putnam 1967) des états mentaux fait en sorte qu'ils admettent possiblement une forme d'*explication* réductionniste, dans la mesure où nous pouvons identifier et étudier les mécanismes qui en sont responsables.

Par contre, le problème difficile concerne plutôt les concepts phénoménaux (e.g. quand je pense « *Cette douleur est vraiment déplaisante* » lorsque je suis chez le dentiste). Chalmers (1995, 1996) soutient que ces concepts ne sont pas des concepts fonctionnels. Selon Chalmers, c'est ce que démontre l'*argument de la concevabilité des zombies*.

En philosophie de l'esprit, on entend par « zombies » des créatures qui ont été imaginée afin de mettre en évidence certains problèmes quant à la relation entre le caractère phénoménal et le monde physique. Selon l'histoire qu'on en fait habituellement, les zombies seraient des créatures *comme nous en tout point*, à l'exception du fait qu'ils seraient totalement dénués de conscience phénoménale. Ainsi, une copie de moi molécule-par-molécule pourrait être un zombie, malgré le fait qu'elle soit identique à moi autant du point de vue de sa composition physique que de son comportement.

Convaincu par les arguments faisant appel aux zombies, certains auteurs ont affirmés que les concepts d'états phénoménaux sont plutôt des concepts « recognitionnels » [*recognitionnal*] (Carruthers 2000, p. 66 et suiv., voir aussi Loar 1990). Par cette

expression, Carruthers réfère à l'idée toute cartésienne selon laquelle les états mentaux ayant un caractère phénoménal « sont ceux que nous pouvons reconnaître en nous-même, de manière non-inférentielle, directement [*straight off*] », en vertu de la manière dont ils se présentent à nous subjectivement (ibid., p. 14).

Ainsi, pour conclure, les problèmes qualifiés de « faciles », sont désignés ainsi non pas parce que nous pouvons présentement en fournir une explication satisfaisante, mais parce que nous savons que ces problèmes pourront être *éventuellement* expliqués avec les méthodes traditionnelles des sciences cognitives, c'est-à-dire en comprenant le fonctionnement des mécanismes qui leur sont sous-jacents, ce qui n'est pas le cas du « problème difficile ».

Selon Joseph Levine (1983, 1993, 2001 chap. 3), le fait que la conscience phénoménale échappe à une explication en termes fonctionnels ou causaux fait en sorte nous faisons face à un déficit ou à un fossé explicatif [*explanatory gap*] infranchissable¹¹. Plus précisément, Levine fait valoir que peu importe la quantité d'information *objective* que nous pourrions accumuler sur l'activité du cerveau, la structure physique des neurones, ainsi que sur leurs diverses opérations chimiques ou électriques, celle-ci sera toujours insuffisante pour expliquer *pourquoi* tel ou tel phénomène objectif (e.g. un changement physique dans le cerveau) produit ou génère tel ou tel phénomène subjectif (e.g. un sentiment de douleur particulier dans le genou). Il faut distinguer cette position très forte sur la possibilité d'expliquer la conscience, d'une position plus bénigne, selon laquelle il est évident qu'il y a

¹¹ « What seems to be responsible for the explanatory gap, then, is the fact that our concepts of qualitative character do not represent, at least in terms of their psychological contents, causal roles. Reduction is explanatory when by reducing an object or property we reveal the mechanisms by which the causal role constitutive of that object or property is realized. Moreover, this seems to be the only way that a reduction could be explanatory. Thus, to the extent that there is an element in our concept of qualitative character that is not captured by features of its causal role, to that extent it will escape the explanatory net of a physicalistic reduction. » (1993, p. 134).

actuellement un déficit d'explication (après tout, nous ne savons toujours *aujourd'hui* comment l'activité neuronale donne lieu systématiquement à la conscience phénoménale).

L'argument du déficit d'explication, quant à lui, découle de la conception de l'*explication* adoptée par Levine. D'après lui, l'explication doit avoir une forme déductive, c'est-à-dire qu'on doit être capable de déduire la nature de l'*explanandum* à partir d'information sur la nature de l'*explanans*. Lorsque deux propositions sont dans une telle relation déductive, elles sont *dans une relation nécessaire*. Or, selon lui, il n'y a pas de telle relation entre les propositions à propos de l'activité neuronale et celles à propos de notre expérience phénoménale. La conséquence étant que, selon Levine, nous ne pouvons expliquer cette identité psycho-physique et restons ainsi toujours avec un « déficit d'explication » (comme on peut le voir, à bien des égards l'argument de Levine s'apparente au *knowledge argument* de Frank Jackson (1982)).

CHAPITRE II

LE « MODÈLE TRADITIONNEL » EN SCIENCE COGNITIVE

La stratégie généralement utilisée pour présenter le modèle éactif de la cognition est de le mettre en opposition avec le modèle classique, souvent aussi qualifié de « cartésien »¹². Évidemment, par « cartésien », on ne fait pas référence ici au dualisme des substances tel qu'on le retrouve dans la philosophie de Descartes et auquel pratiquement plus personne n'adhère. Cependant, malgré le fait que ce cartésianisme ne consiste pas en un dualisme ontologique, le modèle traditionnel continue à entretenir une distinction entre les phénomènes mentaux et les phénomènes physiques, faisant ainsi des phénomènes mentaux un domaine d'étude distinct de celui des sciences naturelles.

Dans cette perspective, plusieurs philosophes ont repris la distinction entre *théories personnelles* et *théories subpersonnelles*, introduite initialement par Dennett (1969, p. 90-6). Alors que les théories personnelles attribuent des états mentaux à des agents en tant que tels, les théories subpersonnelles s'intéressent aux mécanismes causaux qui permettent aux agents d'entretenir de tels états mentaux, *ces deux types de théories étant deux façons différentes de décrire un même système*. En effet, nous pouvons décrire le comportement d'un système et son rapport à son environnement en adoptant

¹² Pour un aperçu des critiques adressées au modèle traditionnel, voir entre autres : Dreyfus 1972, Varela *et al.* 1991, chap. 7, Clark 1997, Haugeland 1998, Petitot *et al.* 1999, Hurley 1998, Noë 2004 chapt 1-2, Wheeler 2005 chap. 1-4, Thompson 2004, 2007 chap. 1, van Gelder 1995. Enfin, précisons que d'un point de vue purement historique, cette référence au « cartésianisme » n'a pas été introduite par ceux qui critiquent le modèle classique, mais bel et bien par ceux qui y adhèrent. Dans son *Cartesian Linguistics* (1966), par exemple, Chomsky voit d'ailleurs Descartes comme un ancêtre de la grammaire générative.

un vocabulaire *intentionnel*, notre description se situe alors au niveau personnel; par ailleurs, en adoptant une perspective *mécaniste-causale*, nous pouvons aussi décrire ce système en se concentrant sur l'interaction entre les différentes parties de celui-ci (ses organes, etc.), nous avons alors affaire à une description des processus subpersonnels. Cette distinction entre les phénomènes mentaux et les phénomènes physiques est d'autant plus renforcée par « le contraste entre sémantiques et syntaxes, espace des raisons versus espaces des causes, ou encore entre vocabulaire intentionnel et vocabulaire physique » que l'on retrouve abondamment dans la littérature (Haugeland 1998, p. 208) (pour plus de détails, voir le troisième chapitre).

Ce chapitre sera divisé en trois sections. Dans la première, je présente rapidement le modèle traditionnel en insistant particulièrement sur la notion de représentation mentale. Ensuite, dans la section suivante, je présente les théories généralement proposées en philosophie de l'esprit pour expliquer la conscience. Ce bref exposé devrait permettre de comprendre en quoi ces théories de la conscience se conforment au modèle traditionnel. La troisième et dernière section porte sur certaines critiques qui ont été adressées au modèle traditionnel. Celles-ci serviront à mettre la table pour les deux chapitres qui suivront.

2.1 Le modèle traditionnel et la notion de représentation mentale

Ce *cartésianisme matérialiste*, si l'on peut dire, se manifeste aujourd'hui de plusieurs façons dans notre manière d'expliquer l'esprit, et ce, tout particulièrement au sein du représentationnalisme et de la théorie computationnelle de l'esprit, qui sont deux postulats centraux des sciences cognitives. Mais les restes de cartésianisme ne se limitent pas à cela. L'habitude de séparer perception, cognition et action de même que l'idée selon laquelle l'esprit cause l'action sont aussi des legs du cartésianisme donnant lieu à des problèmes qui persistent toujours en sciences cognitives.

Le **représentationnalisme**, aussi appelé *théorie représentationnelle de l'esprit*, est la théorie selon laquelle la cognition peut être expliquée en postulant des représentations mentales; cette théorie trouvant son origine dans la *révolution cognitive* et la fin du béhaviorisme (pour plus de détails, c.f. Greenwood 2009, chap. 12). En refusant de concevoir l'introspection comme une méthode scientifique valide, les béhavioristes évitaient, en règle générale, le langage mentaliste, et donc, de postuler des états mentaux (i.e. des représentations réalisées dans le cerveau). Par contre, avec la naissance des sciences cognitives au milieu des années cinquante, plusieurs chercheurs ont accepté le principe méthodologique selon lequel sans états mentaux, il est impossible d'expliquer le mental. Ce *tournant méthodologique*, si l'on peut dire, implique à son tour une *exigence ontologique* : rendre compte de la nature de ces représentations mentales dans un langage compatible avec les sciences.

La notion de représentation, qui viendra jouer un rôle important dans les chapitres suivant, mérite que nous qu'on s'y attarde quelques instants, c'est ce que je vais faire dans les prochains paragraphes.

Lorsque les philosophes parlent de « représentations mentales », ils font habituellement référence à différents types d'états mentaux (des exemples typiques sont les croyances, les désirs, les intentions, les perceptions, les hallucinations, etc.). Un aspect important de plusieurs de nos états mentaux est le fait qu'*ils sont à propos de quelque chose*, c'est-à-dire qu'*ils sont dotés d'intentionnalité*, et qu'*ils représentent le monde comme étant de telle ou de telle façon* (dans ce mémoire, j'utilise *intentionnel* et *représentationnel* de manière interchangeable). Les états mentaux ne sont évidemment pas les seules choses à être des représentations; les phrases et les images, notamment, sont aussi des représentations. En ce sens, les représentations sont des entités avec des propriétés sémantiques (elles ont possiblement une référence, un contenu, une valeur de vérité, des conditions de satisfaction, etc.).

Le concept de représentation mentale est aussi un construit théorique central en sciences cognitives et un des objectifs principaux de la discipline est d'expliquer comment de simples phénomènes physiques (e.g.. l'activité neuronale) peuvent représenter le monde. Comprendre comment nous pouvons entretenir des états représentationnels et quelle est la nature de ces états est important, dans la mesure où ceux-ci déterminent notre comportement (après tout, nous agissons habituellement conformément à nos croyances et à nos désirs, états qui représentent comment le monde *est* et *devrait être*) On notera aussi qu'historiquement, la plupart des termes mentalistes utilisés en philosophie s'avèrent être des représentations mentales (on pensera ici aux pensées, aux idées, aux impressions, aux concepts, etc.).

À première vue, l'expression « A représente B » peut laisser croire que cette relation représentationnelle est *entre deux choses réelles*, puisqu'on accepte généralement qu'une relation entre deux choses implique que celles-ci *existent toutes deux*. Il n'est pas évident que cela s'applique aux *objets intentionnels*, c'est-à-dire les *objets visés* par nos représentations, car nous pouvons penser des objets inexistantes ou impossibles (le carré rond étant un exemple célèbre d'objet intentionnel impossible).

Ce qui permet aux représentations mentales d'être intentionnelle, c'est le fait que les états mentaux ont un *contenu*. La notion de contenu est notoirement imprécise en philosophie de l'esprit et souvent tenue pour primitive, mais nous ne pouvons nous en tenir à cela, puisque les questions sur lesquelles nous nous pencherons dans les prochains chapitres concernent justement la nature du contenu représentationnel.

Conformément à la majorité des philosophes, nous dirons que c'est en vertu de son contenu représentationnel qu'un état mental, comme une croyance, représente un aspect du monde. En ce sens, le contenu est la manière dont un objet, une propriété ou un état de chose est donné à un sujet dans son expérience ou à sa pensée (Cussins, 2003, p. 133). En effet, lorsque j'ai une perception ou une croyance, ma relation à

l'objet intentionnel se fait *selon un point de vue particulier* (Searle 1992, chap. 7). C'est ce que John Searle a appelé l'*aspectual shape*¹³.

Tel que je l'entends ici, le contenu détermine aussi les conditions de satisfaction de nos états représentationnels. Si je crois, par exemple que *le dernier film de Charlie Kaufman joue à mon cinéma préféré ce soir*, je pense à plusieurs choses (au film, à mon cinéma préféré, etc.), mais je ne pense pas à ces choses de manière complètement aléatoire : je pense à une situation particulière ou à un état de chose possible. On voit ainsi que c'est en vertu de son contenu qu'une croyance est vraie ou fausse (la situation représentée par ma croyance selon laquelle *le dernier film...* peut ou pas s'avérer).

Par ailleurs, la croyance que *le dernier film...* est distincte de la croyance que *le chat est sur la table*. C'est pourquoi on affirme souvent que le contenu est ce qui permet l'*individuation* d'un état mental particulier, c'est-à-dire de déterminer son *identité* et de le distinguer des autres états mentaux. Nos états mentaux peuvent aussi avoir le même objet intentionnel tout en ayant un contenu représentationnel différent, l'exemple de Frege à propos d'*Hesperus* et de *Phosphorus* en est un cas célèbre. Empruntant la formulation heureuse de J. J. Valberg, Tim Crane conçoit le contenu perceptuel comme étant *ce que nous mettrions en mots, si nous avions les mots pour nécessaires pour le faire* (1992, p. 39). Cette manière de voir le contenu sied parfaitement au contenu perceptuel, puisqu'il est souvent pris pour acquis que notre expérience perceptuelle est plus riche et plus détaillée que ne l'est notre répertoire

¹³ Plusieurs termes ont été utilisés dans l'histoire de la philosophie pour désigner le fait que nos états mentaux se font selon une certaine perspective. Comme le souligne Tim Crane (1992, p. 38), l'expression « aspectual shape » est utile car elle n'est pas liée à une théorie particulière de l'intentionnalité, comme peuvent l'être les termes de *Sinn* chez Frege ou de *Noema* chez Husserl. Par ailleurs, rapprocher signification et contenu mental peut donner l'impression que ce dernier a une forme propositionnelle calquée sur l'expression linguistique; il n'est cependant pas clair qu'il en aille ainsi avec le contenu perceptuel.

conceptuel (Evans 1982, p. 229 pour une défense de la position opposée, c.f. McDowell 1994).

Enfin, il y a une différence entre *croire que le dernier film...*, *désirer que le dernier film...*, *imaginer que le dernier film...*, etc.. La différence entre ces différents états mentaux concerne le *mode intentionnel*, qui détermine la relation que nous entretenons avec le contenu : nous pouvons percevoir, croire, désirer quelque chose, etc. L'expression « attitude propositionnelle » est aussi souvent utilisée, mais conformément à John Searle je préfère parler en terme de *mode intentionnel*, afin de ne pas présupposer que tout contenu mental est de nature propositionnelle.

Une autre distinction importante, est celle existant entre le *contenu* représentationnel et son *véhicule* représentationnel. Le véhicule d'un contenu représentationnel est le substrat physique qui *a* ou *exprime* ce contenu, e.g. un état neuronal, un ensemble de sons, un morceau de papier, etc.. Une façon de comprendre cette distinction est par la voie d'un exemple : la représentation *le chat est sur la table* peut être représentée de plusieurs façons différentes, que ça soit par l'entremise d'une phrase ou d'une image. Dans les deux cas, la même situation est représentée, mais par l'entremise d'un véhicule différent.

Ainsi, alors que le *contenu* constitue ce qui est représenté, le *véhicule* est l'ensemble de propriétés intrinsèques qui permettent à une représentation particulière (une occurrence) de remplir ses fonctions représentationnelles. Lorsque nous avons un état mental conscient, nous sommes conscient de ce qu'il représente, mais pas de son véhicule, qui est l'activité neuronale. En sciences cognitives, les chercheurs font appel à plusieurs types de véhicules représentationnels pour rendre compte du contenu de nos états mentaux, que ce soit des *computations* sur des symboles, des réseaux de neurones, des systèmes dynamiques ou des états neuronaux.

J'ai mentionné plus haut que de parler de représentations mentales implique à son tour une *exigence ontologique*, c'est-à-dire expliquer à la fois comment les représentations peuvent jouer un rôle causal et comment des entités physiques peuvent avoir des propriétés sémantiques. Là-dessus, la **théorie computationnelle**, qui est une des hypothèses les plus influentes en sciences cognitives, peut être considérée comme une manière de satisfaire cette exigence. En effet, selon celle-ci, la cognition peut être expliquée en terme de *computations sur des symboles*, c'est-à-dire que l'esprit considéré comme étant une machine qui emmagasine, transforme et manipule des structures porteuses d'information (des représentations) selon des règles syntaxiques, et donc en fonction de leur propriétés causales plutôt que de leur propriétés sémantiques. On comprend donc que la distinction entre véhicules et contenus est motivée par l'espoir de pouvoir rendre compte de manière naturaliste, et donc dans un vocabulaire indépendant de l'esprit ou du langage, du fait que nos représentations mentales entrent dans des relations causales les unes avec les autres (e.g. ma croyance perceptuelle selon laquelle il y a de la bière dans le réfrigérateur peut causer un désir d'être désaltéré).

La théorie computationnelle de l'esprit renforce, par contre, l'idée selon laquelle la cognition peut être expliquée en faisant pratiquement abstraction de ce qui se passe à l'extérieur de l'esprit-cerveau. Traditionnellement, les diverses disciplines au sein des sciences cognitives ont adopté une conception de l'esprit selon lequel celui-ci consiste en un ensemble de processus de traitement de l'information, dont la relation avec le monde extérieur était, d'un point de vue théorique, relativement sans importance. Ainsi, les systèmes perceptuels et moteurs étaient considérés comme accessoires dans la compréhension de la cognition centrale (le langage, le raisonnement, etc.). Ceux-ci étaient conçus, en effet, comme étant simplement des périphériques responsables des *input* et des *output* du système. Par ailleurs, selon le modèle classique, la perception, la cognition et l'action sont aussi conçues comme étant indépendantes l'une de l'autre. Dans cette perspective, la perception est la fonction qui vient en premier,

alors que l'action lui est dérivée ou subordonnée. La cognition, quant à elle, est conçue comme étant le point « central » de l'esprit.

Plusieurs chercheurs en sciences cognitives ont ensuite adopté la position d'après laquelle perception, cognition et action sont décomposables hiérarchiquement en différents *modules*, qui sont des processus spécialisés réalisant des tâches bien précises; la cognition, servant ici d'*intermédiaire* entre la perception et l'action. Autrement dit, dans cette perspective, la perception est responsable d'extraire de l'information de l'environnement et de construire un modèle de celui-ci, qui sera ensuite utilisé par la cognition, celle-ci étant la cause du comportement. La métaphore de Hurley est une bonne illustration de cette manière de voir l'esprit : « perception et action ne sont pas seulement séparées l'une de l'autre, mais aussi séparées des processus cognitifs de haut niveau. L'esprit est une espèce de sandwich, où la cognition en est la garniture » (1998, p. 401).

2.2 Bref aperçus des théories contemporaines de la conscience

Dans le chapitre précédent, nous avons fait un survol des différentes acceptations du terme « conscience » en philosophie, de même que des problèmes philosophiques soulevés par la conscience. Cette section, quant à elle, vise à offrir un (trop) bref sommaire des différentes théories de la conscience en philosophie de l'esprit. Pour des raisons d'espace, je ne pourrai entrer dans les détails de ces théories. Cependant, cette esquisse rapide devrait suffire à comprendre en quoi ces théories de la conscience se conforment au modèle traditionnel de la cognition.

Depuis le regain d'intérêt des philosophes pour la conscience, les débats ont principalement porté sur les tentatives d'expliquer la conscience de manière réductionniste. On dira ici qu'une théorie est « réductionniste » si elle essaie d'identifier, ou d'expliquer, la conscience en termes qui ne la présupposent pas. Par

exemple, certains ont essayé de réduire la conscience à une relation particulière entre certains états mentaux, alors que d'autres ont plutôt tenté de l'identifier à des processus d'information ou encore à des états physiologiques ou neurologiques.

En philosophie, nous pouvons regrouper les diverses approches du problème de la conscience en trois grandes catégories, que je vais présenter tour à tour. D'abord, il y a les théories représentationnalistes, qui visent à réduire le caractère phénoménal de nos états mentaux à des propriétés intentionnelles ou représentationnelles. Ensuite, il y a des approches qui tentent de diminuer l'importance du problème de la conscience : en ce sens, la notion de conscience phénoménale ferait davantage référence à un ensemble de confusions conceptuelles qu'à une réelle entité digne d'étude. Enfin, il y a ceux qui affirment que la conscience ne peut tout simplement pas être expliquée de manière réductionniste, malgré tout nos efforts en ce sens, nous seront toujours pris avec un déficit d'explication.

L'approche la plus répandue en philosophie de l'esprit est sans conteste celle visant à réduire la conscience à une forme ou une autre de représentations mentales, ce que je désigne ici par le terme de « représentationnalisme ». Selon les théories représentationnelles de la conscience (Harman 1990, Dretske 1995, Tye 1995), le caractère phénoménal d'une expérience peut être totalement réduit à ce que l'expérience représente. Ainsi, selon le représentationnalisme, l'amertume propre à mon expérience du café n'est pas une entité mentale particulière ayant la propriété de l'amertume, mais est plutôt une *représentation* de la propriété objective qu'est l'amertume.

La majorité de ceux qui adhèrent à ces théories le font parce qu'ils pensent que la notion de représentation est plus facile à expliquer de manière physicaliste (c'est-à-dire en utilisant des termes qui ne présupposent aucune propriétés sémantiques) que ne peut l'être la notion de conscience phénoménale.

Le représentationnalisme n'est cependant pas sans critiques, plusieurs philosophes ont d'ailleurs jugé que la conscience implique nécessairement une forme de méta-représentation d'ordre supérieur (d'où, *higher order theories*). Contrairement aux théories représentationnalistes « unilévélites » [*first order*] mentionnées plus haut, les théories d'ordre supérieur soutiennent qu'une représentation est consciente seulement lorsqu'elle est l'objet intentionnel d'un état mental d'ordre supérieur (i.e. une méta-représentation). Ainsi, selon ces théories, la conscience implique nécessairement une forme de conscience de soi.

Spécifier la nature de cet état d'ordre supérieur a par contre donné lieu à de nombreuses controverses. Selon Rosenthal(1990) l'état d'ordre supérieur s'apparenterait à un jugement assertorique, c'est-à-dire à une pensée [*thought*], alors que selon Lycan celui-ci s'apparenterait davantage à la perception (2004).

Dans la deuxième catégorie, Daniel Dennett (1991) est fameux pour avoir tenté de dissoudre le problème de la conscience. Selon lui, la distinction entre conscience phénoménale et conscience d'accès ne tiendrait tout simplement pas la route, notre expérience se limitant totalement à ce que nous pouvons possiblement en dire dans des rapports verbaux (Dennett 2003, Cohen & Dennett 2011). Il n'y aurait par conséquent pas de réelle distinction à faire entre les problèmes faciles et le problème difficile : une fois tous les problèmes faciles expliqués, la conscience sera bel et bien expliquée.

Il faut cependant noter que malgré sa tentative de marginaliser le problème de la conscience, la théorie qu'il propose a les traits d'une théorie d'ordre supérieur (pour un exposé systématique de sa théorie, c.f. Carruthers 2000, chapt. 10 et 11).

Finalement, certains philosophes dont il a déjà été question dans le chapitre précédent (Chalmers 1995, 1996, Levine 1983 et McGinn 1989, 1991, 1993 Nagel 1974) ont défendu la thèse selon laquelle les deux approches précédentes sont vouées à l'échec puisque malgré toutes les tentatives de réduction nous ne pouvons échapper au déficit d'explication.

2.3 Quelques critiques adressées au modèle traditionnel

Maintenant que nous avons présenté le modèle traditionnel de la cognition, de même que les stratégies les plus populaires pour traiter de la conscience en philosophie de l'esprit, nous pouvons passer aux diverses critiques et théories allant dans le sens inverse et qui ont vu le jour au début des années quatre-vingt-dix. Celles-ci, habituellement regroupées dans l'ensemble, très large, des *théories de l'esprit incarné*, dont l'énaclivisme fait partie, souscrivent à l'idée selon laquelle les processus cognitifs nécessitent une interaction entre l'agent *incarné* et son environnement immédiat. Cette idée, certes très vague, regroupe cependant une multitude de thèses différentes qui ne sont pas partagées par tous les chercheurs qui défendent de telles théories. Dans un article remarquablement clair, Margaret Wilson (2002), par exemple, en dénombre six :

1. La cognition est située. La cognition est une activité qui prend place dans le contexte d'un environnement physique et social particulier et qui implique à la fois la perception et l'action. Par « cognition située », on fait ici référence à des processus cognitifs où **a.** l'information perceptuelle parvient de manière continue afin d'informer ces processus cognitifs et où **b.** l'activité motrice transforme l'environnement de manière pertinente pour la réalisation de la tâche en cours (ibid., p. 626). Alors que cela est évident dans le cas de certaines tâches cognitives, comme la conduite automobile, il n'en va pas de même dans le cas de tâches plus « abstraites ».

2. *La cognition se déroule dans des contraintes temporelles.* Pour comprendre la cognition, il faut observer comment elle réagit dans le contexte d'interaction en *temps réel* avec l'environnement. En mettant l'accent sur cet aspect, il est question de faire référence à certains « points faibles des modèles traditionnels en intelligence artificielle, qui peuvent généralement contruire et manipuler des représentations internes d'une situation donnée à loisir » en ne tenant pas compte des contraintes temporelles qui sont pourtant si importantes à la survie d'organismes vivants devant survivre dans la nature (ibid., p. 627). C'est encore à cet aspect temporel auquel on renvoie lorsqu'il est question de *goulot d'étranglement représentationnel*, c'est-à-dire lorsque des situations demandent des réponses rapides et adaptées et « qu'il n'y a tout simplement pas assez de temps pour construire un modèle détaillé [*full blown*] de l'environnement, duquel tirer un plan d'action » (ibid. p. 628).

3. *Les agents déchargent un certain fardeau cognitif dans leur environnement.* Parce que nos ressources cognitives sont limitées, notamment notre mémoire de travail et notre attention, nous exploitons notre environnement de manière à réduire notre fardeau cognitif. Pour ce faire, nous emmagasinons ou manipulons de l'information dans l'environnement, plutôt que « dans notre tête », information que nous récupérons seulement lorsque certaines tâches cognitives le demandent.

4. *L'environnement est une partie du système cognitif.* Le flux d'information entre l'esprit et l'environnement de l'agent est si dense et continu qu'afin d'*étudier* la nature de l'activité cognitive, il faut considérer le système constitué par son esprit, son corps et son environnement.

5. *La cognition est pour l'action.* La fonction de l'esprit est de guider l'action, et les processus cognitifs comme la perception et la mémoire doivent être interprétés dans la perspective de leur contribution au comportement approprié à la situation.

6. *La cognition hors ligne [off-line] est incarnée.* Même lorsque l'agent est découplé de son environnement, certains processus mentaux (e.g. l'imagerie mentale, la mémoire de travail, la mémoire épisodique, le raisonnement) sont toujours fondés dans des mécanismes qui ont évolué pour permettre l'interaction avec l'environnement.

Les thèses selon lesquelles l'esprit est un phénomène *incarné, situé* et *enchâssé* [*embedded*] dans son environnement est cependant triviale, personne ne semblant prêt à affirmer le contraire. Il est, en effet, évident que les causes du comportement, de même que des événements mentaux, sont aussi à chercher en dehors du cerveau de l'agent. Cela dit, la question visant à déterminer si les phénomènes mentaux peuvent être *compris indépendamment* du corps et de l'environnement de l'agent est plus controversée, tout comme celle visant à établir quel devrait être l'objet d'étude des sciences cognitives. Ainsi, ce n'est pas parce que la cognition est le fait d'un système distribué constitué d'un agent, de son corps et de son environnement (thèse ontologique), qu'il faille pour autant en conclure que l'objet d'étude des sciences cognitives doivent être ce même système (thèse épistémologique). C'est justement ce qu'a John Haugeland en tête, lorsqu'il se demande si une conception désincarnée de l'esprit est *intelligible*. Il affirme en effet :

we can grant the holist thesis that mind would be impossible in the absence of body and/or world, and still ask whether it can be *understood* as a distinct and well-defined subsystem within the necessary larger whole. (1998, p. 211, je souligne).

Autrement dit, là où réside encore la controverse est lorsqu'on se demande jusqu'où nous pouvons pousser les thèses incarnées, situées et enchâssées et quelles conséquences métaphysiques nous devons tirer de cet état de chose. Dans les sections précédentes j'ai introduit le modèle traditionnel de la cognition, de même que les théories de la conscience qui ont reçu le plus d'attention en philosophie de l'esprit. J'ai ensuite présenté les grandes lignes des critiques adressées à ce modèle. Partant de ces considérations générales, les deux chapitres qui suivront visent à approfondir plus particulièrement une des théories de l'esprit incarné, l'approche éactive, qui a été sujette à de nombreuses discussions ces dernières années.

CHAPITRE III

L'APPROCHE ÉNACTIVE, UNE MENACE POUR LES NEUROSCIENCES ?

Dans ce chapitre, j'évalue les conséquences pour l'explication en sciences cognitives et neurosciences de l'externalisme éactif, qui est la thèse selon laquelle le *contenu perceptuel*¹⁴ est réalisé par des processus qui dépassent les simples limites du cerveau. Ce questionnement est justifié par le fait que l'externalisme éactif semble non seulement *contre-intuitif*, mais aussi incompatible avec le modèle explicatif de ces sciences et qui consiste essentiellement à décomposer un mécanisme en ses composantes (que l'on considère habituellement comme étant localisées dans le cerveau).

Puisque les discussions entourant l'externalisme renvoient à une longue controverse en philosophie de l'esprit et des sciences cognitives, je commence par faire une taxinomie des différents types d'externalismes. J'abonde ici dans le même sens que Susan Hurley (2010), lorsqu'elle affirme que, même si l'externalisme éactif fait violence à nos intuitions (pré)théoriques sur la relation entre la conscience et le monde physique : nous ne pouvons rejeter cette thèse sans des justifications formulées en bonnes et dues formes, car nous n'avons, à ce jour, aucune idée de la manière dont le cerveau génère la conscience phénoménale (**section 1**). Cela fait, j'évalue ensuite si l'externalisme caractéristique de l'approche éactive est compatible avec le modèle explicatif dominant en biologie, en neuroscience et en sciences cognitives : l'explication mécaniste (**section 2**). L'objectif de ce chapitre est donc de

¹⁴ Comme il a été précisé plus haut, par contenu perceptuel j'entends autant l'aspect représentationnel que le caractère phénoménal de nos états perceptuels.

discuter des enjeux relatifs à l'explication scientifique soulevés par l'approche énative, sans pour autant aborder de front la théorie énative en tant que telle. En ce sens, une part importante des considérations tenues ici sont possiblement applicables aux autres types d'externalismes qui ont vu le jour en sciences cognitives ces dernières années sous la bannière de « l'esprit étendu ».

3.1 Introduction et contexte

Une des idées reçues les plus répandues chez les philosophes et les scientifiques qui s'intéressent à l'esprit est probablement celle voulant que l'esprit soit *réalisé*, d'une manière ou d'une autre, *par des entités et des processus localisés dans le cerveau*. En effet, même les philosophes qui défendent des arguments externalistes à l'égard de l'individuation du contenu représentationnel (l'externalisme psychologique, voir plus bas) tendent à accepter, du moins implicitement, la *thèse de la survenance neuronale* selon laquelle les *processus cognitifs* qui *permettent* [*enable*] aux agents d'entretenir des états mentaux, ont pour substrat des états neuronaux. Le terme « survenance » [*supervenience*] est une notion technique de la métaphysique contemporaine. On dira qu'il y a une relation de survenance entre un ensemble de propriétés d'ordre supérieur A et un ensemble de propriétés de niveau inférieur B lorsque les propriétés A ne peuvent changer sans que les propriétés B changent aussi. Autrement dit, dans le cadre de la question qui nous concerne, nous pouvons dire que le phénomène mental M survient sur le phénomène physique P lorsqu'il « est impossible que deux événements s'accordent sous toutes leurs caractéristiques physiques [...] tout en différant sous leurs caractéristiques psychologiques. » (Davidson 1980, p. 336). Dorénavant, lorsque je dirai qu'une propriété d'ordre supérieur, e.g. un état mental, est *réalisée* par un état neuronal, il sera alors question d'une relation de survenance.

Retournons, maintenant, à l'idée largement partagée selon laquelle le mental survient sur le neuronal. C'est cette idée qui motive ce que Jerry Fodor a appelé le *solipsisme*

méthodologique et que Stephen Stich a appelé le *principe d'autonomie*¹⁵, à savoir que les seules propriétés pertinentes pour l'explication en psychologie ou en neurosciences sont celles qui entraînent une différence dans les propriétés physiques internes de l'organisme.

Cette conception de l'esprit et de l'explication en psychologie découle à son tour d'un dilemme qui est non moins généralement accepté : *ou bien* l'esprit est *réalisé par* l'activité cérébrale, *ou bien* il faut accepter le dualisme des substances de Descartes. La position dualiste cartésienne étant jugée pour le moins inconfortable dans le contexte naturaliste actuel, la majorité des auteurs ont, par conséquent, adopté la thèse selon laquelle l'esprit, et par extension, la conscience, sont générés par le cerveau (le naturalisme, ou le physicalisme, tel que je l'entends ici, étant la thèse d'après laquelle tout ce qui existe est réalisé physiquement et est gouverné par les lois de la physique ; la manière précise de spécifier ces deux notions est, bien sûr, matière à controverses).

D'autre part, ces dernières années, plusieurs philosophes s'identifiant aux thèses de l'esprit incarné ont présenté divers arguments épistémologiques ou ontologiques dans le but de rejeter cette dichotomie. Susan Hurley (2010), notamment, a fait valoir que la dichotomie entre le cerveau et la *res cogitans* n'était pas valide, mais relevait plutôt du faux dilemme. En effet, selon elle, nos concepts mentaux (e.g. esprit, perception, conscience, croyance, etc.) ne réfèrent pas à des propriétés du cerveau, mais désignent

¹⁵ Je laisse ici de côté les critiques formulées par Stich contre le solipsisme méthodologique de Fodor et je m'en tiens à l'idée générale. Stich présente ainsi son principe d'autonomie : « The basic idea of the principle is that the states and processes that ought to be of concern to the psychologist are those that supervene on the current, internal, physical state of the organism. [...] If we respect the autonomy principle, then the fact that a pair of organisms have different histories or that they are in significantly different environments will be irrelevant to a psychological theory except insofar as these differences make a difference to the organism's current, internal, physical state. » (1983, p. 164). Ce principe semble si évident à ses yeux, que dans l'article où il l'avait introduit initialement (1978), Stephen Stich ne s'était pas senti obligé de fournir un argument en faveur de celui-ci : selon lui, ce principe est en fait partie prenante des pratiques des psychologues. Ces derniers l'acceptent presque unanimement comme étant une évidence (1978, p. 714, note 5).

plutôt l'organisme actif en tant que tel en interaction avec son environnement. Par cette affirmation, Hurley est loin de nier que le cerveau soit un élément *nécessaire* pour qu'il y ait esprit, elle ne fait qu'amoindrir son rôle dans l'explication des phénomènes mentaux; autrement dit, selon elle, si l'activité neuronale est nécessaire au mental, elle n'est pas suffisante.

Dans cette perspective, la thèse de l'**externalisme actif**¹⁶ est la thèse selon laquelle moyennant des conditions particulières, certains états mentaux ou processus cognitifs sont réalisés en dehors du cerveau, que ce soit dans le corps ou l'environnement immédiat de l'organisme. Contrairement à l'externalisme psychologique, l'externalisme actif concerne les véhicules de nos représentations mentales (voir le deuxième chapitre). En règle générale, cette thèse part du constat que l'organisme est actif et capable d'agir sur son environnement et d'y recueillir l'information requise pour effectuer les tâches cognitives nécessaires à sa survie; à l'inverse, postuler des états internes afin d'expliquer la cognition perd de l'importance.

Conformément à cette thèse, Susan Hurley et ses collaborateurs ont défendu ce qu'on pourrait appeler l'*externalisme énaactif*, qui est la thèse métaphysique d'après laquelle « l'expérience visuelle n'est pas une occurrence dans l'esprit qui est produite par l'activité neuronale », mais serait plutôt *constituée* par l'activité habile [*skillfull*] de l'organisme dans son environnement (O'Regan & Noë 2001, p. 1018). Autrement dit, non seulement le *contenu* de l'expérience perceptuelle est déterminé par des événements extérieurs au cerveau (e.g. Dretske 1996), mais *les mécanismes qui donnent lieu ou réalisent* cette expérience s'étendent aussi partiellement dans l'environnement. Susan Hurley, Alva Noë et Kevin O'Regan reconnaissent que cet

¹⁶ J'emprunte, ici, le terme *externalisme actif* à Andy Clark et David Chalmers (1998). Cette même idée s'est aussi retrouvée dans littérature sous diverses appellations, dont *vehicle externalism* (Hurley 1998) ou *enabling externalism* (Hurley 2010). L'expression que je choisis est plus élégante que les deux dernières lorsque traduite en français.

« externalisme radical » peut paraître contre-intuitif à première vue. Pour cette raison ils invitent les objecteurs potentiels à ne pas rejeter cette hypothèse en se fondant uniquement sur leurs intuitions quant à la nature de l'esprit, mais à considérer plutôt le potentiel explicatif de cette thèse :

If someone really has no conception of how neural or internal functional properties—or indeed any others—could explain phenomenal qualities, then how can he be so confident that if phenomenal qualities can be explained, it must be internal factors that do the job? What is so magical about the boundary around internal factors? (Hurley 2010, p. 104).

C'est ce qui les porte à affirmer que le débat sur l'externalisme est une question *empirique* qui ne sera pas réglé (uniquement) par des réflexions philosophiques.

En ce sens, l'approche éactive n'est pas seulement une thèse métaphysique, mais est aussi associée à une *thèse sur la nature de l'explication des phénomènes mentaux* : en effet, si l'on accepte de considérer l'hypothèse que l'expérience perceptuelle n'est peut-être pas constituée exclusivement par des processus neuronaux, il faut aussi accepter qu'elle ne pourra être *expliquée* en s'en tenant seulement au *niveau* neurologique. L'approche éactive s'inscrit donc dans la lignée de la psychologie écologique de James Gibson¹⁷, pour qui : « la physique, l'optique, l'anatomie et la physiologie décrivent des faits, mais pas des faits *au niveau approprié* pour l'étude de la cognition » (1979, p. xiii, je souligne). La conséquence de cette posture théorique est que les sympathisants de la psychologie écologique ont concentré l'essentiel de leurs efforts non pas à découvrir les mécanismes neuronaux, mais plutôt à décrire la relation entre l'organisme et son environnement, ce qui, selon eux, permet réellement d'expliquer les phénomènes qui les intéressent :

¹⁷ Pour être exact, Gibson et les partisans de l'approche écologistes défendent une thèse plus forte que Hurley et compagnie. Pour Gibson, faire référence à l'environnement est nécessaire à l'explication du mental, alors que pour Hurley *il pourrait* être nécessaire.

the specialized neural mechanisms essential to vision have been of relatively little interest to ecological psychologists. Vision would be impossible without the brain, but it would be equally impossible without the optic array and the mobile organs of vision. (Neisser 1999, p. 255).

En mettant l'essentiel du pouvoir explicatif dans la relation entre l'organisme et son environnement, autant la psychologie écologique que l'approche énaactive semblent à première vue proposer une conception de l'explication qui soit incompatible, non seulement avec le *principe d'autonomie*, mais aussi avec le modèle d'explication prédominant en biologie et en neuroscience, à savoir, l'explication mécaniste. Voilà pourquoi certains philosophes ont opposé une fin de non-recevoir à l'approche énaactive, faisant valoir que celle-ci ignorait le succès des recherches visant à découvrir les corrélats neuronaux de la conscience (NCC) (Prinz 2012 p. 171, Prinz & Jack 2004). Le modèle d'explication mécaniste¹⁸ en philosophie des sciences avance que les phénomènes sont expliqués en *identifiant* et en *décomposant* les mécanismes qui en sont responsables. Cette décomposition implique notamment d'identifier les *opérations* qui donnent lieu à ces phénomènes et de les associer à des entités physiques *localisables*.

C'est justement cette décomposition qui *peut sembler engendrer une tension* avec l'énaactivisme puisque le réflexe naturel est de chercher ces mécanismes *dans* l'organisme lui même (e.g. dans la boîte crânienne, dans le lobe occipital, etc.);

¹⁸ Au cours des dernières années, quelques philosophes ont commencé à s'intéresser à la manière dont les scientifiques expliquent des phénomènes dans les sciences de la vie et en neurosciences. Si la réflexion philosophique sur le sujet est en grande partie descriptive, il y a aussi un aspect normatif, dont l'objectif est de proposer des principes sur les bonnes pratiques de recherche. Carl Craver, par exemple, insiste sur le fait que la conception manipulationniste de la causalité qui fonde son modèle explicatif permet de distinguer les simples corrélations des causes *pertinentes*. Dans ces pages, je me contente essentiellement de *décrire* en quoi consiste l'explication mécaniste en l'opposant notamment au modèle d'explication déductif-nomologique sans pousser l'analyse philosophique sur ce sujet plus loin. Enfin, je mentionne que, bien que je suive principalement l'analyse de William Bechtel et de ses collaborateurs, d'autres perspectives ont aussi été adoptées (e.g. Glennan 1996, Craver 2007).

l'internalisme étant d'ailleurs souvent la « position par défaut » adoptée par les scientifiques (McCauley & Bechtel 2001, Poirier & Faucher 2009, Bechtel & Richardson 2010). Cela étant dit, l'approche énaactive n'est pas nécessairement incompatible avec l'explication mécaniste, dans la mesure où cette dernière reconnaît l'importance des sciences d'ordre supérieur, et ce, malgré le fait qu'elle adhère à une certaine forme de réductionnisme. Plus précisément, selon l'explication mécaniste, le pouvoir explicatif réside dans l'interaction entre les niveaux d'explication distincts : c'est en décomposant un mécanisme que l'on *comprend pourquoi* il donne lieu au phénomène qui nous intéresse. William Bechtel (2008 chap. 4, 2009, §3) a d'ailleurs manifesté une certaine ouverture à la psychologie écologique et aux thèses de l'esprit étendu en insistant sur le fait qu'il faille rendre compte de la nature des *entrées [input]* que reçoit l'organisme de la part de son environnement. Pour que l'explication soit satisfaisante, donc, chaque niveau est pertinent : le chercheur doit non seulement morceler le mécanisme en une hiérarchie de parties et d'opérations, mais il doit aussi découvrir l'organisation entre ces composantes, organisation qui n'est pas toujours séquentielle. Il doit, de plus, comprendre comment le mécanisme s'intègre dans un réseau plus large de causes, que se soit au sein de l'organisme ou de l'environnement. L'énaactivisme et l'approche mécaniste sont donc compatibles, puisqu'il n'est pas exclu *a priori* que les mécanismes pertinents à l'explication des phénomènes cognitifs aient des composantes localisées en dehors du cerveau.

3.2 De l'externalisme psychologique à l'externalisme énaactif

Pendant plusieurs années, et encore aujourd'hui dans une moindre mesure, le débat entre individualistes et externalistes fut le pain quotidien des philosophes dont le domaine d'étude était la philosophie de l'esprit. Il n'est donc pas étonnant que celui-ci ait occasionné une littérature considérable, formant aujourd'hui l'arrière-plan des arguments donnés en faveur de l'externalisme énaactif.

Tel que je l'ai déjà mentionné à plusieurs reprises, le débat entre internalisme et externalisme concerne la nature des représentations mentales, notion que j'ai introduit dans le deuxième chapitre. De plus, bien que la paire de termes « externalisme » et « internalisme » soient des termes largement utilisés en philosophie de l'esprit et en sciences cognitives, ils ne sont pas pour autant être définis explicitement à chaque fois.

Pour commencer, une façon intuitive, mais peut-être légèrement inexacte, de concevoir la distinction entre externalisme et l'internalisme consiste à affirmer que l'externalisme est la position selon laquelle une propriété n'est pas *intrinsèque* à un système, mais dépend plutôt de ce qui se passe à l'extérieur de celui-ci, ce qui en fait une propriété *relationnelle*. C'est d'ailleurs la manière dont Tyler Burge introduit la notion dans son tout récent ouvrage :

Individualism is the view that all of an individual's representational mental kinds are constitutively independent of any relation to a wider reality. *Anti-individualism* maintains that many representational mental states and events are constitutively what they are partly by virtue of relations between the individual in those states and a wider reality. (2007, p. 3).

On pourrait par contre opposer à cette manière de concevoir l'internalisme et l'externalisme le fait que certaines propriétés peuvent être extrinsèques ou relationnelles tout en étant « internes » : le fait que je sois citoyen d'un pays donné est une propriété relationnelle (dans ce cas-ci, elle dépend notamment de l'existence de cet état), sans pour autant m'être extérieure. Je retournerai sur cette distinction plus en détail un peu plus bas.

Par ailleurs, comme nous pouvons le voir dans cet extrait, en philosophie de l'esprit les débats à propos de l'externalisme portaient principalement sur l'*individuation* du contenu représentationnel, où ce qu'on a appelé l'**externalisme psychologique**. Tel

que je l'ai déjà mentionné, par individuation, je fait ici référence à l'*identité* des états mentaux : le fait que ma croyance soit à propos *du chat sur la table* plutôt qu'à propos *du livre sur la table* est dû à son contenu. Ceux qui défendent des thèses externalistes à l'égard du contenu mental soutiennent donc que celui-ci est en partie déterminé par des propriétés extrinsèques à l'agent, c'est-à-dire son *contexte*. Par là, on entend notamment l'environnement physique ou social dans lequel l'agent est enchâssé, de même que son histoire causale (e.g. Burge 1979, 1986). L'intuition centrale derrière ce type de thèse étant que « les états représentationnels ne peuvent se produire dans un *vacuum* » (Burge 2007, p. 3). En effet, puisque les états mentaux sont *intentionnels*, c'est-à-dire *à propos de quelque chose dans le monde*, ils sont aussi *relationnels*, et donc, en partie déterminés par des propriétés qui sont extérieures aux agents. Cependant, contrairement aux théories de l'esprit énoncées et étendues, ces arguments n'ont jamais remis en question le fait que les processus cognitifs qui *permettent* aux agents d'entretenir des états mentaux (i.e. les véhicules représentationnels) se trouvent, eux, dans le cerveau¹⁹.

Toutes les positions externalistes accentuent le rôle de l'environnement pour le mental, mais en ayant toutefois recours à des conceptions distinctes de l'environnement. Essentiellement, l'externalisme psychologique se demande ce qui se

¹⁹ Il n'est pas clair que l'externalisme psychologique de Putnam-Burge entre en opposition avec la thèse de la survenance neuronale telle que définie plus haut. Dans une publication récente, par exemple, Burge regrette d'avoir laissé croire que son anti-individualisme est un rejet de la thèse de la survenance locale. Comme il le rappelle d'ailleurs à propos de l'expérience de pensée de la Terre Jumelle, « the main point of the thought experiment was not to establish that local supervenience fails, but to establish that reference and 'meaning' depend not on the character of an individual's body but on relations to the wider environment. [...] So some twin-earth thought experiments might be taken to support rejection of local supervenience, while others need not be. » (2007, p. 19, note §27). Il se trouve, en effet, que « In the third place (and most importantly), individualism and anti-individualism are not fundamentally about supervenience, but about the natures of mental states, their correct individuation conditions. » (2006, p. 153, note §2). En ce sens, Burge ne laisse que très peu d'indices sur l'externalisme des véhicules représentationnels qui nous intéresse ici. Dans sa discussion de *Vision* de David Marr, par exemple, il affirme que les représentations postulées par Marr vont dans le sens de sa théorie anti-individualiste, sans affirmer pour autant que les processus cognitifs qui donnent lieu à la vision s'étendent dans l'environnement.

passerait avec les propriétés sémantiques des états mentaux d'un agent si celui-ci était dans un autre monde possible. L'environnement est donc conçu de manière très abstraite. De leur côté, ceux qui défendent une version ou une autre de l'externalisme actif adoptent une conception gibsonienne de l'environnement, où la relation entre ce dernier et l'organisme qui l'habite est *mutuelle et causale* (Gibson 1979, chap. 1).

À terme, mon objectif ici est à la fois de proposer une taxonomie des différents types d'externalismes, mais aussi de faire valoir en quoi l'externalisme énonctif diffère des autres positions externalistes défendues jusqu'à présent. Pour ce faire, je vais m'inspirer de la taxinomie introduite par Susan Hurley (2010), qui a réparti les types d'externalismes selon deux axes (avec cependant quelques variations dans les détails). Elle oppose, d'une part, l'externalisme concernant le contenu représentationnel à l'externalisme concernant le caractère phénoménal et, d'autre part, l'externalisme concernant l'*individuation* du contenu de nos représentations (ce que j'appelle ici l'externalisme psychologique) à l'externalisme concernant les *véhicules* de ces représentations.

Avant de passer aux arguments en faveur de l'externalisme psychologique, je dois faire une dernière précision. Tel que je l'ai mentionné plus haut, l'externalisme psychologique concerne l'individuation ou l'identité des états mentaux; en ce sens, nous avons ici affaire à une *thèse ontologique* sur la nature des états mentaux.

L'*externalisme véhiculaire*, quant à lui, tel qu'adopté par certains tenants de la théorie de l'esprit incarné porte *sur la nature de l'explication des phénomènes mentaux*. Compris en ce sens, l'internalisme et l'externalisme sont des doctrines théoriques à propos de la localisation des composantes d'un mécanisme particulier, mécanisme donnant lieu au phénomène étudié (c.f. Poirier et Faucher 2009, section 1). De sorte que, l'*internaliste* maintient que les composantes responsables d'un phénomène donné seront trouvées dans le système original; alors que l'*externaliste* défend que certaines

composantes seront dans le système original, mais que d'autres seront à l'extérieur de celui-ci. À titre d'exemple, en neurosciences, les chercheurs s'intéressant à la vision pensaient initialement que le système responsable de la vision se trouvait dans l'aire de Brodmann 17 (V1); plusieurs recherches subséquentes firent cependant comprendre que certains aspects de la vision devaient trouver leur explication en dehors de cette aire cérébrale. Par la suite, dans les années quatre-vingt-dix, les chercheurs oeuvrant dans ce domaine ont identifié jusqu'à 32 aires cérébrales répondant à des stimuli visuels (Bechtel 2009 p. 545 - 549). Les notions de composantes et de mécanisme seront expliquées plus en détail lorsque j'aborderai l'approche mécaniste. Cette manière de concevoir la distinction internalisme/externalisme permet d'éviter les problèmes rencontrés lorsqu'on la rabat sur la distinction entre propriétés intrinsèques et propriétés relationnelles.

3.2.1 L'externalisme psychologique

Initialement, les arguments externalistes à propos de l'individuation du contenu mental trouvent leur origine dans les arguments externalistes en sémantique²⁰ et s'opposent à la thèse connue sous le nom d'*individualisme* ou d'*internalisme*.

Un certain nombre de philosophes ont suggéré que le contenu de nos états mentaux dépendait entièrement de ce qui se passe dans notre cerveau, et que par conséquent, il n'était pas déterminé par ce qui se passe dans le monde extérieur, sinon pour les effets de ce dernier sur notre cerveau sous forme d'*entrée* [*input*]. Ce que cela veut dire,

²⁰ Faute d'espace, je ne peux poursuivre plus loin sur cette question. Putnam (1975) et Kripke (1980) ont défendu des théories causales de la référence contre les théories descriptives. Soutenir une théorie descriptive dans le cas des noms propres, consiste à dire qu'un nom propre réfère par l'entremise du contenu descriptif qui lui est associé. L'entité (unique) qui satisfait la description est donc celle à laquelle réfère le nom propre. D'un autre côté, l'idée centrale derrière les théories causales est qu'un nom réfère à l'entité à laquelle il est connecté suite à un acte baptismal (e.g. désormais x sera appelé « John »). La référence est ensuite transmise de locuteur en locuteur. Le nom fait donc référence à l'entité qui se trouve au début de la chaîne causale.

c'est que si un scientifique en arrivait à créer une copie molécule-par-molécule de mon cerveau et à le maintenir en vie dans une cuve, tout en le stimulant artificiellement de sorte qu'il serait sujet aux mêmes événements chimiques ou électriques que mon cerveau actuel, ce cerveau artificiel aurait exactement les mêmes états mentaux que moi. Ceux qui acceptent cette idée sont individualistes²¹ à l'égard du contenu mental, alors que ceux qui la rejettent sont externalistes.

Comme je l'ai mentionné plus haut, la position individualiste est généralement considérée comme étant très intuitive. Cela étant dit, ceux qui défendent une forme ou une autre d'externalisme n'ont pas besoin de démontrer que la position individualiste est incohérente : en règle générale, la stratégie utilisée consiste à développer un argument similaire à celui de la *Terre Jumelle* employé par Hilary Putnam dans *The meaning of 'meaning'* (1975). À cet effet, Susan Hurley (2010, p. 147 note 1) note avec raison que l'individualisme a un fardeau de la preuve plus lourd que les diverses positions externalistes. Ceux-ci, en effet, peuvent se contenter de fournir des contre-exemples exposant les limites des positions individualistes, alors que les internalistes doivent pouvoir soutenir une thèse qui puisse s'appliquer à tous les états mentaux visés.

L'argument de la Terre Jumelle repose sur une expérience de pensée dans laquelle deux individus qui sont des copies molécule par molécule (les « jumeaux ») se retrouvent dans deux environnements ou dans deux mondes identiques, sauf pour ce qui est d'une différence significative pour la vie cognitive de l'agent (habituellement l'objet intentionnel de certains états mentaux). Par exemple, la formulation classique que l'on retrouve dans Putnam 1975 nous demande d'imaginer un monde identique au

²¹ Je réfère ici la conception de l'individualisme proposé, et éventuellement critiquée, par Tyler Burge « According to individualism about the mind, the mental natures of all a person's or animal's mental states (and events) are such that there is no necessary or deep individuating relation between the individual's being in states of those kinds and the nature of the individual's physical or social environments. » (1986, p. 3-4).

nôtre excepté sous un aspect particulier : que la substance qui emplie les rivières et les océans, et qui entre dans la composition des organismes n'est pas la molécule H_2O , mais plutôt la molécule XYZ (c'est la « Terre Jumelle »). Ces deux molécules sont similaires à première vue, toutes deux sont inodores, ne goûtent rien, sont transparentes etc., et ne peuvent être distinguées sans une fine analyse chimique. Si l'on imagine une situation où les jumeaux affirment tous les deux que « l'eau de mer est salée » et que l'on prend pour acquis qu'aucun des jumeaux ne connaît la différence entre H_2O et XYZ, celui qui est sur Terre aura une croyance qui est vraie, alors que celui qui est sur la Terre jumelle aura une croyance qui est fausse. Autrement dit, *les externalistes affirment que les jumeaux font référence à un objet différents lorsqu'ils disent ce même énoncé*. En effet, ils expriment des croyances différentes, même s'ils affirment le même énoncé, car une même croyance ne peut être à la fois vraie et fausse. Or, les jumeaux étant des copies molécule par molécule, le contenu de leurs états mentaux ne peut être déterminé par ce qui se passe dans leur cerveau respectif : il est plutôt déterminé par l'histoire causale et le monde dans lequel ils se situent²².

Dans son article, l'argument de Putnam porte sur les espèces naturelles [*natural kinds*], mais le même type d'argument peut aussi être formulé dans le cas croyances auto-référentielles, des croyances à propos de notre propre localisation spatio-temporelle et de nos croyances à propos des autres personnes (c.f. Stich 1973, p. 706–8). Il semble donc plausible que l'argument de Putnam puisse s'étendre à une variété de contenus représentationnels. Et en effet, aujourd'hui, l'influence de ces arguments fait en sorte que l'externalisme psychologique est une thèse largement partagée en philosophie de l'esprit. Plusieurs philosophes acceptent donc aujourd'hui les conclusions de l'argument de la Terre Jumelle : le *contenu* d'un état mental n'est pas

²² Burge (1979) a aussi présenté des arguments influents à propos de la contribution de notre environnement *social* dans la détermination de notre contenu mental, notamment dans ce qu'il appelle la *déférence sémantique*.

uniquement déterminé par ce qui se passe dans la tête de l'agent, mais aussi par ce qui se passe dans son environnement. Précisons, cependant, que cette forme d'externalisme n'entre pas directement en conflit avec les principes méthodologiques des neurosciences, puisque les véhicules représentationnels restent des états neuronaux. Enfin, l'externalisme psychologique paraît moins contre-intuitif lorsqu'on considère le fait que les états mentaux ne sont pas les seules entités à voir leur identité constituée de manière relationnelle. Pour citer encore une fois Tyler Burge :

The identity of a heart depends on its function in the whole body—on its relations to parts of the body outside the heart. In a crudely analogous way, the identities of some mental kinds depend on those kinds' relations to entities beyond the individual's body. They depend on cognitive function, on obtaining information, in an environment in something like the way the kind *heart* depends for its individuation on the function of the heart in the body that contains it. (1989, p. 278).

Ainsi, l'*identité* des états mentaux, tout comme celle de nos organes, est définie par le rôle causal que jouent ceux-ci au sein d'un système plus large.

Dans sa version initiale telle que défendue par Putnam, Kripke et Burge, l'externalisme psychologique concerne seulement la signification et le contenu de nos états représentationnels comme nos croyances ou nos désirs. Poursuivant dans la même voie, d'autres auteurs ont aussi étendu le même raisonnement au caractère phénoménal qui caractérise certains de nos états mentaux, comme nos perceptions. Ainsi, Dretske (1995), Tye (1995) et Lycan (1996), par exemple, ont élaboré diverses formes de théories représentationnalistes de la conscience²³, dont l'idée centrale

²³ Évaluer les forces et faiblesses des théories représentationnelles de la conscience nous mènerait à trop loin tout en n'étant pas nécessaire à l'élaboration de mon argument : ce qui est important de saisir, est le fait que l'externalisme du caractère phénoménal défendu par Dretske, Tye et Lycan repose ultimement sur le même type d'argument que l'externalisme psychologique. Les deux arguments principaux en faveur de l'externalisme phénoménal sont les suivants. D'une part, comme je l'ai dit dans le chapitre précédent, réduire l'aspect phénoménal du mental à des propriétés représentationnelles semble faciliter la possibilité d'une naturalisation de la

consiste soit à identifier le caractère phénoménal avec le contenu représentationnel, soit à affirmer que le caractère phénoménal *survient* sur ce-dernier. Dretske, Tye et Lycan sont représentationnalistes à propos du caractère phénoménal dans le sens où ils affirment que le contenu représentationnel rend compte *totalemment* du caractère phénoménal : les propriétés dont nous faisons l'expérience (les qualia) *sont* les propriétés telles qu'elles sont représentées. En adoptant à la fois les arguments externalistes à la Putnam et une position représentationnaliste à l'égard du caractère phénoménal, ils adhèrent ainsi à une forme d'externalisme psychologique à propos du caractère phénoménal de nos expériences. Puisque que les externalismes actif ou énonactif ont recours à des arguments d'une toute autre forme, il nécessaire ici de s'étendre sur l'externalisme présenté par ces auteurs.

3.2.2 L'externalisme actif

Farouchement opposées au « neurocentrisme » inhérents aux sciences cognitives traditionnelles, les diverses positions regroupées sous l'étiquette « d'externalisme actif » ont des visées radicalement différentes de celle de l'externalisme psychologique tel qu'exposé dans la section précédente. En conséquence, ceux et celles qui défendent l'externalisme actif ne font pas appel aux expériences de pensée dans le style de la *Terre jumelle*. En effet, comme nous venons de le voir, l'externalisme psychologique est la thèse selon laquelle le contenu est en partie déterminé par l'environnement physique ou social dans lequel l'agent est enchâssé ;

conscience, puisque plusieurs théories naturalistes des représentations mentales sont aujourd'hui proposées. En ce sens, les théories représentationnelles de la conscience suivent deux étapes : d'abord, réduire le contenu phénoménal au contenu représentationnel, ensuite, réduire le contenu représentationnel à des processus descriptibles en termes physiques. D'autre part, les représentationnalistes invoquent souvent l'argument de *transparence de l'expérience* en faveur de leur théorie. Par transparence de l'expérience, on veut dire que lorsque nous « introspectons » notre expérience, la seule chose dont nous prenons connaissance c'est *ce à propos de quoi* est notre expérience : « Look at a tree and try to turn your attention to intrinsic features of your visual experience. I predict you will find that the only features there to turn your attention to will be features of the presented tree. » (Harman 1990, p. 667).

l'externalisme actif ou véhiculaire, quant à lui, porte sur les *processus* ou les *mécanismes* qui permettent aux agents d'entretenir des états avec un contenu représentationnel ou un caractère phénoménal particulier (Hurley 2010, p. 101 et suiv.). Les arguments soutenant l'externalisme actif insistent sur le rôle *actif* de l'organisme enchâssé dans son environnement ; activité supposément nécessaire à la constitution de certains processus cognitifs. Ainsi, par *externalisme actif* j'entends :

externalisme actif : La cognition n'est pas un processus se déroulant seulement dans l'esprit ou le cerveau de l'organisme, mais est plutôt distribuée à travers les différentes parties prenantes de la situation, incluant le cerveau, le corps et l'environnement. Les véhicules de certains processus cognitifs sont donc réalisés *partiellement* en dehors du cerveau ²⁴.

Cela étant dit, pour ceux qui s'intéressent à l'explication en sciences cognitives ou en neurosciences, la question qui saute maintenant à l'esprit est la suivante : si les structures environnementales peuvent entrer dans constitution de l'esprit, comment discriminer les processus qui sont cognitifs de ceux qui ne le sont pas ? Le cognitivisme avait une réponse simple : la cognition est l'ensemble des processus d'information qui se déroulent dans le cerveau. En effet, par opposition au béhaviorisme, dont l'objet d'étude est le comportement observable des agents, le cognitivisme postule des représentations internes inobservables pour rendre compte de ces comportements. Ainsi, avec la fin du béhaviorisme à la fin des années cinquante et la montée du cognitivisme, l'*explanandum* de la psychologie est passé du comportement des agents aux représentations internes donnant lieu à la mémoire, à la résolution de problème, etc.; la cognition étant définie comme étant la manipulation,

²⁴ Je laisse la thèse de l'externalisme actif aussi imprécise délibérément. Mon point est justement de faire valoir que de nombreux chercheurs ont adopté une variante plus ou moins forte de l'externalisme actif, parmi ceux-ci on compte notamment Beer 1995, Clark 1997, Clark et Chalmers 1998, Hurley 1998, Menary 2007, Thelen & Smith 1994, Varela Thompson et Rosh 1991, van Gelder 1995, Wheeler 2005

la transformation ou le stockage de ces représentations. Comme on peut s'en douter, ceux qui défendent l'externalisme actif affirment qu'il faut remettre en question cette conception de la cognition.

Mark Rowlands souligne judicieusement à cet effet que les processus que l'on considère comme étant le propre de la cognition ou de l'esprit sont habituellement choisis non pas à l'aide de critères précis, mais plutôt par ostension. Ainsi, on inclut traditionnellement parmi ceux-ci la perception, la mémoire, la raison, le langage, le comportement intelligent, etc. Afin de déterminer quels sont les processus que nous pouvons qualifier de « cognitifs » sans faire de pétition de principe, ceux qui défendent l'externalisme actif affirment que la cognition est ce qui permet d'exécuter des tâches cognitives, e.g. : percevoir le monde, se souvenir d'informations déjà mémorisées, avoir la capacité à utiliser cette information dans des raisonnements (Rowlands 2003, p. 157 et suiv.).

Les théories qui adhèrent à l'externalisme actif sont généralement tributaires des travaux de Rodney Brooks (1991a, b) en robotique de même que des travaux de Dana Ballard (1991) sur la « vision animée » et les saccades visuelles. Ces travaux tendent à montrer que les organismes n'ont pas besoin de se construire des représentations détaillées (i.e., à haute résolution) pour guider leur comportement, mais peuvent se contenter d'aller chercher l'information dont ils ont besoin de manière répétée. Comme le dit la formule célèbre introduite par Brooks, les agents peuvent utiliser le monde comme étant son propre meilleur modèle [*the world as its own best model*]. Dans la même veine, ceux qui défendent l'externalisme actif affirment qu'il n'est pas nécessaire de postuler des processus cognitifs internes gourmands en ressources pour expliquer le comportement, puisque l'organisme peut manipuler son environnement afin de résoudre les problèmes qui se pose à lui. C'est ce que Rowlands (1999), et Menary (2007) après lui, ont appelé la *thèse de la manipulation de l'environnement*

[*manipulation thesis*]²⁵. Dans ce contexte, on ne s'étonnera pas que les actions corporelles, de même que les actions épistémiques, y jouent un rôle central.

Ceux qui défendent l'externalisme actif ont habituellement eu recours à deux types d'arguments : celui du fonctionnalisme étendu [*extended functionalism*] et celui de l'intégration cognitive²⁶. Bien que distincts des arguments en faveur du fonctionnalisme étendu, ceux en faveur de l'intégration cognitive sont parfois présentés conjointement. Je vais présenter ces deux arguments tour à tour.

Le fonctionnalisme étendu. En philosophie de l'esprit, on entend par *fonctionnalisme*, la thèse selon laquelle ce qui fait de quelque chose un état mental d'un type particulier, par exemple, le fait d'être une croyance plutôt qu'un désir, n'est pas sa constitution physique, mais plutôt le rôle causal qu'il joue dans le système dont il fait partie (e.g. Putnam 1967). L'idée, encore largement partagée, selon laquelle les états mentaux sont sujets à la réalisation multiple est au coeur de l'argument en faveur du fonctionnalisme ; la *réalisation multiple* étant la thèse selon laquelle des phénomènes d'ordre supérieur, comme les états mentaux ou les tornades, peuvent être réalisés physiquement de manière différente.

Pour justifier cette thèse, Putnam a recours à l'exemple de la douleur : l'être humain n'est pas la seule espèce qui semble éprouver de la douleur, les autres mammifères, de même que les reptiles, les mollusques, (voire même les martiens) semblent éprouver de la douleur. Étant donné les différences entre les cerveaux de chacune de ces espèces, il est peu probable qu'il y ait quoique ce soit de commun dans les divers états

²⁵ La définition originale telle qu'on la retrouve chez Rowlands : « Cognitive processes are not located exclusively in the skin of cognising organisms because such processes are, in part, made up of physical or bodily manipulation of structures in the environments of such organisms. » (1999, p. 23).

²⁶ Cette discussion est en grande partie inspirée de Sutton 2010, Menary 2010 et Kirchhoff 2012, tout en rejetant leur distinction entre « première vague » et « seconde vague » d'arguments. Je m'éloigne aussi de leur propos dans les détails.

neuronaux qui donnent lieu à la douleur chez chacune de ces espèces. En ce sens, le fonctionnalisme est une réaction aux théories matérialistes (comme la théorie de l'identité) qui proposaient une relation trop étroite entre les états mentaux et les états neuronaux (e.g. le douleur = l'activation de la fibre-C, pour reprendre une formule fameuse chez les philosophes). Le fonctionnalisme est donc une manière de caractériser les états mentaux de manière assez abstraite pour qu'ils ne soient pas attachés à des structures neuronales particulières ; cette abstraction permettant ensuite de faire des généralisations intéressantes qui serviront à expliquer ou prédire le comportement (e.g. ceux qui éprouvent de la douleur tendent à vouloir s'éloigner du stimulus). (Pour en savoir davantage sur la relation entre fonctionnalisme, réalisation multiple et explication en psychologie, voir Fodor 1974).

Ceux qui défendent le fonctionnalisme étendu, comme Clark et Chalmers (1998), font valoir que si nous acceptons le fonctionnalisme traditionnel, alors nous n'avons pas de bonnes raisons d'en refuser la version *étendue*. En effet, si l'on admet que les états mentaux sont définis par leur rôle fonctionnel, il est tout aussi injustifié de refuser qu'ils puissent être réalisés dans un cerveau en silicone, que de refuser qu'ils soient (partiellement) réalisés dans l'environnement de l'agent²⁷. Clark et Chalmers affirment par conséquent qu'on ne peut décider arbitrairement que le mental se limite à ce qui se trouve dans le cerveau, puisque c'est cette frontière même qui est remise en question. Ils affirment ainsi une *parité fonctionnelle* entre les processus cognitifs localisés dans le cerveau et ceux qui lui sont extérieurs et qui consistent en la

²⁷ Un contre-argument célèbre à la théorie de l'esprit étendu défendu par Clark et Chalmers et qui fait appel au fonctionnalisme étendu, est celui proposé par Adams et Aizawa (2010). Ces derniers affirment que c'est une chose que dire que quelque chose est lié causalement à un processus cognitif dans le cerveau lorsque l'agent effectue des tâches cognitives, s'en est une autre que de dire que cette chose *fait partie* de l'esprit. Adams et Aizawa défendent que la marque du mental est l'*intentionnalité non-dérivée* et que les représentations externes, les photographies, par exemple, n'ont qu'une *intentionnalité dérivée*. Autrement dit, selon Adams et Aizawa, une photographie est à *propos* d'un paysage donné que parce qu'elle est interprétée par un agent (Contrairement à ce qu'en dit Clark (2010), ils n'affirment pas que tous les processus qui se déroulent dans le cerveau sont des processus mentaux).

manipulation, l'exploitation et la transformation de certaines structures d'information adéquates dans l'environnement. Pour citer Clark :

If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process.(Clark 2010, p. 44).

Plusieurs exemples ont été proposés dans la littérature afin d'illustrer cette idée. Je vais m'en tenir ici à l'exemple classique d'Otto et d'Inga introduit dans l'article de Clark et Chalmers cité plus haut et dont l'objectif est de montrer que nos *croyances* peuvent être réalisées à l'extérieur du cerveau. Les auteurs décrivent le cas suivant, impliquant Inga, une personne normale, et Otto, souffrant d'Alzheimer. Selon l'histoire, nos deux protagonistes apprennent qu'il y a présentement une exposition intéressante au *Museum of Modern Art* de New York (MoMA). Voulant s'y rendre, Inga réfléchit pendant un instant, se souvient que celui-ci se trouve sur la 53^e rue, et entame son chemin vers le MoMA, conformément à sa croyance sur son emplacement géographique. Pour les auteurs, il est évident qu'avant même d'apprendre l'existence de cette exposition, Inga avait déjà une croyance *dispositionnelle* à propos de l'endroit où se trouve le MoMA. Celle-ci est devenue *occurrente* seulement lorsque Inga a consulté sa mémoire.

Otto, de son côté, souffre d'Alzheimer et doit donc avoir recours à l'information qui se trouve dans son environnement afin de pouvoir donner libre cours à ses activités quotidiennes (comme plusieurs autres personnes dans la même condition). Il traîne donc avec lui un cahier où il note les nouvelles informations qu'il juge pertinentes. Dans les faits, le cahier de note de Otto joue pour lui le rôle d'une sorte de mémoire externe qui remplace sa mémoire biologique. Ainsi, à chaque fois qu'il doit récupérer de l'information déjà acquise, plutôt que de consulter sa mémoire biologique (défaillante) il consulte son cahier. Entendant parler de l'exposition au MoMA et

disirant la voir, il consulte son cahier là où il avait indiqué l'adresse du musée, et se rend sur le 53^e rue.

Les auteurs affirment que si l'on accepte le principe de parité fonctionnelle, il faut aussi accepter de considérer Otto et son cahier comme un système cognitif à part entière :

For in relevant respects the cases are entirely analogous: the notebook plays for Otto the same role that memory plays for Inga. The information in the notebook functions just like the information constituting an ordinary non-occurrent *belief*; it just happens that this information lies beyond the skin. (1998, p. 33-34, je souligne).

Clark et Chalmers précisent, par contre, cette affirmation en insistant sur le fait que les états qui s'étendent partiellement dans l'environnement ne peuvent être considérées comme de réelles croyances si elles ne remplissent pas certaines conditions bien précises. Par exemple, si l'information inscrite dans le cahier d'Otto était trop souvent *inaccessible* à celui-ci, il y aurait un problème, car celle-ci ne pourrait jouer le même rôle que nos croyances dans la conduite de notre comportement. En effet, cette information doit être disponible dans la plupart des situations pertinentes pour être considérée comme étant une croyance au sens propre du terme (p. 36-7). Dans une autre publication, Clark précise la nature de couplage cerveau-environnement pour qu'un état étendu puisse être qualifié de « croyance » : **a.** celui-ci doit être disponible, fiable et utilisé dans des situations pertinentes [*reliably available and typically invoked*], **b.** automatiquement accepté, dans le sens où les croyances de l'agent ne doivent pas lui paraître suspectes et finalement **c.** facilement accessible lorsque nécessaire. (Clark 2010, p. 46). Pour une discussion plus approfondie du fonctionnalisme étendu, c.f. Clark 2008 p. 86 et suiv., Wheeler 2010.

L'intégration cognitive. Le type d'externalisme actif discuté jusqu'à présent était basé sur le principe de parité, c'est-à-dire sur la similarité ou la dissimilitude fonctionnelle entre des états se trouvant à l'extérieur du cerveau et des états traditionnellement considérés comme étant cognitifs. En ce sens, le point de départ de l'argument fonctionnaliste reste ce que nous considérons communément comme étant de la cognition.

Cela dit, contrairement à l'argument précédent, plusieurs auteurs (e.g. Menary 2007) ont présenté des arguments alternatifs, mais pas nécessairement incompatibles, au fonctionnalisme étendu. L'argument de l'*intégration cognitive*, où l'accent est mis sur la *complémentarité* entre l'environnement et l'agent. Menary remarque à cet effet que la motivation principale en faveur de « l'intégration cognitive » n'est pas le principe de parité, qui étend l'esprit en dehors de la boîte crânienne, mais plutôt le « fait brut » de notre corporalité (ibid., p. 228–9). Ce fait brut vient justifier ce que Rowlands a appelé la *thèse manipulationniste*. Pour Menary (2010), par exemple, ce qui fait de certains objets de l'environnement des composantes de notre esprit, c'est le fait qu'elles soient des parties intégrantes dans la résolution de tâches cognitives.

Ceux qui défendent ce type d'arguments rejettent l'idée qu'il doive y avoir une forme de similarité fonctionnelle entre ce qu'on considère communément être des états mentaux et les états mentaux étendus. Ce qui importe, plutôt, c'est l'intégration des processus étendus et des processus internes au sein d'un système cognitif hybride. L'idée derrière « l'intégration cognitive » est qu'il y a une étroite interaction causale entre les processus neuronaux internes et les processus environnementaux externes : ainsi, en insistant sur la *complémentarité* des véhicules internes et externes, de même que sur leur intégration au sein de d'un tout cognitif, Menary en conclut que le système complexe constitué à la fois de l'agent *actif* et de son environnement doit être considéré comme étant un système cognitif à part entière.

Un retrouve un bel exemple d'intégration cognitive dans un article publié par John Sutton récemment (2010, p. 201 et suiv.). S'inspirant des travaux d'Evelyn Tribble (2005), qu'il affirme être l'application de la théorie de l'esprit étendu la plus intrigante et ayant eu le plus de succès jusqu'à présent, Sutton s'intéresse à la manière dont les acteurs apprenaient leurs rôles au *Globe Theater* à l'époque shakespearienne. Les acteurs devaient, en effet, apprendre un nombre impressionnant de rôles en un temps très limité. Par exemple, entre 1594 et 1597, un acteur important comme Edward Alleyn a dû maîtriser pas moins de soixante-et-onze rôles, dont une cinquantaine était nouvellement appris. Comment expliquer une performance aussi impressionnante ? Une hypothèse serait que les acteurs devaient apprendre tous leurs rôles par coeur, en répétant inlassablement leurs répliques. Evelyn Tribble, quant à elle, soutient que cette idée n'est pas réaliste : nos capacités cognitives étant bien trop limitées pour accomplir un tel exploit. Elle propose plutôt une autre hypothèse, en s'intéressant à la manière dont les différents outils et pratiques du *Théâtre* permettaient de former une structure cognitive, qui bien que contraignante et limitée, permet aussi des performances extraordinaires (Tribble 2005, p. 142, cité par Sutton 2010). Par exemple, selon Tribble, les artifacts présents sur scène servaient d'indices et faisaient en sorte d'amener les acteurs à la bonne place au bon moment, sans que ceux-ci aient à apprendre leur texte au complet.

L'intégration cognitive permet aussi une nouvelle interprétation de l'exemple d'Otto donné plus haut. On considère que le cahier fait partie de l'esprit d'Otto non pas parce qu'il y a similarité fonctionnelle entre les croyances réalisées biologiquement et celles dont le véhicule représentationnel est constitué de marques inscrites dans le cahier, mais parce qu'il est *intégré systématiquement* avec ses états internes dans l'accomplissement de diverses tâches cognitives.

Nous voyons ainsi que notre capacité à modifier notre environnement physique et social nous permet d'étendre nos capacités cognitives, ce que Clark appelle des

échafaudages cognitifs [scaffoldings]. Ceux-ci donnent ensuite lieu à d'autres modifications de l'environnement, générant un nombre toujours croissant de capacités cognitives, menant à d'autres modifications environnementales et ainsi de suite. Le langage en est l'exemple poussé à l'extrême : par le langage nous pouvons stocker et manipuler une masse d'information que nous ne pourrions jamais mémoriser ou manipuler dans notre cerveau (c.f. Clark 1997, 2008).

Cette complémentarité entre processus internes et externes a particulièrement intéressé les chercheurs (e.g. Beer 2000, van Gelder 1995, 1998, Port & van Gelder 1995) ayant recours à la théorie des systèmes dynamiques en mathématiques et qui ont aussi défendu des arguments similaires, en mettant l'accent sur le fait qu'il y a un « couplage dynamique » se déroulant en temps réel entre l'agent et son environnement. La motivation première derrière l'utilisation des systèmes dynamiques en sciences cognitives est d'avoir un *langage commun* permettant de décrire l'interaction entre les différentes variables d'un système complexe. Avec de tels outils mathématiques, on peut décrire abstraitement tous les éléments du système, éléments qui sont de nature très différente : d'un côté nous avons l'activité neuronale, de l'autre nous avons l'environnement dans lequel l'organisme est situé. Ces théories permettent donc de montrer comment des structures complexes peuvent émerger de la coopération d'une multitude de composantes individuelles.

Cette tendance à faire valoir que la complémentarité entre l'agent et son environnement afin d'étendre l'esprit en dehors de la boîte crânienne n'est pas sans susciter de certaines critiques. Adams et Aizawa (2010) mettent le doigt sur une des difficultés majeures de l'externalisme actif, à savoir, qu'il n'est pas facile de fournir un argument expliquant pourquoi des structures externes qui sont importantes dans la résolution de tâches cognitives doivent être considérées comme étant des parties *intégrantes de l'esprit en tant que tel*. Peu importe à quel point on met l'accent sur la complémentarité entre les véhicules externes et les processus neuronaux, cela ne veut

pas dire pour autant qu'il faille les considérer comme formant un système cognitif à par entière. Une des sources de ce problème est problématiquement le fait qu'il n'y pas, à ce jour, une « théorie bien établie de ce qu'est exactement le cognitif » (ibid. 2010, p. 68).

3.2.3 Externalisme éactif

Les arguments présentés plus haut en faveur de l'externalisme actif concernent l'aspect cognitif ou représentationnel de nos états mentaux, mais délaissent l'aspect phénoménal qui caractérise pourtant une part importante de notre vie mentale. À bien des égards, les arguments justifiant l'externalisme éactif s'inscrivent dans la même veine que ceux en faveur de l'intégration cognitive. Tel que je l'ai dit brièvement dans l'introduction, j'entends ici par *externaliste éactif*, la thèse suivante :

externalisme éactif : l'externalisme éactif est la thèse selon laquelle le contenu de notre expérience perceptuelle, incluant son *caractère phénoménal*, est constitué à la fois de **a.** processus *internes*, principalement liés à l'acquisition et à l'activation des connaissances des régularités sensorimotrices, mais aussi **b.** de processus *externes* où l'agent sonde son environnement et y explore les structures porteuses d'informations.

On voit, par conséquent, que l'externalisme éactif *est une forme spécifique d'externalisme actif* portant sur les propriétés représentationnelles et phénoménales de notre expérience perceptuelle. Tous deux ont recours à des arguments de nature similaire.

Plus précisément, d'après l'approche éactive, la perception visuelle est constituée de deux ingrédients : d'*attentes* et d'*habiletés*. Ces attentes concernent la manière dont l'influx sensoriel varie relativement à nos mouvements; selon les éactivistes, la

perception requiert d'avoir une connaissance implicite de ces attentes. En plus de ces attentes, un ensemble d'*habiletés* nous permettant d'agir sur le monde, c'est-à-dire de sonder et d'explorer les structures d'information qui se trouvent dans notre environnement sont nécessaires à la perception. La perception est donc un processus hybride, où le monde est considéré comme étant une source d'information que les agents explorent activement par des actes de perception, actes de perception qui se font par l'intermédiaire de savoir-faire sensorimoteurs (ce à quoi les énéactives réfèrent lorsqu'ils parlent de « connaissances des contingences sensorimotrices »).

L'approche énéactive sera expliquée plus en détail au prochain chapitre; cela étant dit, ce bref aperçu clarifie l'idée selon laquelle l'approche énéactive conçoit la perception comme étant une activité de l'organisme dans son environnement. Les tenants de l'énéactisme adoptent ainsi les arguments en faveur de l'intégration cognitive, présentés dans la section précédente : leur point de départ est le fait brut de notre incarnation. Par contre, certains ont remis en question le fait que l'approche énéactive soit compatible avec le fonctionnalisme étendu (e.g. Rowlands 2009). En effet, comme le note Rowlands, alors que le fonctionnalisme implique une certaine forme d'indifférence à l'égard de la manière dont les processus sont réalisés, l'importance que l'approche énéactive porte au corps de l'organisme semble exiger un rejet du fonctionnalisme.

Susan Hurley (2010) admet que l'externalisme des véhicules à propos de l'expérience phénoménale est la forme d'externalisme la plus radicale et la moins intuitive d'externalisme. Cependant, la taxinomie qu'elle propose entre les différentes formes d'externalismes permet de mettre en lumière un phénomène intéressant : malgré le déficit d'explication entre notre compréhension des processus neuronaux et l'expérience phénoménale, nos intuitions internalistes restent toujours aussi fortes. En effet, si nous n'avons pas la moindre idée de la manière dont l'activité neuronale donne lieu à l'expérience phénoménale, sur quel base pouvons-nous rejeter

l'externalisme éactif? Elle en conclut, donc, avec raison, que le fait que l'externalisme éactif soit contre-intuitif n'est pas, en soi, une preuve que cette thèse est fausse.

De plus, l'objectif de Susan Hurley dans cet article est de nous inviter à ne pas commettre ce qu'elle appelle l'*erreur causale-constitutive*. Cette erreur consiste à s'opposer à l'externalisme (én)actif en affirmant que les arguments externalistes donnent un rôle *constitutif* aux facteurs externes alors qu'ils n'ont, en fait, « qu'un rôle causal ». Selon Hurley, ceux qui commettent cette erreur n'ont pas d'argument ou de critère pour fonder la distinction causale-constitutive à part la distinction entre ce qui est interne et ce qui est externe à l'organisme (ibid., p. 106, pour un exemple de cela, voir Prinz 2006).

Cela dit, parce que l'externalisme éactif est une forme d'externalisme actif, la critique formulée par Adams et Aizawa, déjà mentionné dans la section précédente, semble aussi poser problème l'approche éactive. En effet, il ne suffit pas de maintenir qu'il y a une riche interaction causale entre certains processus internes et externes pour montrer que le système pertinent dans l'étude de la cognition, de la conscience ou de la perception implique aussi l'environnement. Après tout, comme le rappelle Clark, une dépendance au contexte est toujours la règle, et non l'exception (1998). Les prochaines sections sur l'explication mécaniste rendront ces considérations plus explicites.

3.3 L'externalisme et l'explication en sciences cognitives et en neurosciences

Dans les sections précédentes j'ai dressé un portrait des différentes variétés d'externalismes. On constate, ce faisant, qu'alors que l'externalisme psychologique à la Putnam concernant l'individuation du contenu représentationnel ou du contenu

phénoménal et peut sembler, à première vue, contre-intuitif, il n'est pas moins compatible avec le *principe d'autonomie*.

Comme je l'ai dit plus haut, ce principe guide implicitement la recherche en neuroscience et implique que les seules propriétés pertinentes pour l'explication en psychologie ou en neurosciences sont celles qui entraînent une différence dans les propriétés physiques internes de l'organisme. Micheal Tye, par exemple, défend une position externaliste à propos du caractère phénoménal, mais affirme cependant que la thèse de la survenance neuronale « n'est pas directement menacée par [sa] position. » (1992, p. 172), autrement dit, l'explication des phénomènes mentaux par l'analyse des mécanismes neuronaux qui les rendent possibles n'est pas directement remise en cause par les diverses formes d'externalismes psychologiques.

Or, il en va tout autrement des externalismes actif et éactif, dont l'objectif est justement de remettre en question la thèse de la survenance neuronale : le contenu perceptuel est une propriété de l'organisme en tant que tel, et non d'un processus se déroulant dans son cerveau. Conséquemment, le niveau d'explication adéquat pour expliquer la conscience ne peut ni être le niveau neuronal, ni être le niveau cognitif, mais doit plutôt être celui constitué du cerveau, du corps et de l'environnement de l'organisme.

Dans cette section, j'introduis d'abord la question des niveaux d'explication en distinguant le *niveau personnel* du *niveau subpersonnel* (3.3.1). Cette distinction est pertinente pour notre enquête, et le prochain chapitre rendra cela d'autant plus clair, car pour les éactivistes la conscience est une propriété se trouvant au niveau personnel, ne pouvant aucunement être réduite au niveau subpersonnel (Noë & Thompson 2004). J'y présente ensuite, dans les grandes lignes, le modèle de l'explication mécaniste (3.3.2). La question des niveaux d'explication concerne le rôle des neurosciences et des sciences biologiques pour la philosophie de l'esprit et est

étroitement liée à la question des niveaux d'organisations (e.g., on considère intuitivement que la physique des particules opère à un niveau inférieur que celui de la chimie moléculaire).

3.3.1 Niveau personnel et niveau subpersonnel

La distinction entre les niveaux d'explication *personnel* et *subpersonnel* a été introduite initialement par Daniel Dennett dans *Content and Consciousness* et a été reprise ensuite par plusieurs philosophes, dont les partisans de l'approche énaïve (1969 p. 90–6, voir aussi 1978, 2007). Dans les faits, tel que je l'ai dit dans l'introduction, celle-ci reflète une dichotomie traditionnelle que l'on peut reconnaître sous de nombreuses formes au sein des débats philosophiques : celle opposant le mental et le corporel, le dedans et le dehors, la sémantique et la syntaxe, l'espace des raisons et l'espace des causes, le vocabulaire intentionnel et le vocabulaire physique, etc. (c.f. Haugeland 1998). Cette séparation entre deux *domaines* d'entités justifie à son tour *deux entreprises explicatives distinctes*, que je vais présenter tour à tour : **a.** d'une part, nous avons les théories expliquant des phénomènes au niveau personnel et qui sont le propre des sciences humaines; **b.** d'autre part, les théories subpersonnelles relèvent des sciences comme les neurosciences, la biologie ou la physique. Enfin, **c.** je vais conclure cette sous-section avec quelques précisions sur la relation entre ces deux niveaux d'explication.

a. les théories personnelles : Les théories qui se situent au niveau personnel visent à expliquer le comportement d'agents rationnels *en tant que tels* en faisant référence à des états intentionnels (croyances, désirs, intentions, etc.) attribués à l'agent en tant que tel, et non pas à l'une ou l'autre de ses parties (e.g. le cerveau). Un trait distinctif de ce niveau de description est que les concepts utilisés par les théories personnelles, e.g. croyance, désir, douleur, conscience, etc., sont tirés du sens commun ou de la

psychologie populaire [*folk psychology*] et ne sont donc pas, initialement, des termes techniques issus de la science.

Traditionnellement, les théories se situant au niveau personnel évacuent totalement la question de l'*implémentation*, puisque pour qu'un système soit considéré comme étant intentionnel il suffit qu'on puisse prédire et expliquer son comportement en lui attribuant des croyances et des désirs. Pour ce faire, on adopte ce que Dennett appelle la posture intentionnelle²⁸ [*intentional stance*], c'est-à-dire qu'on considère que, tout comme nous, ce système a des intentions et des buts et qu'il obéit aux normes de la rationalité (on *assume* que l'agent est rationnel, mais cela n'exclut pas le fait qu'il puisse parfois agir irrationnellement). Dennett (1987) soutient d'ailleurs que cette insensibilité à l'implémentation est un des atouts majeurs de ces théories, puisque le même comportement intentionnel peut être produit par des systèmes réalisés dans des substrats très différents. Autrement dit, les mécanismes nous permettant d'entretenir des états intentionnels sont comme des boîtes noires, dont le détail de l'implémentation *pourrait être* déterminé par une théorie *subpersonnelle*. Alors que Dennett affirme que le niveau personnel ne peut être réduit au niveau subpersonnel (1987, essai 1), bien au contraire, il est par contre possible d'adopter cette même dichotomie tout en ayant une position réductionniste. En effet, l'objectif de Dennett est seulement d'expliquer comment un système auquel nous attribuons des états intentionnels peut aussi être susceptible d'une explication en des termes qui soient compatibles avec les sciences. Finalement deux conditions sont généralement acceptées pour distinguer les phénomènes au niveau personnel (e.g. une croyance) de ceux au niveau subpersonnel (e.g. une représentation dans le système visuel

²⁸ Il faut distinguer ici l'intentionnal stance, comme étant une stratégie prédictive et explative du comportement utilisée dans la vie de tous les jours et l'intentionnal stance en tant que théorie instrumentaliste à l'égard du statut d'existence des états intentionnels. Il est ici seulement question de la première.

primaire). En effet, un état personnel doit être à la fois accessible à la conscience et être intégré dans un réseau plus large d'inférences.

Dans le prochain chapitre, nous mettrons en lumière comment les énaïvistes récupèrent la distinction entre niveau personnel et niveau subpersonnel. Dans le contexte de cette discussion, il serait peut-être préférable de distinguer les *personnes* des *systèmes intentionnels* (Dennett expose sa conception des personnes dans 1976). Être un système intentionnel implique seulement d'avoir un comportement prédictible et explicable par l'attribution d'états intentionnels. En ce sens, les personnes, mais aussi les animaux, les plantes, les robots et les thermostats peuvent être considérés comme des systèmes intentionnels, *selon le besoin* (pour cette distinction, c.f. Dennett 1976). Enfin, à la défense des énaïvistes, mentionnons que Dennett lui-même confond parfois système intentionnel et personnes.

b. les théories subpersonnelles : Le rôle de la psychologie subpersonnelle, tel que je l'ai souligné précédemment, est d'expliquer comment des systèmes peuvent avoir un comportement intentionnel, comportement qui peut « à première vue semble[r] mystérieux et inexplicable » (Dennett 1987, p. 61). C'est ainsi que Dennett espère expliquer des phénomènes aussi « mystérieux » que la conscience ou l'*ego* :

Intuition, then, proclaims that any subpersonal theory must leave out something vital, something unobtainable moreover with subpersonal resources. [...] *I propose to construct a full-fledged "I" out of subpersonal parts* (1978, p. 206, je souligne).

Plus précisément, une théorie subpersonnelle consiste à identifier les mécanismes qui donnent lieu au comportement d'un système; comportement qui peut ensuite recevoir une interprétation au niveau personnel. Ainsi, alors qu'une théorie personnelle est tournée vers la relation entre l'agent et son environnement (social, naturel, etc.), une

théorie subpersonnelle est tournée vers l'intérieur du système, c'est-à-dire, qu'elle s'intéresse à l'interaction entre ses différentes composantes :

Subpersonal theories proceed by *analyzing a person into an organization of subsystems* (organs, routines, nerves, faculties, components-even atoms) and attempting to explain the behavior of the whole person as the outcome of the interaction of these subsystems. (1978, p. 205, je souligne).

On voit ici qu'une théorie subpersonnelle consiste à identifier le mécanisme qui *donne lieu* au phénomène qui nous intéresse; elle est donc formulée en termes *causaux* ou *fonctionnels*. Le pouvoir explicatif d'une telle théorie provient du fait que l'on tente *de décomposer le mécanisme qui nous intéresse en composantes* pour en comprendre le fonctionnement. Cette décomposition, Dennett l'illustre par la métaphore de *l'homoncule* (latin pour *homonculus*, petit homme): on progresse dans la compréhension d'un phénomène lorsqu'on décompose la personne (l'homoncule suprême, si l'on peut dire) en homoncules de plus en plus simples et de plus en plus spécialisés (1978, p. 216). Ultimement, on en arrive à des homoncules si simples, qu'il n'est plus nécessaire d'en postuler d'autres pour comprendre ce qu'ils font et comment ils le font.

Dans la théorie de Dennett, un même système peut être à la fois l'objet d'une théorie personnelle que d'une théorie subpersonnelle, puisqu'elles décrivent des phénomènes différents. Une théorie personnelle décrit la relation entre un agent rationnel et un état de chose du monde, comme la douleur, alors que les théories subpersonnelles expliquent le comportement des *parties* d'un agent, comme l'activité neuronale. Les premières sont tournées vers l'extérieur, alors que les secondes sont tournées vers l'intérieur. *Les entités postulées par ces théories sont aussi distinctes* : lorsque l'on passe du niveau personnel au niveau subpersonnel, on arrête de parler de la douleur et on s'intéresse plutôt à son implémentation neuronale. Là-dessus, Dennett ne pourrait être davantage explicite :

Abandonning the personal level of explanation is just that: *abandonning* the pains and not bringing them along to identify with some physical event. The only sort of explanation in which 'pain' belongs is non-mechanistic; hence no identification of pains or painful sensations with brain processes makes sense, and the physical, mechanistic explanation can proceed with no worries about the absence in the explanation of any talk about the discrimination of unanalysable qualities. (1969, p. 94).

Comme il le dit lui-même dans l'extrait, lorsqu'on passe au niveau subpersonnel, *l'aspect subjectif de la douleur est complètement délaissé*, puisque le cerveau lui-même ne peut ressentir de la douleur : seule une personne en est capable. Cela étant dit, pour Dennett, la relation entre le niveau personnel et le niveau subpersonnel n'est pas *causale* : les processus subpersonnel sont ce qui permet aux agents d'avoir les propriétés qu'on leur attribue au niveau personnel (des états intentionnels, etc.). En ce sens, chez Dennett, les niveaux personnel et subpersonnel sont bel et bien une manière d'appréhender une seule et même chose par l'entremise de deux descriptions différentes.

c. la relation entre niveaux et le réductionisme : J'ai mentionné plus haut que la distinction entre théorie personnelle et théorie subpersonnelle n'implique pas pour autant que la première soit réductible à la seconde. Pour Dennett, par exemple, la relation entre le niveau personnel et le niveau subpersonnel est *constitutive* et non *causale* : les processus subpersonnel constituent les parties d'un système qui donne lieu au phénomènes décrits par les théories au niveau personnel.

D'ailleurs, un des arguments célèbres contre le réductionisme est celui du *holisme du mental*, auquel Donald Davidson, avec son *principe de charité* (1980) et Daniel Dennett, avec sa présupposition de rationalité [*assumption of rationality*] (1987) souscrivent tous deux. Un des arguments principaux en faveur de cette thèse a été

proposé par Quine (1951), selon lequel nos croyances à propos du monde ne sont pas confirmées individuellement, mais seulement au sein des théories qui les englobent.

Plus précisément, le holisme du mental est là thèse selon laquelle les états mentaux ne peuvent exister isolément : l'identité de nos croyances, de nos désirs, de nos intentions et de nos perceptions étant déterminée par la place qu'ils occupent dans un réseau constitué par d'autres états intentionnels. Par exemple, si j'ai la croyance selon laquelle j'ai eu un accident de voiture hier soir, je dois aussi avoir plusieurs autres croyances relative à cet événement : sur les conditions routières à ce moment là, sur le modèle de ma voiture, sur l'heure à laquelle j'étais sur la route, etc.; de sorte que pour Davidson, la relation entre ces états mentaux est essentiellement *logique* : le contenu d'un état mental ne pouvant être tenu séparé de ce qu'il implique ni de ce par quoi il est impliqué. Comme on peut s'en douter, le holisme du mental demande que les états intentionnels attribués à un agent montrent un haut niveau de *cohérence* : je ne peux croire que *j'ai* eu un accident de voiture, tout en croyant ne pas savoir conduire.²⁹

3.3.2 L'explication mécaniste

Tel que le rapporte John Bickle (2003), l'intérêt des philosophes pour les neurosciences remonte à la publication de l'ouvrage célèbre de Patricia Churchland, *Neurophilosophy* (1986). Dans les mêmes années, certains philosophes des sciences commencèrent à s'intéresser, non plus exclusivement à l'explication en physique, mais

²⁹ Cependant, le holisme du mental est loin d'être une thèse qui fasse consensus. Jerry Fodor et Ernest LePore (1992), par exemple, ont sévèrement critiqué le holisme de Davidson, Dennett et Quine en faisant valoir que soutenir une thèse holiste à l'égard du contenu mental résulterait en l'impossibilité des sciences de l'esprit, que ce soit les sciences cognitives, les sciences sociales ou autre (ibid., p. 15). Selon eux, l'holisme du mental rend les généralisations en psychologie pratiquement impossibles, et donc, par extension, l'élaboration de lois psychologiques. En effet, si le contenu de n'importe quelle croyance d'un agent dépend sur celui de toutes ses autres croyances, alors il est pour ainsi dire impossible que deux agents partagent le même état avec le même contenu.

aussi à l'explication telle qu'on la retrouve dans les sciences de la vie (à savoir, la physiologie, la biologie cellulaire et moléculaire, la génétique, la biologie évolutionnaire et les neurosciences). Il se trouve, en effet, qu'au début des années quatre-vingt le modèle d'explication dominant était encore le modèle *déductif-nomologique* (DN), dont l'idée principale est que l'explication consiste à subsusmer un phénomène particulier sous des lois universelles :

A DN explanation answers the question "*Why* did the explanandum-phenomenon occur?" by showing that the phenomenon resulted from certain particular circumstances, specified in C1, C2, ..., C_k, in accordance with the laws L1, L2, ..., L_r. By pointing this out, the argument shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected*; and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred. (Hempel 1965, p. 337, l'auteur souligne).

Cette conception de l'explication a aussi été adoptée en philosophie de la psychologie, où la tâche du psychologue est conçue comme étant une recherche de lois. (e.g. Fodor 1974). Ce portrait ne correspond cependant pas à la pratique de ces sciences où il est rare que les chercheurs expliquent des phénomènes par l'entremise de lois.

Dans cette section, je commence par introduire la notion de mécanisme. Ceux-ci ne constituent pas l'explication en tant que telle, mais sont des entités réelles qui sont représentés par l'explication mécaniste

La notion de mécanisme. Le modèle d'explication mécaniste en philosophie des sciences avance que les phénomènes sont expliqués en *décomposant* les mécanismes qui en sont responsables. Qu'est-ce qu'un mécanisme ? Dans un des premiers ouvrages philosophiques consacrés à l'explication mécaniste, Bechtel et Richardson nous en offrent la définition suivante :

A *machine* is a composite of interrelated parts, each performing its own functions, that are combined in such a way that each contributes to producing a behavior of the system. A *mechanistic explanation* identifies these parts and their organization, showing how the behavior of the machine is a consequence of the parts and their organization". (2010, p. 17, je souligne).

En utilisant le terme « mécaniste », Bechtel et Richardson visent à souligner le fait que ce type d'explication conçoit le comportement des systèmes de manière analogue à celui d'une machine développée par la technologie humaine (ibid., p. 17). Plus précisément, cette citation met en évidence le fait qu'un mécanisme est une structure qui performe une fonction en vertu de ses *composantes* c'est-à-dire de ses *parties* (ou *entités*) et de ses *opérations* (ou *activités*), de même que de leur organisation; ce même mécanisme pouvant être responsable d'un ou plusieurs phénomènes distincts (Bechtel & Abrahamsen, 2005). Le coeur, par exemple, est responsable de pomper le sang dans le système circulatoire, mais il est aussi responsable d'autres phénomènes, comme le fait de faire un certain bruit (phénomène cependant beaucoup moins intéressant à étudier).

Les *parties* sont les *composantes physiques du mécanisme*. Ce sont elles qui encourent, donnent lieu, ou subissent les opérations et les activités au sein du mécanisme : pour reprendre notre exemple, le coeur est composé de diverses parties, dont des valves, des ventricules, des artères, des veines, etc.. Au sein du mécanisme, ce sont elles qui ont des propriétés physiques à proprement parler, comme une localisation, une taille, une masse, une structure et une orientation (Craver 2007, p. 5–6). Quant à elles, les *opérations* sont les *composantes causales*³⁰ du mécanisme et sont *productrices*, dans le sens où « elles ne sont pas de simples corrélations, ni de simples séquences temporelles et qu'elles peuvent être exploitées dans le but de

³⁰ Comme le souligne Craver (2007, p. 113), les chercheurs en neurosciences n'utilisent pas nécessairement le terme « cause », mais ils utilisent par contre une multitude de termes qui sont de nature causale et que l'on retrouve abondamment dans les manuels en sciences cognitives et en neurosciences : *activate, generate, cause, influence, control, inform, encode, inhibit, excite, modulate, filter, process, recognize, represent, regulate, store*.

manipuler et de contrôler » le mécanisme (ibid.). Chacune des opérations implique au moins une partie physique : « habituellement, il y a une partie active qui initie ou maintient l'opération (et peut être changée par elle), et au moins une partie passive qui est changée par l'opération. » (Bechtel 2005). Enfin, il ne faut pas penser que les différentes parties d'un mécanisme doivent partager la même localisation : dans le cas des mécanismes très complexes, comme ceux responsables de la cognition, les parties sont impossibles à distinguer des parties avoisinantes et sont distribuées largement dans le système, ce qui rend d'autant plus ardue leur localisation.

Comme l'indique Craver (ibid.), ces différentes composantes sont *organisées* « spatialement, temporellement, causalement et hiérarchiquement ». Cette organisation peut prendre la forme d'une séquence linéaire, mais les mécanismes biologiques, de même que ceux que l'on retrouve dans les sciences de l'esprit tendent à avoir une organisation plus complexe, non linéaire, où plusieurs niveaux sont impliqués.

Enfin, un mécanisme est *situé* dans un environnement particulier, et son comportement dépend des conditions environnementales. Le type d'information que peut obtenir un organisme à l'aide de son système perceptuel permet de guider la recherche lorsqu'on tente de découvrir le fonctionnement d'un mécanisme.

L'explication mécaniste. Cela étant dit, les mécanismes ne constituent pas en eux-mêmes l'explication; ceux-ci sont des systèmes réels que l'on retrouve dans la nature. Ce sont eux que nous devons expliquer. L'explication, quant à elle, est une *activité épistémique*, dont un aspect central consiste à déterminer qu'elles sont les composantes du système, de même que ce qu'elles font (ibid., p. 18). L'explication mécaniste repose sur deux *heuristiques* principales : la *décomposition* et la *localisation* (Bechtel & Richardson 2010). Par heuristique, on entend des règles rapides, intuitives et cognitivement économiques [*rules of thumb*] guidant nos

décisions et nos actions, tout en prenant en compte nos limites cognitives (e.g., le fait que notre mémoire, notre attention et notre patience soient limitées).

La première étape de l'explication mécaniste ne consiste pas à identifier le mécanisme, mais plutôt à bien *décrire* la phénomène à expliquer : un mécanisme est toujours le mécanisme d'un *phénomène donné*. Il y a ainsi plusieurs manières de découper la réalité, en fonction de ce qui nous intéresse : les frontières du mécanisme, c'est-à-dire, ce qui en fait partie et ce qui n'en fait pas partie sont déterminées en référence au phénomène à expliquer. Cela est d'autant plus important en neurosciences, où les diverses parties du cerveau sont en interactions étroites. Craver (2007, chap. 4) souligne l'importance d'une description adéquate du phénomène à expliquer : une erreur typique consiste à assumer que plusieurs phénomènes distincts sont en fait un seul et même phénomène. La caractérisation du phénomène n'est évidemment pas finale, puisque le phénomène pourra recevoir une caractérisation plus précise au fur et à mesure de l'avancement des recherches.

C'est aussi le choix du phénomène à expliquer qui va justifier à quel point il faut décomposer le mécanisme en composantes plus simples. Expliquer les mécanismes donnant lieu à la perception visuelle, par exemple, peut demander de décomposer le mécanisme jusqu'à la biologie moléculaire, mais il ne sera probablement pas nécessaire de remonter jusqu'aux particules atomiques. Dans le cas de la conscience, cette question est encore plus controversée : alors qu'aujourd'hui, plusieurs chercheurs affirment que l'explication de la conscience se fera au niveau neuronal, par la découverte de corrélats neuronaux, d'autres (e.g. Penrose 1989) ont défendu la thèse radicale selon laquelle il faut descendre à un niveau encore plus fondamental, celui des phénomènes quantiques³¹.

³¹ Pour un discussion critique et concise des enjeux métaphysiques liés à la physique quantique, c.f. Maudlin 2003.

Bechtel identifie deux types de décompositions : la *décomposition fonctionnelle* et la *décomposition structurelle*, dépendamment qu'on s'intéresse aux *opérations* ou aux *parties* du mécanismes (2005). Ces deux formes de décomposition peuvent se dérouler simultanément, même si en général se sont des chercheurs dans des domaines différents qui s'en occupent. La décomposition fonctionnelle consiste à partir du fonctionnement global du mécanisme et à le morceler en opérations plus simples qui pourraient donner lieu au comportement de celui-ci (e.g. Cummins 2000). Ce type de décomposition prend souvent la forme d'un diagramme [*box-and-arrow*] où l'organisation des diverses opérations est illustrée *indépendamment* de leur réalisation matérielle.

Dans le cadre de l'explication mécaniste, la décomposition fonctionnelle est insuffisante. En effet, ce modèle d'explication est *constitutif* : elle explique le comportement du mécanisme en faisant référence aux opérations effectuées *par les parties*, qui sont des *choses physiques réelles qui doivent ultimement être identifiées par les chercheurs*. Cela est dû au fait que l'approche mécaniste n'étudie pas des phénomènes abstraits, mais se penche plutôt sur des phénomènes particuliers et concrets, e.g. le rôle de l'aire V1 dans le fonctionnement du système visuel. Ce faisant, le but ultime est d'identifier par quelles parties sont réalisées les différentes opérations du mécanisme, c'est ce que l'on appelle la *localisation*.

Pour comprendre comment les parties et les opérations réussissent à donner lieu au phénomène que l'on cherche à comprendre il faut *recomposer* le mécanisme et déterminer l'organisation existant entre les diverses parties et opérations caractérisant le mécanisme dont il est question. Le flux linéaire de nos pensées peut nous inciter à croire que le monde obéit à ce même ordre séquentiel, mais avec le développement des neurosciences nous savons désormais qu'il n'en est rien : au sein du système visuel, par exemple, nous savons qu'il y a une certaine activité neuronale partant d'aires visuelles plus avancées (MT) et allant vers l'aire visuelle primaire (V1),

phénomène qu'on désigne par le nom de *feedback*. Même si le rôle de ce phénomène est controversé au sein du mécanisme donnant lieu à la conscience, son existence est largement reconnue aujourd'hui (c.f. Baars & Gage 2010, Bechtel 2009).

Par ailleurs, reconnaître l'existence de ce phénomène est encore plus facile lorsqu'on adopte une représentation pictoriale (un diagramme, un modèle) du mécanisme plutôt qu'une représentation uniquement linguistique comme c'est le cas pour les déductions logiques du modèle D-N. D'un point de vue purement cognitif, le simple fait de représenter un mécanisme dans l'espace permet de mieux saisir son fonctionnement. Cela dit, la représentation d'un mécanisme ne s'arrête au fait d'en présenter l'organisation spatiale, il faut aussi représenter les processus (i.e. les opérations) entre les différentes parties. L'opération pouvant être représentée par des flèches, celles-ci servant à indiquer l'interaction entre des composantes particulières. Enfin, alors qu'habituellement nous nous représentons le changement au sein d'un mécanisme en *simulant dans notre tête* l'interaction entre les composantes, aujourd'hui cette interaction peut être représentée par des graphiques animés à l'ordinateur.

3.3.3 Explications mécaniste, réductionnisme et niveau d'explication

Le modèle d'explication mécaniste tel que proposé par Bechtel, Richardson et Craver contraste autant avec ceux qui défendent l'*autonomie de la psychologie* (Fodor 1974), c'est-à-dire l'idée selon laquelle le mental est *irréductible* à d'autres niveaux d'explication, qu'avec ceux qui défendent une forme extrême de réductionnisme, tel John Bickle (2003), selon lequel l'explication aux niveaux neurobiologiques rend les niveaux d'ordre supérieur, comme le niveau psychologique, sans raison d'être et n'ayant, pour ainsi dire, qu'une valeur heuristique (c.f. *inter alia* chap. 3). Dans les prochains paragraphes j'espère expliciter ce contraste; l'explication mécaniste étant,

d'une certaine manière, à *mi-chemin* entre l'autonomie du mental d'un Jerry Fodor et le réductionnisme dévergondé [*ruthless*] d'un John Bickle³².

Ce que j'entends par là est que l'explication mécaniste est une forme de réductionnisme qui reconnaît la pertinence des différents niveaux explicatifs. Initialement, lorsque les philosophes commencèrent à s'intéresser à l'explication mécaniste, la réduction était principalement conçue comme étant *interthéorique*, où une théorie à la portée plus spécifique est réduite à une théorie plus générale (e.g. Nagel 1961 chap. 11, Churchland & Churchland 1990). Un exemple classique d'une telle réduction est celle de la réduction des trois lois de Kepler à la mécanique newtonienne.

La raison pour laquelle le réductionnisme donne parfois lieu à une certaine forme d'inquiétude est qu'on conçoit la réduction comme étant le fait de vouloir tout expliquer *par le niveau le plus fondamental* ou *par la théorie la plus générale*, qui est aussi souvent celui de la physique (ou dans certains cas, celui de la neurobiologie). Or, cette conception du réductionnisme présuppose que l'on puisse rendre compte complètement d'un phénomène en s'en tenant *uniquement*, soit à des lois de nature très générale, soit aux composantes ultimes de la réalité.

C'est précisément cette idée qui va à l'encontre de l'esprit de l'explication mécaniste tel que proposé par les auteurs mentionnés plus haut, idée selon laquelle différents types de relations causales se retrouvent dans les interactions entre différentes composantes, elles-mêmes se trouvant à divers niveaux explicatifs. Ainsi conçue, l'explication s'étend sur plusieurs niveaux et intègre les découvertes de divers champs d'études (Craver 2007, p. 2).

³² Dans les faits, il me faut préciser que l'approche favorisée par Bickle (2006), bien que réductionniste, est aussi explicitement mécaniste. Sa position métaphysique sur la relation interniveau est par contre différente de celle des auteurs mentionnés plus haut.

Comme nous l'avons vu, l'approche mécaniste implique de montrer comment le mécanisme donne lieu à certains phénomènes non seulement par l'interaction des différentes composantes, mais aussi par le fait qu'il est situé dans un contexte approprié. Elle est donc intrinsèquement *multileveliste*³³, le fardeau explicatif n'étant pas porté uniquement par les composantes des niveaux inférieurs.

En philosophie des sciences, la manière de distinguer les différents niveau d'explication est loin de faire consensus : on retrouve en compétition plusieurs conceptions différentes de ce qu'est un « niveau d'explication ». On peut d'abord concevoir les différents niveau à partir de la *taille* des entités concernées, nous avons ainsi les atomes d'un côté et le système solaire de l'autre. On peut aussi découper la réalité en niveau d'explication en se conformant à la séparation entre les différentes *disciplines scientifiques* : la physique et la chimie opèrent à des niveaux différents, puisqu'elles sont des disciplines différentes. (Craver 2007, Bechtel 2008, p. 144 et suiv.). Dans une perspective mécaniste, la relation entre niveau d'explication est *constitutive*, c'est-à-dire de nature méréologique (Craver 2007, Bechtel & Richardson 2010), la méréologie étant la théorie s'intéressant à la relation entre le tout et ses parties. Plus précisément, le comportement d'un mécanisme est conçu comme étant

³³ Du point de vue de la philosophie des sciences contemporaines, un des enjeux importants touchant l'approche éactive est celui de l'*émergence* et de la *causalité descendante* [*downward causation*]. Un phénomène est dit émergent lorsqu'il atteint un tel degré de complexité qu'il commence à manifester de nouvelles propriétés qui ne pas partagées par ses composantes. Certaines propriétés, comme la capacité à catégoriser ce qui est donné dans la perception, sont des propriétés qui ne sont pas possédées par les neurones elles-mêmes. C'est l'interaction entre ces neurones qui donne lieu à ce phénomène. La notion d'émergence est très populaire en philosophie et ailleurs, mais elle engendre plusieurs problème de nature conceptuelle. Certains, comme Kim (1999), ont fait valoir que nous ne pouvons séparer la notion d'émergence de celle de causalité descendante. Ainsi, Thompson (2007), par exemple dans son récent ouvrage, défend explicitement une forme d'*émergentisme* qui implique aussi de la causalité descendante. Pour une critique de la notion de causalité descendante, et par le fait même, de la notion d'émergence, voir les nombreux textes de Kim sur le sujet (1984, 1989, 1992, 1999). Bien que très intéressante, nous ne pouvons poursuivre plus loin cette piste réflexion pour des raisons d'espace.

« émergeant » dans la mesure ce qui donne lieu à celui-ci, c'est l'*organisation* entre les composantes et non pas contribution individuelle des différentes composantes.

3.4. Conclusion partielle : éaction et explication mécaniste

Rendre compte des conséquences de l'approche éactive pour l'explication en psychologie n'est pas chose aisée; pas facile, non plus, de fournir un argument qui irait dans la sens de la thèse de Prinz (2012), selon lequel l'éactivisme est incompatible avec la méthodologie scientifique la plus élémentaire. Je soutiens ici qu'à première vue, l'approche éactive n'est pas, en soi, incompatible avec l'explication mécaniste, mais qu'il se présente cependant un certain nombre de problèmes, auxquels les éactivistes doivent donner réponse. D'abord, l'approche éactive est antiréductionniste, alors que l'explication mécaniste adhère à une certaine forme de réductionnisme. En fait, pour les éactivistes, ce qu'on cherche expliquer se trouve au niveau personnel, mais ils doivent tout de même spécifier comment la recherche en neurosciences participent à l'explication de ces phénomènes.

Les philosophes qui s'intéressent à l'explication en neuroscience (e.g. Bechtel & Richardson 2010, Craver 2007) répondent explicitement à cette question : le pouvoir explicatif provient du fait qu'un mécanisme est décomposé en une hiérarchie de mécanismes plus simples, et ce, jusqu'au moment où le phénomène qui nous intéresse nous est intelligible. En soi, l'explication mécaniste n'exclut pas qu'un processus cognitif puisse être réalisé (i.e. localisé) partiellement dans l'environnement de l'agent. Cependant, ceux qui adhèrent à l'externalisme éactif doivent nous dire, précisément, comment nous pouvons localiser et décomposer ces mécanismes. La localisation et la décomposition étant les deux heuristiques de recherche principales de l'explication mécaniste, l'éactivisme sera incompatible avec la pratique en neuroscience tant que ses partisans n'auront pas dissipé les doutes sur cette question. Aussi, bien que les éactivistes ne soient pas hostiles aux neurosciences dans

l'explication de la perception (e.g. Thompson, Lutz, Cosmelli 2005), ils doivent tout de même expliciter le rôle qu'elles doivent jouer dans l'explication de celle-ci.

CHAPITRE IV

L'APPROCHE ÉNACTIVE ET LE PROBLÈME DES REPRÉSENTATIONS

Ce chapitre est divisé en cinq sections. Dans la première je fais une brève mise en contexte de l'approche énaactive, en insistant principalement sur la manière de concevoir la vision et la conscience en sciences cognitives et en neurosciences. Dans la deuxième section je présente ensuite l'approche énaactive de la perception, avant de poursuivre avec les conséquences qu'on peut en tirer pour l'expérience perceptuelle, dans la troisième section. La quatrième section portera sur l'antireprésentationnalisme énaactif. Finalement, dans la cinquième section je vais exposer les conséquences de l'approche énaactive pour l'explication de la conscience, de même que pour la distinction entre les diverses modalités sensorielles.

La majeure partie des théories de la vision que l'on retrouve dans les manuels en sciences cognitives ou en neurosciences présentent notre expérience visuelle comme étant *l'aboutissement* d'un ensemble de computations sur des représentations, processus se déroulant principalement dans les aires visuelles du cerveau. En conséquence, les énaactivistes affirment que la plupart des chercheurs en neurosciences souscrivent à la thèse du *substrat neuronal minimal*, selon laquelle « pour tout état conscient, il y a un substrat neuronal minimal » qui est *suffisant* pour donner lieu à chacun des aspects de cette expérience (Noë & Thompson 2004). Par ailleurs, les chercheurs en sciences cognitives, en intelligence artificielle et en robotique ont longtemps soutenu que la vision consistait à produire une représentation interne détaillée du monde extérieur, se modèle ayant le fonction de guider le comportement intelligent de l'organisme. Ainsi, à en croire cette théorie, si je suis

capable de m'orienter dans l'Université sans buter contre un mur, c'est parce que mon système visuel a construit un modèle de l'environnement dans lequel j'évolue quotidiennement. Cette conception de la vision et de l'expérience visuelle engendre deux problèmes : un problème « d'ingénierie » (a.) et un problème philosophique (b.)³⁴.

a. Le problème d'ingénierie : Le problème d'ingénierie est occasionné par une antinomie. D'un point de vue phénoménologique, notre système visuel semble nous offrir une expérience à la fois passive, uniforme, stable et riche en détail. Cependant, les données comportementales et neurologiques ne corroborent pas ce portrait de notre système perceptuel. En effet, l'information provenant de nos yeux est non seulement continuellement interrompue par les saccades visuelles et nos clignements d'yeux, mais en plus de cela, la répartition des photorécepteurs sur la rétine n'est pas uniforme, seule une petite région de la rétine (la fovéa) étant dotée d'une haute résolution. Il est donc vrai que « le monde que nous *percevons* [...], le monde dont nous sommes *conscients*, ne porte pas la marque de ses origines représentationnelles » (Akins & Winger 1996 p. 173). Malgré cet état de choses, comment notre cerveau en arrive-t-il à donner lieu à notre riche phénoménologie visuelle ? La réponse courante à cette question est de postuler l'existence de plusieurs *mécanismes de compensation*. Cela pose par contre certains problèmes, dans la mesure où ces mécanismes sembleraient computationnellement exigeants, et donc, demander un certain temps d'exécution dont ne disposent pas toujours les organismes qui doivent interagir en temps réel avec leur environnement, par exemple lorsqu'un animal doit fuir un prédateur (e.g. Ballard 1991).

³⁴ Cette conception de l'état des lieux en sciences cognitives est une prémisse, souvent implicite, justifiant les propos des énoncivistes (e.g. Pessoa et al. 1998, Thompson et al. 1999, Noë 2001, 2004 chap. 1). Dans leur critique adressée à la *Theory of pure vision*, Churchland, Ramachandran et Sejnowski (1994), s'attaquent à une cible similaire, mais en précisant toutefois que la *theory of pure vision* relève de la caricature. Les énoncivistes n'ont cependant pas une telle prudence. Dans les pages qui suivent je prends pour acquis que cela est un portrait fidèle de ce que pensent les scientifiques qui s'intéressent à la vision, sans m'attarder à vérifier si cela est effectivement le cas.

b. Le problème philosophique : Le problème philosophique, quant à lui, est celui du déficit d'explication touchant le caractère phénoménal de notre expérience, i.e. les qualia (Levine 1983, voir aussi Chalmers 1995, 1996) : il survient lorsqu'il est question de déterminer *pourquoi* cette activité neuronale *donne lieu* à une expérience particulière. Nous avons vu dans le première chapitre que l'intuition derrière l'argument de Levine était que même si nous en savons de plus en plus sur le fonctionnement du cerveau, intuitivement, il semble toujours y avoir un déficit explicatif entre notre expérience et les représentations et computations internes qui sont censées en être responsables (c.f. le deuxième chapitre). En d'autres mots, il semble « manquer un ingrédient » nous permettant de rendre compte de cette identité psychophysique. La manière dont Levine, Chalmers et bien d'autres formulent le problème de la conscience fait en sorte que la solution à celui-ci, s'il en est, devra être conceptuelle. Comme nous le verrons plus loin, bien que les énaïvistes refusent d'aborder le problème dans ces termes, ils sont d'accord avec Levine lorsqu'il affirme que le programme de recherche visant à trouver les corrélats neuronaux de la conscience ne pourrait aucunement combler le déficit d'explication (Hurley & Noë 2003, p. 160–1).

Dans ce chapitre, je présenterai les motivations théoriques et empiriques motivant l'approche énaïve de la *perception*³⁵ et du *contenu perceptuel* telle que défendue principalement dans O'Regan et Noë 2001 et Noë 2004 en me concentrant sur un aspect central de leur théorie : leur *antireprésentationnalisme*. Celui-ci est justifié par leur *posture sensorimotrice*, qui se veut d'abord et avant tout une solution au problème philosophique des qualia, mais aussi au problème d'ingénierie.

³⁵ Les énaïvistes utilisent souvent les termes « vision », « perception », « contenu perceptuel », voire même « conscience », de manière interchangeable, ce qui semble indiquer que leur stratégie devrait avoir une portée globale. Pour rester clair, je m'en tiendrai ici seulement aux termes de « perception » et « contenu perceptuel ». Bien que les énaïvistes s'intéressent principalement à la vision et au toucher, leur thèse doit minimalement s'appliquer à toutes les modalités sensorielles, faute de quoi elle sera inconsistante.

L'énactivisme propose une conception alternative de notre expérience perceptuelle, selon laquelle voir est une *manière d'agir* qui s'apparente au toucher. Par cette théorie, les énéactivistes ne font pas que remettre en question la séparation traditionnelle entre perception et action, mais aussi et surtout le recours aux représentations dans l'explication de la vision et du contenu perceptuel. En effet, les énéactivistes font valoir que pour voir, il n'est pas nécessaire que le cerveau de l'organisme se construise une représentation détaillée du monde extérieur, car l'organisme *sait* comment aller chercher les détails dans son environnement immédiat lorsqu'il en a besoin.

D'un point de vue explicatif, le coeur de la stratégie énéactive consiste ainsi à mettre l'accent sur un type bien précis de connaissances, les *connaissances des contingences sensorimotrices*, plutôt que sur des représentations internes (dans ce mémoire, un état mental est *interne* si les états neuronaux et corporels suffisent pour y donner lieu). Les contingences sensorimotrices étant, en bref, la relation nomique entre, d'une part, le mouvement de l'organisme percevant et, d'autre part, la stimulation sensorielle; les énéactivistes défendant que ces contingences sensorimotrices sont précisément ce en quoi consiste la différence entre les diverses modalités sensorielles : la vision et le toucher étant caractérisés par des contingences sensorimotrices différentes (pour une discussion philosophique intéressante sur la relation entre le toucher et la vision, c.f. O'Shaughnessy 1989, Martin 1992).

À de nombreux égards, cette conception active de la vision ressemble beaucoup à la théorie de la « vision animée » proposée par Dana Ballard dix ans plus tôt, de même qu'aux théories de Brooks (1991a, b) et de Dennett (1991), sans parler d'autres influences plus lointaines, comme celle de MacKay (1962, 1967), Gibson (1979), voire même de Merleau-Ponty (1945). À ce point, il est justifié de se demander en quoi consiste l'originalité de l'approche énéactive. En présentant l'approche énéactive, je

soutiens que cette approche est d'abord et avant tout une théorie du contenu perceptuel, qui consiste à adopter une stratégie similaire à celle de Dennett³⁶ à l'égard de la conscience (i.e. des qualia) tout en allant plus loin : les processus qui donnent lieu à la conscience sont non seulement distribués dans le cerveau, ils sont aussi distribués dans le corps et dans l'environnement de l'organisme percevant.

Avec sa théorie connue sous le nom de *Multiple Draft Model*, Dennett défend la thèse selon laquelle la conscience est à la fois distribuée spatialement et temporellement dans le cerveau (c.f. Dennett 1991 et Dennett et Kinsbourne 1992). Autrement dit, peu importe ce que nous savons sur le fonctionnement du cerveau, il sera toujours impossible de trouver un point précis dans celui-ci où un contenu particulier devient conscient. Ce faisant, on diminue aussi le fardeau explicatif des théories de la conscience, puisqu'il n'est plus nécessaire de trouver en quel endroit les représentations mentales deviennent conscientes. S'il en est ainsi, selon Dennett, c'est parce que la conscience est *constituée* uniquement par l'activité de différents processus neuronaux qui fonctionnent en parallèle et que rien de plus n'est nécessaire pour rendre un contenu conscient.

C'est précisément cette thèse que remettent en cause les énaïvistes : la conscience n'est pas quelque chose qui se produit dans le cerveau, mais plutôt quelque chose que l'organisme *fait*. Par conséquent, l'originalité de la stratégie adoptée par les énaïvistes consiste à mettre l'accent sur *l'utilisation* des connaissances des contingences sensorimotrices plutôt que sur des états, des processus ou des propriétés internes aux organismes dans le cadre du problème de l'expérience perceptuelle.

³⁶ Pour une discussion intéressante de la théorie de Dennett dans une perspective énaïve, c.f. Pessoa et al. 1998, p. 744–6, de même que O'Regan et Noë 2001, p. 965. Voir aussi Dennett 2001, p. 225–6

Une opposition qui vient rapidement à l'esprit est celle voulant que les *connaissances* des contingences sensorimotrices soient des représentations motrices, et que par conséquent, elles soient des représentations internes au même titre que les autres représentations postulées par les sciences cognitives traditionnelles (voir le deuxième chapitre pour la notion de représentation mentale). On peut cependant répondre à cette opposition en affirmant que les connaissances sont des propriétés de l'organisme en tant que tel, qu'elles sont attribuées au niveau personnel, et donc, ne sont pas des propriétés internes à proprement parler.

Par ailleurs, comme plusieurs chercheurs en sciences cognitives qui adhèrent aux théories incarnées de la cognition, les énaïvistes proposent une critique radicale de la distinction entre systèmes perceptuels et systèmes moteurs. Cela dit, contrairement à ce que laisse penser Jesse Prinz (2006 p. 8, 2012), le point central de l'approche énaïve ne peut se réduire à une variante de la thèse de l'esprit incarnée selon laquelle l'expérience perceptuelle est partiellement constituée par des représentations dans le système moteur, par opposition à des représentations du système visuel uniquement. L'idée principale de l'approche énaïve est plutôt que « le substrat de la conscience – en particulier de la conscience perceptuelle visuelle – semble ne pas se fonder à la division cerveau-corps-monde. » (Noë & Thompson 2004, p. 26, voir aussi Noë 2001 p. 47). C'est précisément la raison pour laquelle, ils ne fournissent pratiquement jamais de données neurologiques pour asseoir leur thèse : au final, il semble que contrairement à ce qu'ils disent parfois, leur argument n'est pas de nature empirique, mais est plutôt de nature conceptuelle c'est-à-dire qu'il vise à poser un regard nouveau sur la conscience et la manière de l'expliquer (ils sont, sur ce point, fidèles à l'approche initiée par Varela, Thompson et Rosch en 1991).

Conformément aux interprétations de Rowlands (2003), Block (2005), Prinz (2006, 2012) et Clark (2008), j'interprète donc l'approche énaïve comme étant une *thèse métaphysique* sur le statut du *contenu de notre expérience perceptuelle* plutôt qu'une

théorie de la perception en tant que telle, ce qui semble conforme aux propos que O'Regan et Noë tiennent dans leur article de 2001 (malgré une forte emphase sur l'expérience visuelle). De plus, je justifie cette interprétation de la manière suivante : d'une part, l'approche énative ne peut être seulement une théorie vision, car par définition elle doit pouvoir s'appliquer à toutes les modalités sensorielles (Hurley & Noë 2003, Noë 2004), d'autre part, sauf exception (Noë 2004, chap. 3), les énativistes sont très rarement explicites sur le problème traditionnel de la perception tel qu'on le retrouve dans l'histoire de la philosophie depuis Descartes (au moins), sinon pour rejeter la théorie des *sense-data*, théorie qui est aujourd'hui passée de mode (par contre, pour une défense récente, voir Robinson 1994). Par « problème traditionnel de la perception », je fais référence au débat philosophique classique visant à déterminer quel est *l'objet direct ou immédiat* de la perception. Comme le dit si bien J. J. Valberg (1992), ce problème prend la forme d'une antinomie. D'une part, la perception semble être une ouverture sur le monde : en ce sens, ce que nous percevons, ce sont des objets, des propriétés ou des faits du monde. D'autre part, si nous suivons un certain raisonnement à propos des erreurs perceptuelles (i.e. à propos des illusions et des hallucinations), nous sommes portés à conclure que l'objet de la perception n'est pas une partie du monde, c'est-à-dire un objet externe, mais plutôt un objet mental comme les *sense-data*.

Bien que O'Regan et Noë se prononce rarement sur cet enjeu, nous pouvons cependant douter que l'approche énative soit complètement neutre à cet égard, dans la mesure où elle a très certainement quelque chose à dire sur la nature de l'objet de la perception. Une discussion portant sur la relation entre l'énativisme et le problème classique de la perception serait intéressante, mais sort malheureusement du cadre fixé ici.

4.1. Les approches classiques dans les sciences de la vision et de la conscience

Mon objectif dans cette section est de mettre en contexte la théorie éactive en la mettant en contraste avec des approches auxquelles elle s'oppose. Pour ce faire, il sera d'abord question du modèle classique de la vision. Nous passerons ensuite au programme de recherche visant à trouver des corrélats neuronaux à la conscience. Bien que ces enjeux soient distincts, l'approche éactive présente dans chacun de ces cas la même stratégie, qui consiste à mettre l'accent sur les connaissances sensorimotrices et leur utilisation, plutôt que sur certains processus, états ou propriétés internes à l'agent (respectivement des computations, des représentations et des propriétés phénoménales), qui faisaient jusqu'à présent obstacle à une explication satisfaisante et complète. Ici, l'utilisation du concept de représentation est particulièrement critiquée par les éactivistes. Cela dit, nous verrons dans les sections subséquentes que ce que les éactivistes rejettent ce n'est pas la notion générale de représentation mentale, mais plutôt l'idée que le système visuel sert à construire un modèle détaillé du monde extérieur guidant le comportement de l'agent. Après tout, postuler des connaissances sensorimotrices semble demander une forme ou une autre de représentations mentales. Ainsi, ce que les éactivistes font, au final, c'est donner un rôle plus modeste aux représentations dans l'explication de phénomènes cognitifs plutôt que de les éliminer totalement. (C'est précisément sur ce point que l'éactivisme–sensorimoteur de Noë, O'Regan et Hurley, dont nous discutons ici, doit être distingué de l'éactivisme radical³⁷ tel qu'on le retrouve dans Hutto et Myin 2012).

³⁷ Il est intéressant de constater que Hutto et Myin adoptent exactement la même stratégie que O'Regan et Noë : soutenant que la notion de représentation mentale donne lieu à la majorité des problèmes en philosophie de l'esprit, les auteurs proposent d'abandonner complètement cette notion.

4.1.1 L'internalisme et la science de la vision

La meilleure façon de saisir en quoi consiste l'approche éactive est de la mettre vis-à-vis la thèse à laquelle elle s'oppose, à savoir la *conception photographique* [*snapshot conception*] de la vision, thèse qui est aussi une certaine forme de représentationnalisme. Selon celle-ci, nous sommes aujourd'hui spontanément portés à croire que « l'oeil est comme une caméra et que la vision est un processus quasi photographique » (Noë 2004, p. 36–7). Par là, on entend que notre champ visuel est comme une *image*, aussi détaillée en son centre que dans sa périphérie, exactement comme peuvent l'être les photographies.

Nous constatons à quel point nous adhérons spontanément à la conception photographique dans la vie de tous les jours, lorsque nous sommes mis face aux limitations et aux « défauts » de notre système visuel. Dans les prochains paragraphes, je vais aborder deux cas où ces limitations sont particulièrement saillantes : celui de la *cécité au changement* et de la *cécité inattentionnelle* (a.) et celui des saccades visuelles (b.) :

a. La cécité au changement : La cécité au changement a été mise en évidence dans plusieurs travaux, notamment dans ceux de Simons et Chabris (1999). Ces derniers empruntent un protocole expérimental d'abord développé par Ulrich Neisser pour démontrer l'importance de l'attention pour la perception. Dans ce protocole, les sujets s'engagent dans une tâche cognitive continue qui leur demande de se concentrer sur un aspect particulier de la scène, tout en ignorant les autres aspects de celle-ci. Cependant, à un certain point pendant la tâche, un événement inattendu que la majorité des sujets ne peuvent rapporter se produit, et ce, même si les observateurs qui ne sont pas engagés activement dans la tâche peuvent le constater. Dans le cas de Simon et Chabris, un vidéo montrant six personnes (trois ayant un dossard blanc, trois ayant un dossard noir) jouant au basketball dans une aire confinée. Les auteurs

demandent aux sujets de compter le nombre de passes entre les joueurs de l'équipe au dossard blanc; la tâche étant difficile, les sujets doivent se concentrer. Pendant ce temps-là, en plein milieu de la tâche, une personne déguisée en gorille marche en plein milieu du groupe de joueurs, s'arrête, se bombe le torse et finalement sort de l'aire de jeu. (Pour une discussion plus approfondie de ce phénomène, c.f. Mack & Rock 1998).

b. les saccades visuelles : Les travaux de John Grimes (1992), sont aussi très éloquentes. Grimes a démontré en laboratoire que nous sommes littéralement aveugles pendant les saccades visuelles, à un point tel que nous sommes incapables de détecter des modifications flagrantes effectuées sur une image lorsque le changement en question se déroulent lors d'une saccade visuelle. Nous savons aujourd'hui que la vision est un processus cognitif complexe nécessitant l'interaction de nombreuses composantes. La rétine, où se trouvent les photorécepteurs dans l'oeil, est une des premières composantes impliquées dans ce processus; les photorécepteurs ne sont cependant pas distribués de manière uniforme sur celle-ci. En effet, la fovéa est la région relativement restreinte où la résolution est la meilleure, c'est-à-dire la région la plus sensible aux menus détails. Pour cette raison, l'organisation du système visuel des primates est faite en sorte qu'afin de percevoir une scène visuelle dans le détail, il est nécessaire d'orienter la fovéa vers diverses parties de la scène de manière séquentielle. C'est ici que les saccades visuelles entrent en jeu : les primates déplacent leurs yeux d'un lieu à l'autre par l'entremise de mouvements très rapides appelés *saccades* (ibid., p. 90).

Je reviendrai sur ces deux cas dans les prochaines sections, car ils occupent une place non négligeable dans les arguments en faveur de l'approche énactive.

La conception photographique de la vision, en conformité avec notre phénoménologie du sens commun, s'oppose cependant à un état de choses généralement accepté par les

scientifiques de la vision : notre système visuel est imparfait et l'*input* sensoriel provenant de nos rétines est pauvre. En effet, les exemples des « défauts » qui caractérisent notre système visuel sont nombreux³⁸, le cas de la tache aveugle [*blind spot*] étant probablement celui le plus célèbre. Pour saisir ce phénomène, un exemple, peut-être imparfait, consiste à comparer la tache aveugle à l'angle mort lors de la conduite automobile : pour le conducteur, la *tache aveugle* (l'angle mort) est la région qu'il ne peut percevoir derrière la voiture avec aucun de ses rétroviseurs (Baars & Gage 2010). Similairement, nous avons une tache aveugle dans chaque oeil : dans la région où le nerf optique quitte la rétine, nous n'avons aucun photorécepteur.

Tout cela ne fait qu'accentuer la disparité entre la richesse de notre expérience perceptuelle et le détail de nos représentations : de ces considérations sur la nature de notre système visuel « le problème qui émerge pour une théorie de la vision consiste à expliquer l'apparente continuité de l'expérience perceptuelle malgré sa discontinuité actuelle. » (Noë 2001, p. 44).

Autrement dit, cela soulève une question à laquelle tentent de répondre les scientifiques qui oeuvrent dans le domaine de la vision : comment avec des moyens aussi limités, notre système visuel en arrive-t-il à créer une expérience du monde qui nous semble si nette et détaillée ? Les chercheurs oeuvrant dans le domaine de la vision ont habituellement présupposé que celle-ci dépendait d'une représentation interne et qu'il existait un *mécanisme de compensation* (voire plusieurs); ils ont donc développé différentes méthodes afin de trouver et de mesurer le travail de ces mécanismes. Pour expliquer le fait que nous ne voyons pas la tache aveugle, par exemple, un nombre important de scientifiques s'intéressant à la vision ont postulé des mécanismes de *complétion visuelle*. En général, les scientifiques qui oeuvrent dans le domaine de la vision utilisent les termes de *complétion visuelle* ou de

³⁸ Pour un bon aperçu des divers « problèmes » ou « défauts » touchant notre système visuel, voir O'Regan 2011 chap. 1.

remplissage [filling-in] pour faire référence à des situations où des sujets affirment que quelque chose est présent dans une région de leur champ visuel, alors que dans les faits (i.e. physiquement), cette chose en est absente. Notons cependant que bien que dans la littérature la tache aveugle soit considérée comme un cas paradigmatique de complétion visuelle (celle-ci se produit aussi dans le cas de scotome), ce choix est controversé³⁹.

L'exemple classique de théorie de la vision qui souscrit à la conception photographique est la théorie mise de l'avant par David Marr (1982). Celui-ci a développé la première *théorie computationnelle* de la vision, qui fut aussi une des théories de la vision les plus influentes au cours des trente dernières années, notamment chez les philosophes. Trouvant son origine dans la théorie de l'information, une théorie computationnelle de la vision vise non seulement à déterminer quelles informations sont extraites de l'image rétinienne, mais aussi, et surtout, à déterminer comment cette information est utilisée et quelles computations opèrent sur elle (Ballard 1994, p. 111).

Selon Churchland, Ramachandran et Sejnowski (1994), la théorie de Marr se conforme à trois présupposés théoriques qui forment ce qu'ils appellent la *Theory of pure vision* (bien que ces auteurs décrivent la théorie de la vision pure comme étant une « caricature », la théorie élaborée par David Marr s'y conforme fortement). D'abord, que ce que nous percevons à chaque instant implique une *représentation élaborée* du champ visuel et que le but de la vision est de construire dans le cerveau un *modèle détaillé* du monde extérieur tel qu'il se présente à nos yeux. Ensuite, que la

³⁹ Certains chercheurs ont critiqué l'utilisation de la tache aveugle comme exemple type de la complétion visuelle, puisqu'il y a encore des controverses quant à la description phénoménologique de ce phénomène. En effet, la tache aveugle se trouve à la périphérie de notre champ visuel, là où notre acuité est la plus limitée (Davis et Driver 2003, p. 128-9). Pour une discussion plus approfondie sur le phénomène de la complétion visuelle, voir le volume édité sur sujet par Pessoa et de Weerd (2003).

vision est un processus qui se déroule en *diverses étapes*, dont le résultat principal est d'extraire et d'intégrer une quantité grandissante d'information sur divers aspects qui se présentent dans le champ visuel, et ce, jusqu'à ce que le système visuel puisse produire une représentation interne. Enfin, que les étapes les plus avancées dans le processus de traitement d'information *dépendent* des niveaux les moins avancés, sans que l'inverse ne soit vrai; le système visuel ayant donc une la forme d'une *structure séquentielle linéaire*.

Conformément à cela, Marr postule trois étapes⁴⁰ de représentations du monde visuel qui sont créées et manipulées par le processus donnant lieu à la vision : l'esquisse primaire [*primal sketch*], l'esquisse 2½D et le modèle 3D. Plus précisément, Marr conçoit la vision comme étant un processus débutant avec l'image rendue disponible par la rétine, que l'on peut concevoir, par analogie, comme étant une matrice de pixels, à partir de laquelle est *inféré* ou *construite* un ensemble de *représentations internes de plus en plus et détaillées*; le but de ce processus étant d'extraire les propriétés objectives relatives à la forme géométrique des objets perçus.

De cette matrice de pixels, certains processus computationnels produisent l'*esquisse primaire* qui « rend explicite de l'information importante à propos de l'image à deux dimensions, principalement les changements d'intensité » au sein de celle-ci de même que « leur distribution et organisation géométrique » (ibid., p. 37). Le principe opérant ici est que ces changements d'intensité correspondent aux arêtes des objets présents dans le champ visuel.

L'*esquisse 2½D*, quant à elle, « rend explicite l'orientation et la profondeur des surfaces visibles » de même que les discontinuités dans leurs contours à partir des informations sur les arêtes fournies par l'esquisse primaire (ibid.). Autant dans le cas

⁴⁰ À ne pas confondre avec les trois niveaux d'explications (computationnel, algorithmique et implémentation) que Marr propose dans le même ouvrage.

de l'esquisse primaire que de l'esquisse 2½D la perspective est du point de vue de l'observateur.

Enfin, le modèle 3D « décrit les formes et leur organisation spatiale » en adoptant un point de vue impersonnel, et donc, centré sur les objets plutôt que sur les perspectives visibles du point de vue de l'observateur.

Bien que la théorie de Marr eu un certain succès, elle reste silencieuse sur plusieurs enjeux importants, comme certains de ses détracteurs l'on remarqué. D'abord, du point de vue de l'organisme, sa conception de la vision est fondamentalement *passive* ou *contemplative*, malgré le fait que les computations internes soient des *activités*. Cela dit, Marr évacue complètement le contexte comportemental de l'organisme dans son environnement, de même que son importance pour la vision. Ensuite, cette théorie ne tient aucunement compte des aspects caractéristiques du système visuel humain, notamment en ce qui concerne les saccades visuelles et de la fovéa, là où la résolution est la meilleure (Ballard 1996, p. 112 et suiv.).

4.1.2 Internalisme et conscience : la recherche des corrélats neuronaux

Dans le deuxième chapitre de ce mémoire, j'ai présenté à grands traits la théorie incarnée de la perception, de même que les critiques généralement adressées au modèle traditionnel. On y a aussi vu que la majorité des philosophes tentent d'expliquer la conscience en la réduisant à des représentations mentales. En neurosciences, les chercheurs ont principalement eu comme objectif la recherche de *corrélats neuronaux de la conscience*, c'est-à-dire l'activité neuronale minimale suffisante pour donner lieu à telle ou telle expérience. Cette section vise à présenter brièvement les présupposés « internalistes » derrière ce programme de recherche.

Tel que je l'ai déjà dit, concernant l'expérience perceptuelle en tant que telle, l'approche énactive s'oppose à la conception commune selon laquelle le caractère phénoménal de notre expérience est compris comme étant *l'occurrence de quelque chose dans notre cerveau* (O'Regan & Noë 2001, p. 960). La conception photographique de la vision et cette conception de l'expérience vont de pair, car dans la première la vision est un processus où une représentation est construite, alors que dans la seconde l'expérience est conçue comme étant le résultat de ce même processus : une représentation dans notre cerveau où les différentes propriétés de notre expérience correspondent à l'activation de certains neurones.

Dans son article maintenant célèbre, David Chalmers définit les corrélats neuronaux (NCC) de la manière suivante :

An NCC (for content) is a minimal neural representational system N such that representation of a content in N is sufficient, under conditions C, for representation of that content in consciousness. (2000, p. 31).

Déterminer quels sont les corrélats neuronaux de la conscience n'est cependant pas chose facile. Comme le remarque Chalmers (1998), on retrouve dans la littérature au moins une vingtaine de propositions de corrélats neuronaux. Malgré leurs différences, ces diverses théories partagent toutes un but commun : expliquer comment la structure de l'activité neuronale peut donner lieu aux multiples aspects de notre expérience consciente. Pour se faire, l'heuristique guidant la recherche en neurosciences consiste à trouver une structure neuronale isomorphe à la structure de notre expérience.

Dans un article visant à faire valoir que la recherche de NCC est voué à l'échec, Noë et Thompson (2004a, b) prétendent que la vaste majorité des théories de la conscience en neurosciences s'inscrivant dans ce programme de recherche reposent à la fois sur

deux présupposés théoriques important : celui de l'*isomorphisme analytique*, et celui du *bridge locus*. Je vais aborder tour à tour chacun de ceux-ci.

isomorphisme analytique : *L'isomorphisme analytique* « est essentiellement une doctrine conceptuelle ou méthodologique à propos de la forme d'explication adéquate en neurosciences cognitives. » (Pessoa et al 1998, p. 726). Cette doctrine concerne la nature des corrélats neuronaux de la conscience, puisqu'elle implique qu'il doit y avoir une *correspondance structurelle* une à une entre [*one-one mapping*] entre la structure de notre expérience à celle du substrat neuronal. Autrement dit, chaque aspect de notre expérience doit être en relation isomorphe avec une activité neuronale bien précise dans le cerveau. Bien que cet isomorphisme ne doivent pas nécessairement être spatial ou topographique, les deux manières les plus intuitives de concevoir l'isomorphisme, il est reste néanmoins primordial à l'explication en neurosciences. En effet, sans une telle similarité structurelle entre l'activité neuronale et l'expérience perceptuelle, leur relation serait tout à fait mystérieuse.

Une conception naïve de cet isomorphisme serait de dire que les propriétés du monde dont nous faisons l'expérience sont représentées par des propriétés similaires dans le cerveau, e.g. si j'ai une expérience d'un objet rouge, alors dans mon cerveau il y a quelque chose de rouge (Dennett, par exemple, tourne en dérision cette idée, lorsqu'il parle de « peinture mentale » ou de *figment*). Une conception moins naïve consisterait à dire que les propriétés des objets sont représentés par une activité neuronale ayant certaines similarités, comme une figure spatialement étendue dans le cortex visuel. Une analogie peut s'avérer ici éclairante : les couleurs de l'image qui se trouvent sur l'écran de mon ordinateur correspondent non pas à des couleurs dans la mémoire celui-ci, mais plutôt à une matrice de pixel. Ces derniers sont plutôt emmagasinés dans la mémoire sous forme numérique. Il va sans dire que la majorité des chercheurs en sciences cognitives n'adhèrent pas explicitement à ces idées, et Dennett le sait très bien. Cela dit, ces présupposés théoriques continuent d'inspirer notre façon de

concevoir le système visuel de même que notre expérience visuelle (cette discussion est partiellement inspirée de Akins & Winger 1992, p. 175–9).

En rejetant l'isomorphisme analytique, les énéactivistes admettent, comme Dennett, qu'il ne faut pas accepter naïvement le portrait que nous dresse l'introspection de notre expérience perceptuelle : il est impossible d'inférer quelque chose sur la nature du contenu neuronal à partir de la structure de notre contenu perceptuel, notamment parce que notre expérience est stable et séquentielle, alors que les processus qui en sont responsables sont instables et opèrent en parallèle (Dennett 1991, Dennett & Kinsbourne 1992, voir aussi Hurley 1998, introduction).

bridge locus : La seconde idée est celle selon laquelle il y aurait un *bridge locus*, et que par conséquent, dans l'étude de l'expérience perceptuelle nous devrions identifier une étape finale dans le cerveau où il y a un tel isomorphisme entre le contenu de l'expérience du sujet et l'activité neuronale :

Most visual scientists probably believe that there exists a set of neurons with visual system input, whose activities form the immediate substrate of visual perception. We single out this one particular neural stage, with a name: the bridge locus. The occurrence of a particular activity pattern in these bridge locus neurons is necessary for the occurrence of a particular perceptual state; neural activity elsewhere in the visual system is not necessary. The physical location of these neurons in the brain is of course unknown. However, we feel that most visual scientists would agree that they are certainly not in the retina. For if one could set up conditions for properly stimulating them in the absence of the retina, the correlated perceptual state would presumably occur (Cité dans Hurley & Noë 2003, note 6).

Ce *bridge locus* peut être comparé à la ligne d'arrivée d'une course : avant d'atteindre ce point, les représentations neuronales sont inconscientes, mais une fois passée la ligne d'arrivée elles deviennent conscientes. Autrement dit, les représentations qui parviennent au *bridge locus* serait suffisante pour donner lieu à l'expérience, peu

importe l'activité que l'on retrouve ailleurs dans le système visuel, par exemple, par l'entremise de stimulation sensorielle.

La popularité des protocoles expérimentaux reposant sur la *rivalité binoculaire* démontre bien à quel point les chercheurs en neurosciences s'intéressent à la conscience souscrivent à ces présupposés théoriques. La rivalité binoculaire est un phénomène visuel qui se produit lorsque des stimuli différents sont présentés sur les aires rétiniennes correspondantes des deux yeux. Plutôt que d'avoir une perception stable qui amalgame les deux stimuli, le sujet fait l'expérience des deux stimuli en alternance. L'intérêt pour ce type de protocoles expérimentaux a été popularisés par les travaux que Logothetis et ses collègues ont effectué avec des macaques entraînés à peser sur un bouton lorsqu'ils encourent une expérience particulière (Logothetis & Schall 1989, Leopold & Logothetis 1996, Leopold *et al.* 2003). Avec ces protocoles expérimentaux, Logothetis et cie., semblent nous donner les outils pour dissocier, d'une part, l'activité neuronale suscitée par le stimulus (sur la rétine), et d'autre part, l'activité neuronale correspondant à l'expérience vécue actuellement par le sujet. En effet, l'hypothèse justifiant cette méthode est l'idée qu'il y a une forme ou une autre d'isomorphisme structurel entre l'expérience vécue par le sujet et l'activité neuronale, si tant est qu'on accorde que des animaux puissent être dotés de conscience phénoménale, les termes même du débat étant loin de faire consensus. Certains philosophes, comme Daniel Dennett (1991) ont refusé d'attribuer des états phénoménaux aux animaux, notamment parce que ceux-ci ne sont pas dotés de langage.

Ceux qui oeuvrent en neurosciences considèrent généralement que de trouver de telles corrélations entre l'activité neuronale et l'expérience est la première étape vers la résolution du problème de la conscience, puisqu'elle permettrait supposément de *localiser* précisément les mécanismes responsables de l'expérience perceptuelle dans certaines aires cerveau. Ce travail de localisation a, par exemple, été marqué d'un

franc succès dans le cadre de l'étude neurologique du système visuel chez les primates, où nous savons de mieux en mieux quelles opérations sont effectuées dans des aires spécifiques du cerveau (Bechtel 2009). Ce type d'étude, par exemple, permet d'établir clairement que les aires primaire du système visuel (V1) ne correspondent pas au contenu de notre expérience visuelle.

Comme nous allons le voir dans les sections subséquentes, malgré le succès des travaux de Logothetis et de ses collègues, il n'est pas certains que ces corrélations permettent de combler le déficit d'explication. En effet, malgré une connaissance de plus en plus accrue du fonctionnement du cerveau, il reste toujours un « fossé explicatif » (Levine 1983) entre les propriétés de notre expérience et l'activité neuronale : au-delà des corrélations, nous sommes incapables de comprendre en quoi consiste la relation entre ces deux *relata*. C'est précisément l'hypothèse selon laquelle il y aurait une correspondance entre le contenu neuronal et le contenu de notre expérience que les énéactives remettent en question

4.2 L'approche énéactive, un tour d'horizon

Tel que je l'ai mentionné précédemment, l'énéactisme conçoit la perception comme étant *une manière d'agir* qui *dépend* d'un ensemble d'habiletés ou de connaissances sensorimotrices. Pour citer Noë : « ce que nous percevons est déterminé par ce que nous faisons (ou ce que nous savons faire); [ce que nous percevons] est déterminé par ce que nous sommes *prêts à faire* » (Noë 2004, p. 1). Ce faisant, les énéactives rejettent la séparation traditionnelle entre la perception et l'action. En effet, il est courant dans les manuels en sciences cognitives et en neuroscience de présenter le système perceptuel et le système moteur comme étant séparés et relativement indépendants; en conséquence, d'un point de vue explicatif, on considère habituellement que ces deux systèmes peuvent être étudiés et compris indépendamment l'un de l'autre.

De son côté, en affirmant que l'expérience perceptuelle est *constituée* par un ensemble de connaissances sensorimotrices implicites à l'égard de la manière dont le mouvement affecte la stimulation sensorielle, les énoncivistes rejettent cette séparation. C'est précisément cette que vise à défendre Susan Hurley dans son *Consciousness in Action* (1998, particulièrement dans le chap. 1). Alors qu'on conçoit traditionnellement la relation entre la perception et l'action comme étant *séquentielle* et *instrumentale*, Hurley affirme que leur relation est à la fois *dynamique* et *constitutivement interdépendante* (ibid., p. 11-2). Par «instrumentale» Hurley veut dire que « la structure et le contenu de la conscience ne devraient pas varier lorsque les *inputs* sensoriels restent inchangés. » (ibid., p. 197). En se fondant sur des données empiriques en neurophysiologie, Hurley (chap. 5) défend qu'il y a plutôt du *feedback dynamique*, c'est-à-dire un échange d'information, entre le système perceptuel et le système moteur. Précisons, cependant, que cette relation entre perception et action ne veut pas dire que la perception *sert à guider* l'action, mais plutôt que *la perception est une forme d'activité* (Noë 2004, p. 12).

Le rejet de la séparation entre système perceptuel et système moteur proposé par Hurley et endossé ensuite par Noë et O'Regan n'est pas sans soulever un certain scepticisme. Jesse Prinz, par exemple, remarque, avec raison, que des dommages causés au système visuel ne génèrent pas de déficits moteurs, ce qui semble contredire la thèse voulant que les deux systèmes soient interdépendants (2006, p. 11).

Par ailleurs, les données empiriques utilisées par Hurley sont tirées principalement d'expériences comportementales ou de cas de lésions cérébrales (1998, chap. 5). Pour cette raison, Prinz poursuit en affirmant qu'afin de justifier leur théorie, les énoncivistes devraient plutôt fournir des données neurologiques d'imagerie cérébrale prouvant que les systèmes moteurs sont activés lorsque nous avons des expériences visuelles. Cela

étant dit, il n'est pas clair que l'argument de Prinz touche sa cible. En effet, les énaïvistes ne prétendent pas que les systèmes moteurs font partie des corrélats neuronaux de la conscience.

De la conception énaïve de la perception, découle à son tour une thèse métaphysique sur la nature du *contenu perceptuel* : celui-ci est conçu comme étant *virtuel* (j'utilise ici l'expression de Noë) d'un bout à l'autre [*all the way in*]. Par « virtuel », j'entends ici la thèse selon laquelle le détail du contenu de notre expérience perceptuelle ne se trouverait pas *actuellement* dans notre esprit, mais serait plutôt *disponible* dans notre l'environnement immédiat, là où l'organisme *actif* peut aller le chercher lorsqu'il en a besoin. C'est ce que l'on veut dire lorsque l'on affirme que le contenu perceptuel est *constitué* par la relation étroite entre l'organisme actif et son environnement⁴¹. En ce sens, analyser les arguments de l'énaïvisme, c'est analyser les arguments en faveur de la virtualité du contenu perceptuel. Enfin, notons en passant le fait suivant : Noë et O'Regan défendent la thèse selon laquelle le détail de notre expérience est *totalelement virtuel*; il est cependant tout à fait possible, avec les mêmes arguments, de défendre la thèse plus faible et probablement moins controversée, selon laquelle notre contenu perceptuel serait *partiellement virtuel* : une partie de contenu de notre expérience serait alors stockée dans le cerveau, alors que le reste serait dans l'environnement.

Pour mieux illustrer la conception énaïve de la perception, Alva Noë a recours à de nombreuses analogies, la perception étant comparée à la fois à l'*activité de peindre* et à la *danse*. Je vais présenter ces analogies rapidement rapidement (Noë 2004, p. 222–3., O'Regan et Noë 2001, p. 906).

⁴¹ « [...] perceptual experience, according to the enactive approach, is an activity of exploring the environment drawing on knowledge of sensorimotor dependencies and thought » (Noë 2004, p. 238).

Lorsqu'un artiste peint un tableau, son regard effectue un va-et-vient entre le canevas et le monde; « voir, comme peindre, implique [un] processus étendu dans le temps » où l'agent sonde la scène à répétition (ibid.). Le fait de peindre ne peut être réduit aux états internes de l'agent, comme si celui-ci avait en tête une représentation détaillée de ce qu'il comptait peindre, mais met plutôt en jeu un engagement dynamique entre le peintre, la scène et le canevas.

Tout comme la danse, la vision est le résultat d'une interaction constante entre l'agent et son environnement. Selon l'approche éactive, il en va de même de la vision : la vision est une activité dynamique de l'organisme adroit [*skillful*] dans son environnement qui *se déroule dans le temps*; notre expérience visuelle est donc loin de prendre la forme d'un *instantané* qui puisse être abstrait du flux temporel dans lequel elle s'inscrit.

Conformément à cette idée, les éactivistes défendent que le paradigme de la vision ne doit pas être celui adopté par la conception photographique de la vision; la vision étant plutôt *similaire au toucher*, notre source d'inspiration devrait être celui de l'aveugle qui explore son environnement à l'aide de sa canne et qui, ce faisant, réussit petit à petit, à suivre son chemin dans un espace encombré en sondant son environnement *habilement* (Noë 2004, p. 1). C'est dans cette perspective que O'Regan et Noë affirment que :

To see is to explore one's environment in a way that is mediated by one's mastery of sensorimotor contingencies, and to be making use of this mastery in one's planning, reasoning, and speech behavior. (O'Regan & Noë 2001, p. 944).

Cette conception de la perception peut paraître à première vue contre-intuitive. De fait, nous sommes portés à croire spontanément que la perception est un phénomène *passif* où l'esprit entre en contact avec le monde extérieur. Cela se reflète dans le fait que nous considérons habituellement que la perception et l'action sont à la fois

séparées l'une de l'autre, mais aussi des aspects cognitifs de notre vie mentale d'ordre supérieur, comme le langage, le raisonnement ou les concepts, ce que Susan Hurley a appelé la *conception sandwich de la cognition* (1998, chap. 1, 8–10). Dans l'introduction de ce mémoire, j'ai exposé sommairement les points centraux de la critique menée au cours de ces dernières années à l'encontre de la conception sandwich de la cognition, critique qui a été au coeur des sciences cognitives incarnées. Cela dit, bien que le rejet de la conception sandwich soit un aspect central de l'approche énative, elle n'en est pas pour autant sa caractéristique principale : tel que je l'ai déjà dit à plusieurs reprises, l'énavivisme est d'abord et avant tout une thèse externaliste du contenu perceptuel.

Dans les paragraphes suivants, je vais présenter les concepts principaux de l'approche énative de manière très générale. Je vais ensuite entrer dans les détails dans les sections subséquentes. Je vais d'abord aborder le concept de contingence sensorimotrice (4.2.1.), avant de passer à celui de connaissance des contingences sensorimotrices (4.2.2.), puis, finalement, je vais aborder la relation de dépendance entre ces connaissances et notre expérience (4.2.3).

4.2.1 Les contingences sensorimotrices

Les *contingences* ou *régularités* sensorimotrices sont un ensemble de lois déterminant la relation entre le mouvement et la stimulation sensorielle de l'agent. Celles-ci peuvent être réparties ensuite en deux catégories, celles *relatives aux agents* et celles *relatives aux objets*.

Les contingences sensorimotrices sont déterminées par l'appareillage sensoriel dont l'organisme est pourvu⁴² : de sorte que les contingences sensorimotrices qui nous concernent diffèrent de celles qui caractérisent l'expérience d'une créature dotée d'un sonar, comme les chauves-souris. Selon les énaïvistes, ces contingences sensorimotrices ne font pas que caractériser le type d'expériences perceptuelles que peuvent encourir certaines créatures, elles permettent aussi de différencier les modalités sensorielles les unes des autres. Ce ne serait donc pas,, d'après cette théorie, l'aire neuronale activée qui détermine si mon expérience est visuelle plutôt qu'auditive, comme nous pourrions le croire à première vue. Cette hypothèse est présentée par Hurley et Noë (2003) comme étant un des points forts de l'énaïvisme; je reviendrai sur ce sujet plus bas. Mentionnons ici que les exemples donnés par les énaïvistes concernent presque exclusivement la perception visuelle; en ce sens, ils sont fidèles à la majorité des philosophes écrivant sur la perception, qui se concentrent presque qu'exclusivement sur cette modalité sensorielle. O'Regan et Noë donnent ainsi plusieurs exemples de contingences sensorimotrices distinguant les percepts visuels des percepts provenant d'autres modalités sensorielles. On peut penser ici aux saccades visuelles, que j'ai introduit plus haut, mais aussi au fait que nous ne sommes pas conscient de l'interruption de l'influx sensoriel lorsque nos paupières clignent; que lorsque l'oeil effectue une rotation dans son orbite, la stimulation sensorielle sur la rétine varie de manière spécifique, déterminée par l'ampleur du mouvement, la forme de la rétine, de même que les règles de l'optique oculaire; que lorsque nous nous déplaçons autour d'un objet partiellement dissimulé, les surfaces cachées apparaissent, etc. (O'Regan & Noë 2001, p. 941). Les auteurs restent par contre généralement silencieux sur les contingences sensorimotrices d'autres perceptions, sinon à de rares exceptions, lorsqu'ils traitent du toucher. À ce propos, ils notent quelques uns des traits caractéristiques du toucher : lorsque nous touchons un objet, nous ne le faisons pas d'un certain « point de vue » ou perspective

⁴² « To see *as we do*, you must then have a sensory organ and a body like ours. » (Noë 2004, p. 112).

comme dans le cas de la vision, car, entre autres : notre main peut souvent saisir un objet dans son entièreté ; nous ne touchons pas l'objet à distance ; et, finalement, le toucher n'est pas affecté par les conditions de luminosité. Cependant, plus souvent qu'autrement, les auteurs n'abordent cette modalité sensorielle que dans un contexte où ils tentent de démontrer à quel point la vision s'apparente au toucher (ibid., p. 942).

Le fait que O'Regan et Noë n'aient presque aucun exemple de contingences sensorimotrices pour l'ouïe, et encore moins pour l'odorat et le goût, tend à indiquer une faiblesse importante de leur théorie. Si, en effet, notre expérience d'un bon vin dépend d'une manière ou d'une autre de notre maîtrise des contingences sensorimotrices, alors les énéactives doivent nous expliquer comment cela peut-être possible.

Reprenons l'exemple des saccades visuelles, tel que présenté dans la section précédente. Ce cas est important pour énéactives, car il semble établir le fait que la perception est intrinsèquement *active* plutôt que passive. O'Regan et Noë affirment qu'il n'est pas nécessaire de postuler des mécanismes compensatoires pour les saccades visuelles, puisqu'elles font « partie de ce que c'est que de voir ». En un sens, il est indéniable que les saccades visuelles sont parties intégrantes de notre système perceptuel : nous ne pourrions avoir l'expérience visuelle que nous avons sans elles. Par contre, on peut douter qu'elles soient *nécessaires* pour qu'il y ait expérience perceptuelle, comme le soutiennent les énéactives. Non pas que cela soit impossible, mais si tel est le cas, les partisans de l'approche énéactive doivent soutenir cette thèse à l'aide de données empiriques qui soit probantes (pour une critique similaire, voir Prinz 2008).

4.2.2 Les connaissances ou habiletés sensorimotrices

Dans la section précédente, j'ai introduit le concept de contingence sensorimotrice, de même que l'idée selon laquelle ce sont elles qui distinguent les différentes modalités sensorielles. Cela étant dit, les contingences sensorimotrices ne suffisent pas pour qu'il y ait perception, : il faut aussi que l'organisme *comprenne* comment l'influx sensoriel change en fonction de ses propres mouvements (Noë 2004, p. 1). Ces connaissances sont une forme de *savoir-faire implicites*.

La notion de savoir-faire [*know how*], qui est devenue un enjeu philosophique grâce à Gilbert Ryle (1946, chap. 2, voir aussi 1971), est généralement opposée à celle de connaissance propositionnelle (ou factuelle). Procéder à l'aide d'exemples devrait ici être suffisant pour bien saisir cette distinction. En guise de connaissance factuelle, nous pouvons penser au fait que je sais que Londres est en Angleterre ou que $2 + 2 = 4$, alors que l'exemple paradigmatique de savoir-faire est le fait de savoir faire de la bicyclette. Déterminer si les connaissances propositionnelles peuvent être réduites à des savoir-faire, ou inversement, est matière à discussions, il est par contre généralement admis que la majorité des gens qui savent faire de la bicyclette sont difficilement capables d'explicitement verbalement ce savoir. En effet, la meilleure manière d'enseigner à quelqu'un comment faire de la bicyclette reste toujours de lui *montrer* en en faisant soi-même.

Donnons maintenant un peu plus de précisions sur cette notion de savoir-faire sensorimoteurs. D'abord, ces savoir-faire sont des *habiletés*. Une manière de concevoir ce qu'on entend par là, est de concevoir nos habiletés en termes de *dispositions* : je suis capable de faire de la bicyclette si j'ai la *disposition* à en faire (moyennant certaines conditions, comme le fait d'en avoir une à portée de la main). Ces habiletés sont non seulement *incarnées*, car, comme je l'ai déjà mentionné, elles dépendent de l'appareillage perceptuel dont nous sommes pourvus; elles sont aussi

situées, c'est-à-dire que leur actualisation dépend de conditions environnementales extérieures à l'agent (je ne peux jouer du piano sans piano) (Noë 2005, p. 284 et suiv.).

Jusqu'à présent, j'ai décrit les connaissances des contingences sensorimotrices comme étant le propre d'*agents en tant que tel*, et par conséquent, des propriétés attribuées au *niveau personnel*. Plusieurs passages des textes de Hurley, Noë et O'Regan vont d'ailleurs en ce sens (e.g. O'Regan & Noë 2001, p. 944). Cette interprétation tombe apparemment sous le sens, dans la mesure où l'on considère habituellement que ce sont des agents en tant que tels et non pas des parties de ceux-ci qui connaissent (les énéactives insistent d'ailleurs sur la distinction entre niveau personnel et subpersonnel). Cependant, lorsqu'on s'attarde aux détails, on constate que les auteurs attribuent parfois les connaissances sensorimotrices non pas aux agents, mais au cerveau, ce qui n'est pas sans nuire à l'intelligibilité de leur propos. O'Regan et Noë, par exemple, affirme que « the animal, or its brain, must be "tuned to" these laws of sensorimotor contingencies. » (ibid., p. 943). Si à première vue il est possible d'attribuer un sorte de croyance au cerveau, les partisans de la théorie énéactive doivent cependant préciser ce qu'ils entendent par là.

4.2.3 Dépendance et constitution

Lorsque les énéactives affirment que la perception se fait par l'*intermédiaire* de certaines habiletés, ils veulent dire que la perception est une activité que l'organisme apprend et que cet apprentissage *implique nécessairement* de maîtriser ou de connaître les contingences sensorimotrices propres à chacune de ses modalités sensorielles. C'est-à-dire que les énéactives conçoivent la relation entre la perception et l'action comme en étant une de *dépendance* et de *constitution* plutôt que de *causalité*. En fait, probablement personne ne nierait l'influence *causale* de l'action sur la perception; le plus simple mouvement de nos yeux dans leurs orbites affecte

l'influx sensoriel (pour une critique de la thèse constitutive de l'énactivisme, voir Prinz 2006, 2012). C'est d'ailleurs à cela que fait référence Noë lorsqu'il affirme que :

Perception is an activity that requires the exercise of knowledge of the ways action affects sensory stimulation (Noë 2007, p. 532, c'est l'auteur qui souligne).

Prenons un exemple : lorsque nous faisons de la bicyclette, nous avons recours à un ensemble de connaissances sans même que nous nous en rendions compte. Malgré cela, nous ne pourrions pas pratiquer cette activité si nous n'avions pas appris préalablement à en faire; apprendre à faire de la bicyclette est une chose ardue dans les premiers temps et qui demande un certain effort attentionnel (nous devons apprendre à garder notre équilibre et à évaluer les conséquences de nos mouvements), mais qui devient éventuellement une activité inconsciente. Dans la perception comme dans le fait de faire de la bicyclette, le rôle de ces connaissances intervient en deux temps : d'une part, dans le cadre de l'*extraction* des lois sensorimotrices (c'est la période d'apprentissage) et d'autre part, dans le cadre de leur *usage* effectif dans la perception. C'est précisément le rôle du cerveau que d'*extraire* et ensuite d'*utiliser* les différentes contingences sensorimotrices.

Affirmer que notre expérience perceptuelle dépend constitutivement (plutôt que causalement) de l'exercice de certaines connaissances sensorimotrices est une thèse somme toute, assez forte, puisque cela revient à soutenir que sans l'exercice de ces connaissances, *nous ne pourrions pas avoir d'expérience perceptuelle du tout*. Dans la suite de ce chapitre, j'examine trois types d'arguments en faveur de l'approche énactive : l'argument développemental (4.2.4.), l'argument de la cécité expérientielle (4.2.5.) et l'argument de la plasticité sensorielle (4.5). Les deux premiers, qui sont aussi les moins convaincants, seront abordés dans les deux prochaines sous-section, alors que le dernier sera l'objet d'une plus ample discussion dans une section subséquente.

4.2.4 L'argument développemental

Pour faire valoir le rôle nécessaire des connaissances sensorimotrices pour la perception, Alva Noë propose un *argument développemental*. Partant d'une expérience célèbre élaborée par Held et Hein (1963), les énoncivistes proposent un argument développemental qui vise à montrer qu'il ne suffit pas que l'agent soit bougé (extérieurement) pour qu'il y ait vision normale, il faut plutôt qu'il y ait un mouvement volontaire de la part de celui-ci [*self actuated*] (Noë 2004, p. 13, O'Regan et Noë 2001, p. 1020). Dans cette étude, Held et Hein ont fait une expérience sur deux chatons. Le dispositif de celle-ci fait en sorte que tous deux sont harnachés à un même carrousel. Tous deux reçoivent la même stimulation sensorielle, mais l'un d'eux était harnaché de telle manière que ses pattes touchaient au sol, alors que l'autre était suspendu dans les airs. Autrement dit, le premier pouvait se déplacer à l'aide de ses pattes dans les limites permises par le harnais, alors que le second se trouvait déplacé par le mouvement de l'autre chaton, sans rien n'y pouvoir.

Lorsque les chatons furent détachés, Held et Hein constatèrent que seul le premier pu développer une perception des profondeurs. L'approche énoncive explique ce phénomène de la manière suivante : c'est seulement via le mouvement généré par lui-même que l'agent peut en venir à apprendre et maîtriser les contingences sensorimotrices pertinentes. À bien des égards, cette expérience, quoique saisissante, ne semble pas convaincante. En effet, rien ne laisse croire que le chaton n'avait pas d'*expérience perceptuelle*. Il faudrait plutôt conclure qu'il n'arrivait pas à établir la relation appropriée entre la stimulation sensorielle et ses mouvements, faute de pratique. Autrement dit, cette expérience prouve peut-être que le chaton en question ne maîtrisait pas certaines contingences sensorimotrices, mais elle ne nous permet pas d'en conclure qu'il est une créature dénuée de conscience phénoménale.

4.2.5. Deux types de cécités

Un autre argument en faveur de l'énactivisme repose sur un phénomène que O'Regan et Noë appelle la *cécité expérientielle*. Selon eux, il y a deux types de cécité : la *cécité due à l'endommagement de notre appareillage sensoriel*, conformément à l'usage commun du mot, et la *cécité expérientielle* (Noë 2004, chap. 1). Par là, on entend que l'agent est incapable « d'intégrer les stimulations sensorielles au sein des trames [patterns] de mouvement et de pensée » (ibid., p. 4). De sorte que ceux qui souffrent de cécité expérientielle reçoivent ainsi bel et bien de l'influx sensoriel, mais celui-ci reste *confus* et *ne les informe pas sur le monde* : ils ne perçoivent pas des objets, des personnes, etc. avec lesquels ils pourraient interagir normalement. Autrement dit, il ne suffit pas d'avoir des sensations pour avoir une expérience perceptuelle qui soit *signifiante*, c'est-à-dire dotée d'un contenu représentationnel (les énaïvistes ne parlent pas en terme de contenu représentationnel, mais c'est précisément de cela dont il est question : notre expérience perceptuelle doit être *à propos* de quelque chose, c.f. O'Regan & Noë 2001, p. 943).

Ce type de cécité a été introduit par les énaïvistes pour faire référence à un ensemble de phénomènes; ce n'est donc pas un terme utilisé par les spécialistes oeuvrant dans le domaine de la vision. Dans *Action in Perception*, Alva Noë (2004, p. 4 – 6) mentionne plusieurs cas de cécité expérientielle, il cite, par exemple, Gregory et Wallace (1963, p. 366), qui rapportent l'expérience d'un patient qui vient de subir une opération chirurgicale pour traiter des cataractes. Je me permets de citer le passage dans son entièreté :

S.B.'s first visual experience, when the bandages were removed, was of the surgeon's face. He described the experience as follows: He heard a voice coming from in front of him and to one side: he turned to the source of the sound and saw a "blur." He realized that this must be a face. Upon careful questioning, he seemed to think that he would not have known that this was a

face if he had not previously heard the voice and known that voices came from faces.

Bien établir l'existence de ce phénomène est crucial pour les énoncivistes, et ce, pour deux raisons. D'une part, celle-ci vient supposément mettre à mal le modèle traditionnel de la cognition, selon lequel *input* et *output*, perception et action, sont nettement séparées; d'autre part, cela viendrait confirmer que la thèse centrale de l'approche énoncive, selon laquelle la perception implique une forme de connaissance implicite. Alva Noë introduit la distinction entre ces deux types de cécités précisément pour faire voir ce qui se passe lorsque nous avons un influx sensoriel *sans connaissances* et donc, pour démontrer que la perception *dépend constitutivement* de certaines savoir-faire sensorimoteur, comme il le dit lui même d'ailleurs : « ce qui viendrait infirmer l'approche énoncive serait l'existence de perception en l'absence d'habiletés corporelles et de connaissance sensorimotrice [...] » (2004, p. 12). Dans les cas de cécité expérientielle, le sujet n'a que des sensations *dénuées de contenu représentationnel* [*bleached of content*] (2004, p. 4). Plus précisément, ce point fonde la distinction entre *sensation* et *perception*. Par sensations, les énoncivistes font référence à la manière dont les sens sont affectés par des stimuli, alors que la perception est le *résultat* de la catégorisation d'objets et d'événements dans l'environnement (O'Regan & Noë 2001, p. 943) : percevoir implique « d'intégrer de la bonne manière » l'influx sensoriel par l'entremise d'habiletés corporelles (Noë 2004, p. 4).

La façon dont j'interprète cette « intégration » de l'influx sensoriel est en terme *d'affordances* que nos habiletés sensorimotrices nous permettraient de percevoir. Cette notion a été initialement introduite par James Gibson, mais ensuite récupérée par O'Regan et Noë, qui endossent complètement les propos de Gibson (2001, p. 1019). Voici comment Gibson présente cette notion :

The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb to *afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (1979, p. 127, l'auteur souligne)

On voit dans l'extrait cité qui vient d'être cité que par « affordances », on fait référence aux *possibilités* d'interactions (manipulation, nutrition, etc.) avec divers éléments de notre l'environnement : les affordances sont ce que l'environnement *offre* à un organisme donné. Ces possibilités d'interactions peuvent être soit positives, un biscuit m'offre la possibilité d'être mangé, ou négatives, le feu m'offre la possibilité d'être brûlé. En plus des objets inanimés, les autres êtres vivants offrent aussi des possibilités d'interactions qui peuvent être décrites en terme d'affordances, e.g. être un potentiel partenaire sexuel ou bien un potentiel interlocuteur dans une discussion. Notons aussi que les objets présents dans notre environnement ne sont pas les seuls à offrir des affordances, le « décor », si l'on peut dire, comme les surfaces, offrent aussi des possibilités d'interactions : je peux marcher sur la plancher, nager dans l'eau, etc..

Ce qui est controversé à propos de l'idée d'affordance, c'est que Gibson concevait celles-ci comme étant *perçues* et qu'elles existent dans le monde même lorsqu'elles ne sont pas perçues actuellement. Elles ne sont donc pas des « construits » qui apparaissent au gré des besoins des organismes. Il est cependant important de comprendre que Gibson concevait les affordances comme dépassant la dichotomie entre ce qui est objectif et subjectif : si les affordances sont bien *réelles*, elles sont toutefois relationnelles, c'est-à-dire que les affordances perçues par un organisme dépendent des actions que celui-ci peut faire (une chaise m'offre la possibilité de m'asseoir, mais offre d'autres possibilités d'actions à mon chat). Malgré cela, il est essentiel de comprendre que ces affordances *sont dans les objets eux-mêmes*.

Cela étant dit, retournons aux cas de cécité expérientielle. Nous pouvons expliquer la cécité expérientielle comme une forme d'incapacité à percevoir des affordances. Par contre, même si ces patients n'arrivent pas à percevoir des affordances, il est difficile, encore une fois, de croire qu'ils soient totalement dénués d'expériences phénoménales. Par exemple, faire l'expérience d'une douleur au genou, ou d'une image rémanente [*after image*] est à première vue possible, même si l'on n'a pas certaines connaissances sensorimotrice. On voit ici un exemple probant où la notion d'expérience que nous choisissons détermine le type de théorie de la conscience que l'on va élaborer. Les énaïvistes ont une conception très forte de l'expérience perceptuelle : percevoir, ce n'est pas seulement faire l'expérience de *qualités sensorielles* (e.g. l'expérience de l'amertume du café), c'est aussi, et surtout, entrer en contact avec des objets et un environnement qui permettent de riches interactions.

Enfin, j'en profite pour mentionner que plusieurs auteurs soutiennent que la perception des affordances dépend des neurones canoniques dans le cortex moteur (pour une bonne discussion des évidences en neurophysiologie en faveur de l'approche énaïve c.f. Garbarini et Andenzato 2004). Si tel est le cas, nous pourrions très bien imaginer le cas de personnes capable de percevoir, mais incapable de voir les affordances.

En guise de conclusion partielle, résumons les points principaux de la théorie énaïve de la perception. Selon cette théorie, la perception est *constituée* de deux éléments principaux, à savoir : d'une part, de l'organisme qui *explore* et *sonde* les structures d'information de son environnement à l'aide d'un certain nombre d'*habiletés* sensorimotrices bien spécifiques; d'autre part, la perception implique aussi un ensemble d'*attentes* à l'égard de ce que l'organisme devrait percevoir s'il se déplace d'une façon donnée, attentes qui sont elles-mêmes possibles parce que l'organisme dispose de certaines connaissances sensorimotrices. Quelles sont les conséquences de

cette conception de la perception pour la conscience ? Je vais aborder cette question dans la prochaine section.

4.3 L'approche éactive et le caractère phénoménal de la perception

Dans la section précédente, j'ai fait un bref survol de l'approche éactive de la perception, sans pour autant en tirer des conclusions sur le caractère phénoménal de la perception. Ici, je vais concentrer mon propos autour de trois points importants. D'abord, je vais aborder la conception *bidimensionnelle* de la perception (4.3.1); ensuite, je vais me concentrer sur le phénomène de la présence dans l'absence (4.3.2); enfin, je vais en dire un peu plus sur le rôle des *attentes* dans le cadre de l'expérience perceptuelle (4.3.3).

4.3.1 La conception bidimensionnelle de la perception

La conception bidimensionnelle de la perception est l'idée selon laquelle la perception est un processus qui se déroule en deux étapes, *c'est-à-dire qu'elle nous permet de déterminer comment les choses sont à partir de leurs apparences [looks]*. En ce sens, la conception éactive de la perception conserve l'idée centrale des théories des sense-data, sans pour autant adopter les conséquences épistémologiques ou ontologiques qui lui sont liées (Noë 2004, p. 81).

La plupart des philosophes modernes, dont René Descartes, ont soutenu qu'il faut distinguer entre l'objet dont nous sommes directement conscients dans la perception et l'objet physique que nous croyons percevoir. Notre relation au « monde extérieur » (s'il en est un), se ferait donc par l'entremise d'un *intermédiaire*. Ainsi, l'objet de ma perception ne serait pas *dans le monde*, mais serait plutôt une « idée » (Descartes, Locke ou Berkeley), ou encore une « impression » (Hume), se trouvant dans l'esprit de l'agent percevant. Il en va de même des théories des sense-data défendues

initialement par G. E. Moore et Bertrand Russell à l'aube du vingtième siècle et plus récemment par Howard Robinson (1994). Bien entendu, ces diverses théories de la perception sont fort distinctes et ne peuvent être amalgamées sans faire quelques torts à l'exactitude historique (par exemple, lorsqu'il est question de déterminer si l'*intermédiaire* en question est doté de propriétés représentationnelles ou non), mais elles souscrivent toutes à l'idée d'après laquelle nous ne percevons pas le monde directement.

Ce qui a poussé de nombreux philosophes à croire que la perception se rapporte non pas *directement* au monde, mais plutôt à un objet interne, c'est l'intuition cartésienne derrière l'*argument du rêve* et l'*argument du mauvais génie*, d'après lesquels un cas normal de perception pourrait être impossible à distinguer d'une illusion. Ces philosophes défendent que si notre perception se rapporte directement aux objets du monde, il est impossible de rendre compte de certains cas d'erreur perceptuelle, comme les illusions et les hallucinations. Par conséquent, ce ne sont pas des considérations phénoménologiques, mais plutôt des considérations *épistémologiques*, qui motivent des positions comme le *réalisme indirect*, où l'objet de notre perception est une représentation se rapportant au monde et le *phénoménalisme* où il n'y a que des représentations, mais pas de monde extérieur indépendant de la perception.

Revenons-en à l'approche éactive. Afin d'illustrer le fait que la perception est un processus qui se déroule en deux étapes, Noë (2004, p. 82) adopte un argument initialement développé par Peacocke pour fonder la distinction entre *contenu représentationnel* et *contenu sensationnel* (ce que certains appellent les qualia) (1983, p. 12, mais il faut noter que dans son 1992 Peacocke a changé de position).

L'idée est très simple et peut être illustrée via un exemple : imaginons que je me tiens debout sur une route qui s'étend en ligne droite vers l'horizon. Sur le bord de la route, il y a deux arbres *de la même taille*, l'un à environ cent mètres, l'autre, un peu plus

loin, à environ deux-cents mètres. Mon expérience me *représente* ces deux arbres comme ayant *physiquement* la même taille, mais le premier occupe néanmoins un plus grand espace dans mon champ visuel. D'après Peacocke, cela est une propriété sensationnelle de mon expérience perceptuelle, dans la mesure où elle détermine *ce que cela fait* que de percevoir le paysage qui s'étend devant, sans pour autant être une des propriétés du monde. Ces propriétés sont ce à quoi les énaïvistes font référence lorsqu'ils parlent d'apparences [*looks*] ou de propriétés perspectivales [*p-properties*].

Précisons, cependant, que ce n'est pas parce qu'on conçoit la perception comme étant un processus en deux étapes que notre expérience du monde doit être *indirecte*. C'est justement le type de conclusion métaphysique ou épistémologique que les énaïvistes espèrent éviter. Plutôt, nous percevons le monde *tel qu'il est par l'entremise des apparences/perspectives* (Noë 2004, p. 79 et suiv.). En ce sens, les deux dimensions de notre contenu perceptuel font partie de notre expérience. Lorsque nous nous engageons dans des activités mondaines, nous sommes dirigés vers notre contenu représentationnel, alors que lorsque nous adoptons un regard particulier sur notre expérience, comme lorsqu'on peint une toile ou faisons de l'introspection, alors nous sommes dirigés vers les apparences. Ainsi, autant les propriétés physiques que les apparences sont perçues⁴³. Quel est le statut ontologique de ces propriétés perspectivales ? Contre Peacocke, Alva Noë défend l'idée selon laquelle les apparences ne sont pas des propriétés subjectives : bien que relationnelles, elles sont des propriétés objectives qui peuvent être mesurées mathématiquement. Par là, on veut dire que la taille qu'occupera un arbre dans mon champ visuel peut être mesuré par les lois de l'optique, moindrement que je sache sa taille réelle, nos emplacements

⁴³ « perspectival properties (or P-properties). P-properties are themselves *objects of sight*, that is, things that we see. [...]In normal life we tend to pay little attention to P-properties. Only in special circumstances (e.g., in the task of artistic rendering) do they become salient. » (2004, p. 83).

respectifs, etc.. Nous avons ainsi la *taille physique* de l'arbre et la taille dans le champ visuel, i.e. la taille de l'arbre tel qu'il est perçu de l'endroit où je me trouve.

4.3.2 La présence dans l'absence

Des propriétés perspectiveles dont je viens de parler nous pouvons passer à un autre aspect de la perception : celui de la « présence dans l'absence » (c.f. Noë 2004 chap. 2–3, 2012). Ce que l'on veut dire par cette expression est que lorsque nous percevons visuellement un objet, nous sommes aussi perceptuellement conscient des faces cachées de l'objet. C'est donc dire que nous percevons un objet dans son entièreté et non pas uniquement la face ou l'esquisse qui est disponible du point de vue ou nous nous trouvons actuellement. Le phénomène de la *présence dans l'absence* n'a évidemment pas été introduit par les énéactives, mais remonte aux problèmes classiques touchant la perception. Il est, en effet, loin d'être évident de comprendre comment nous pouvons percevoir objectivement le monde alors que nous ne percevons que des facettes de celui-ci (voir, par exemple, la notion d'*Abschattung* chez Husserl).

Alva Noë affirme expliquer ce phénomène à l'aide de sa conception sensorimotrice de la perception. Par exemple, lorsque je vois une maison, je la vois dans son *entièreté*, même si, du point de vue où je me trouve, je n'appréhende que sa façade (2009, p. 61). Autrement dit, la notion de présence dans l'absence fait référence à la perception de certaines parties d'un objet, parties qui sont dénuées d'*attributs visibles*, phénomène qui est rendu possible parce que nous *savons* comment avoir accès aux parties cachées de l'objet en question (Pessoa et al. 1998, p. 728). En effet, dans *Action in Perception*, Noë explique :

Our perceptual sense of the tomato's wholeness—of its volume and backside, and so forth—consists in our implicit understanding (our expectation) that the

movements of our body to the left or right, say, will bring further bits of the tomato into view. Our relation to the unseen bits of the tomato is mediated by patterns of sensorimotor contingency. Similar points can be made across the board for occlusion phenomena. (2004, p. 63).

En affirmant que les parties cachées sont *présentes*, les éenactivistes affirment non seulement que celles-ci sont perçues, c'est-à-dire qu'elles font partie de notre *phénoménologie*, mais aussi qu'elles sont « disponibles » au sujet percevant précisément parce que nous avons certaines connaissances sensorimotrices.

Notons, encore une fois, que cette manière de concevoir la présence dans l'absence est cependant loin d'aller de soi. En effet, Alva Noë conçoit ce phénomène comme étant *perceptuel* plutôt que comme étant *cognitif*, tels une croyance ou un jugement sans pour autant fournir un argument allant en ce sens.

D'un autre côté, Jesse Prinz (2006) formule aussi deux critiques contre cette explication de la présence dans l'absence. D'une part, il est loin d'être évident que le système moteur soit impliqué dans le processus donnant lieu à la complétion amodale : le système visuel pourrait en être totalement responsable. Contrairement au cas de conscience, nous pouvons douter que les éenactivistes puissent donner, soutenir cette thèse sans fournir de données empiriques qui soit probantes. D'autre part, les arguments faisant appel à notre phénoménologie soulèvent souvent des controverses qui semblent difficiles à résoudre de manière objective, il faut donc procéder avec prudence. Cela dit, Prinz a raison d'affirmer que l'on peut interpréter le phénomène de la présence dans l'absence d'autres manières : nous pourrions tout simplement dire que nous avons des attentes inconscientes et que par conséquent nous serions *surpris* si, à la place de face cachée, il n'y avait rien. Ça ne veut pas dire pour autant que la face cachée *fait partie de notre expérience*.

4.3.3. Le rôle des attentes dans l'expérience perceptuelle

Le phénomène de la présence dans l'absence, dont il a été question plus haut, est étroitement lié à conception dispositionnaliste du contenu perceptuel. Concevoir la perception comme une activité où l'organisme va chercher les détails dont il a besoin pour réaliser les tâches cognitives, découle ainsi une conception *dispositionaliste* du contenu perceptuel : notre expérience est constituée par les *attentes* de l'agent concernant la manière dont le mouvement affecte l'influx sensoriel. On constate cela dans l'exemple de la conduite automobile, que Kevin O'Regan et Alva Noë affectionnent particulièrement :

What it is like to drive a Porsche is *constituted* by all these sensorimotor contingencies and by one's *skillful mastery* of them, – one's confident knowledge of how the car will respond to manipulations of its instruments. (O'Regan & Noë 2001, p. 961, je souligne).

Cette conception dispositionnaliste de la perception a notamment été critiqué par Horgan et Kriegel (2008) : selon eux, la conscience phénoménale est propriété *occurente* et non pas une propriété dispositionnelle comme le soutiennent les énaclivistes.

Par ailleurs, le rôle donné à ces attentes dans l'explication du caractère phénoménal a aussi soulevé d'autres inquiétudes (ce point a été soulevé notamment par Clark 2006, mais aussi, dans une moindre mesure, par Rowlands 2003, 2006). En effet, comme le remarque Andy Clark, nous pouvons interpréter de deux façons distinctes le fait que notre expérience d'une forme géométrique ou d'une couleur soit constituée par diverses *attentes* concernant la relation entre le mouvement et la stimulation sensorielle : ces attentes peuvent être soit au niveau personnel, soit au niveau subpersonnel. Au niveau personnel, elles seraient des propriétés de notre expérience consciente, mais ne pourrait être invoquées pour expliquer le caractère phénoménal,

faute de circularité. Au niveau subpersonnel, elles seraient des propriétés du cerveau. Une telle interprétation aurait par contre le même défaut que le programme de recherche visant à identifier les corrélats neuronaux de la conscience : que notre système perceptuel ait ou non recours à de telles attentes ou prédictions, cela ne nous aide aucunement à résoudre le déficit d'explication. Cela dit, bien que les énaïvistes soient parfois ambivalents à propos de l'interprétation qu'ils préfèrent⁴⁴, l'interprétation selon laquelle les attentes se trouvent au niveau personnel semble être celle qui est la plus cohérente. C'est ce que l'on constate dans plusieurs passages portant sur le caractère phénoménal de la perception. Ici, j'en cite deux :

1. To experience a ball, say, as moving toward you is not to have an experience of a special sort of quality (the "something's moving toward one"-quale), just as the experience of the item's being off to the left is not the experience of a special quality or sensation of leftness. When you experience the ball as moving toward you, *you experience it, precisely, as such that, by moving, you would change the character of the way its appearance changes.* (Noë 2004, p. 88, je souligne).

O'Regan et Noë appliquent le même raisonnement à l'égard de l'expérience du rouge :

2. The experience of red, for example, arises when *we know* [...] that, for example, if we move our eyes over a red region, there will occur changes typical of what happens when our non-homogeneously sampling retinas move over things whose color is red. (O'Regan & Noë 2001 p. 963, je souligne).

⁴⁴ **Niveau personnel** : à propos de son texte publié avec O'Regan dans BBS, il affirme que « We proposed that the qualitative character of experience depended on the perceiver's practical knowledge of such patterns, and on the perceiver's ability to use this knowledge to guide thought and action. » (p. 228). **Niveau subpersonnel** : cependant, dans le même texte, les auteurs affirmaient que : « This is because your brain is "tuned" to certain potentialities [...] » (O'Regan & Noë 2001, p. 945). On peut comparer ces passages avec d'autres où la confusion est encore plus évidente « the animal, or its brain, must be "tuned to" these laws of sensorimotor contingencies. » (ibid., p. 943). Ces ambiguïtés sont d'autant curieuses, que les énaïvistes insistent pour qu'on ne confondent pas les niveaux d'explications personnels et subpersonnels. Entretenir des connaissances, c'est quelque chose qu'un *organisme* fait, pas son cerveau.

Ces extraits tendent à montrer deux choses. D'une part, que percevoir, c'est être confronté à un état de chose possible, à une manière dont le monde pourrait être. En ce sens, l'expérience perceptuelle soulève la question visant à déterminer si le monde est tel que présenté dans notre expérience. D'autre part, ces extraits montrent non seulement que les connaissances des contingences sensorimotrices se trouvent au niveau personnel, mais *aussi* que nous en sommes *conscients*. En effet, tel que mentionné dans la dernière phrase du premier extrait, faire l'expérience d'une balle qui fonce sur nous, *c'est faire l'expérience qu'en nous déplaçant nous aurions une expérience différente*. Intuitivement, cette idée pose problème : en effet, on peut douter que nos attentes sensorimotrices aient réellement une trace phénoménologique comme le suggèrent O'Regan et Noë dans ces extraits. Encore fois, si ces attentes n'ont pas, en tant que tel, de caractère phénoménal, alors en quoi peuvent-elles conférer des propriétés expérientielles à nos actes de perception ?

Ces problèmes d'interprétation touchant la théorie éactive démontrent bien que certains aspects devront être développés davantage par ses partisans s'ils espèrent démontrer sa viabilité.

4.4 Antireprésentationnalisme : rejet du *bridge locus*

Tel que mentionné précédemment, la stratégie générale adoptée par les éactivistes pour résoudre les problèmes liés à la vision, à la perception et à l'expérience perceptuelle consiste à mettre l'accent sur l'exercice de savoir-faire sensorimoteurs plutôt que sur des processus, états ou propriétés internes. Cette stratégie les pousse à adopter une forme *d'antireprésentationnalisme* : le système visuel n'aurait pas besoin de produire une représentation détaillée de l'environnement puisque l'organisme a un accès rapide et fiable à l'information contenue dans l'environnement lorsque cela s'avère nécessaire (O'Regan & Noë 2001, p. 963). Comme nous l'avons vu dans la première section, dans le cadre des approches traditionnelles critiquées par

l'énactivisme, l'*explanandum* consiste toujours à postuler l'existence d'une certaine entité dans l'esprit-cerveau de l'organisme percevant (habituellement un type particulier de représentation mentale), le problème étant de concevoir comment de telles entités peuvent exister et comment leur existence permet d'expliquer le caractère phénoménal de notre expérience.

Dans leur article phare publié en 2001, Kevin O'Regan et Alva Noë ont adopté une position somme toute hostile aux représentations mentales. Cette position a cependant été affaiblie par la suite dans leurs publications subséquentes. Dans *Action in perception*, Alva Noë précise effectivement :

No doubt perception depends on what takes place in the brain, and very likely there are internal representations in the brain (e.g., content-bearing internal states). What perception is, however, is not a process in the brain, but a kind of skillful activity on the part of the animal as a whole. (2004, p. 2, je souligne).

En effet, le construit théorique principal de l'approche énaïve, les connaissances des contingences sensorimotrices, est aussi une représentation mentale dans le sens où on l'entend généralement (i.e. un état interne avec des propriétés sémantiques). Malgré ce changement de cap, O'Regan et Noë ne précisent dans aucune de leurs publications en quoi ces représentations sensorimotrices devraient consister et ne donnent aucun détail sur leur rôle dans l'architecture de l'esprit. On peut cependant prendre pour acquis que pour ces auteurs, le rôle attribué aux représentations dans la science de la vision n'est plus central, comme c'est le cas chez Marr, par exemple. Une explication en termes représentationnels ne suffirait donc pas pour rendre compte de la vision : « voir consiste à *utiliser* des représentations, non pas à *avoir* des représentations » (O'Regan et Noë 2001, p. 1017, je souligne).

L'antireprésentationnalisme énaïve trouve sa justification principalement dans les expériences sur la cécité inattentionnelle de même que sur la cécité au changement.

Ces travaux tendent à démontrer que nous remarquons certains changements dans notre scène visuelle seulement si le changement en question affecte la signification de la scène pour l'organisme (Noë 2001, p. 41). À l'origine, les énéactivistes semblaient donner un rôle beaucoup plus décisif à de telles expériences⁴⁵, la cécité au changement étant censée démontrer qu'il n'y a pas de modèle interne. Par contre, dans des publications qui ont suivies⁴⁶ Alva Noë a fait marche arrière et a admis que la cécité au changement ne pas prouver le fait qu'il n'existe pas de représentations détaillées dans l'esprit des agents percevant. Toujours selon Noë, dans le meilleur des cas, la cécité au changement démontre qu'il *pourrait* ne pas y avoir de modèle interne, ou encore, que celui-ci n'est pas nécessaire à la perception.

Cependant, même cette interprétation plus faible des travaux sur la cécité au changement a suscité un certain nombre de critiques, notamment de la part de Pierre Jacob (2006). En effet, les protocoles expérimentaux visant à mettre en évidence la cécité au changement consistent habituellement à alterner deux images où se trouvent une différence flagrante, cette différence étant adroitement dissimulée aux observateurs par ceux qui ont élaboré l'expérience. Cependant, cela ne prouve pas pour autant que nous n'ayons pas de représentation détaillée de la scène visuelle dans notre cerveau. Comme le souligne Jacob, il est possible qu'une représentation détaillée soit produite par le système visuel, mais qu'elle soit « effacée » avant que la comparaison entre les deux images puisse prendre place. Une autre possibilité : les deux représentations détaillées pourraient se trouver dans systèmes neuronaux qui ne soient disponibles pour une comparaison, ou encore être encodées dans un format qui

⁴⁵ Par exemple : « Inattentional blindness [...] *demonstrate* that you do not have all the details of what is seen in consciousness at once » (O'Regan et Noë p. 960), ou encore : « Phenomena such as change blindness *demonstrate* that the brain does not produce models of the detailed environment to support vision [...] » (Noë 2001, p. 46).

⁴⁶ Noë est très clair sur ce point : « Change blindness is evidence, then, that the representations needed to subserve vision *could* be virtual. Change blindness suggests that we don't make use of detailed internal models of the scene (even if it doesn't show that there are no detailed internal representations). In normal perception it seems that we don't have online access to detailed internal representations of the scene. » (Noë 2004, p. 52, l'auteur souligne).

rend la comparaison impossible. En évoquant ces différentes possibilités, mon objectif n'est pas de défendre la thèse selon laquelle une représentation détaillée existe effectivement, mais plutôt de démontrer que la conclusion que tirent les énoncivistes ne va pas de soi. Les énoncivistes doivent donner une réponse satisfaisante à ces critiques, puisque l'antireprésentationnalisme repose directement sur les arguments qui ont recours à la cécité au changement.

4.5 Les modalités sensorielles et l'argument de la plasticité

J'ai mentionné plus haut qu'à de rares exceptions près, les ouvrages publiés sur la conscience ou sur la perception traitent de ces sujets de manière très générale, sans aborder directement les caractéristiques particulières des diverses modalités sensorielles, et lorsqu'ils le font, ils se concentrent presque exclusivement sur la perception visuelle, délaissant ainsi les autres modalités. Ceci est peut-être une des raisons expliquant pourquoi la question visant à déterminer les critères nous permettant de distinguer les modalités sensorielles les unes des autres est si souvent occultée, malgré le fait que ce soit une question philosophique digne d'intérêt (c.f. MacPherson 2011 pour une introduction détaillée de ce problème, de même qu'un ensemble de textes classiques sur la question).

Intuitivement, nous pouvons distinguer les modalités sensorielles par l'entremise de leurs organes sensoriels respectifs. Ainsi, la vision serait le produit des yeux et du système visuel dans son ensemble. Mais cette conception naïve n'est pas sans causer de difficultés, puisqu'il faut ensuite trouver un critère pour distinguer lesdits organes les uns des autres. Sans compter le fait que nos organes sensoriels peuvent être stimulés de manière inappropriée : si je me mets le doigt dans l'oeil, il ne serait pas approprié de dire que je « vois » mon doigt. De plus, adopter un critère aussi fin pour distinguer les modalités sensorielles nous empêcherait d'affirmer que les prothèses venant pallier à un organe sensoriel endommagé remplacent effectivement cet organe.

C'est à cette position que font référence O'Regan et Noë lorsqu'ils critiquent la théorie des « énergies spécifiques des nerfs » de Müller (1838) : que l'on retrouve aujourd'hui reformulée dans le langage des neurosciences. L'idée de Müller, à laquelle les énoncivistes s'opposent, consiste à soutenir que « ce qui détermine l'aspect visuel particulier des sensations visuelles est le fait que ces sensations sont transmises par une voie neuronale particulière (à savoir, celle qui trouve son origine sur la rétine et non dans la cochlée) et qui projette sur des régions cérébrales particulières (essentiellement, l'aire corticale V1) » (O'Regan & Noë 2001, p. 940).

Partant de là, on peut imaginer au moins deux façons alternatives de concevoir cette distinction. On peut d'abord distinguer les modalités sensorielles les unes des autres par le *caractère phénoménal* respectif à chacune d'entre elles : autrement dit, la vision et le toucher seraient caractérisés par des qualia de nature différente. C'est, par exemple, la position que E. J. Lowe défend : « Une expérience peut être qualifiée d'expérience visuelle purement en vertu du caractère phénoménal ou qualitatif qui lui est intrinsèque » (Lowe 1992, p. 80). Lowe justifie sa position en affirmant que, bien que l'expérience visuelle soit habituellement produite par la stimulation des photorécepteurs sur la rétine, la stimulation directe du cortex peut donner lieu à des expériences visuelles comme l'ont montré les travaux célèbres de Wilder Penfield. On peut ensuite distinguer les modalités sensorielles en fonction des *propriétés* auxquelles elles nous donnent accès : la vision nous donnant accès aux formes, aux couleurs, etc. Cependant, une objection naturelle contre cette position est le fait que certaines propriétés peuvent être perçues par plus d'une modalité sensorielle (e.g. la forme géométrique), situation qui concerne surtout la vision et le toucher.

Tel que je l'ai mentionné à de multiples reprises précédemment, les énoncivistes ont formulé une hypothèse pour expliquer ce phénomène et considèrent que cette explication est une des forces en faveur de leur théorie. Plus précisément, selon eux, les modalités sensorielles ne sont individuées ni par les propriétés intrinsèques des

organes impliqués, ni par les aires cérébrales activées, ni par des différences dans leur caractère phénoménal, mais plutôt par les contingences sensorimotrices propres à chacune d'entre elles⁴⁷ (pour une exposition détaillée de cette question, voir Hurley & Noë 2003, voir aussi O'Regan et Noë 2001, p. 941, 1011–2).

À première vue, l'on pourrait croire qu'affirmer que les modalités sensorielles sont individuées par les contingences sensorimotrices consiste à affirmer qu'elles sont individuées par le type d'expériences subjectives auxquelles elles donnent lieu (i.e. les qualia). Du point de vue de la théorie éactive, il faut cependant distinguer ces deux positions, et ce, pour deux raisons. D'abord, les contingences sensorimotrices sont des propriétés objectives qui peuvent être exprimées mathématiquement : ce sont elles qui font que nous avons telle ou telle expérience. D'autre part, les contingences sensorimotrices sont plus abstraites que ce dont nous faisons l'expérience subjectivement, par exemple la perception par TVSS est une forme de vision, même si subjectivement percevoir par TVSS donne lieu à une expérience assez différente de la vision normale.

Cela étant dit, la question des modalités sensorielles est importante pour les éactivistes, et ce pour plusieurs raisons. D'abord, s'ils sont parfois hésitants quant à la capacité de la théorie éactive à expliquer le déficit d'explication « absolu », ils prétendent cependant pouvoir fournir une explication satisfaisante des déficits d'explications comparatifs⁴⁸. Dans leur article de 2003, Hurley et Noë affirment qu'en

⁴⁷ « On the other hand, what *does* differentiate vision from, say, audition or touch, is the *structure of the rules* governing the sensory changes produced by various motor actions, that is, what we call the *sensorimotor contingencies* governing visual exploration. Because the sensorimotor contingencies within different sensory domains (vision, audition, smell, etc.) are subject to different (in)variance properties, the structure of the rules that govern perception in these different modalities will be different in each modality. » (O'Regan & Noë 2001, p. 941).

⁴⁸ S'ils sont hésitants dans Hurley & Noë 2003, dans O'Regan et Noë 2001 ils affirment clairement avoir une solution au déficit d'explication absolu.

plus du déficit d'explication absolu dont j'ai traité longuement le premier chapitre, il y a deux déficits d'explication comparatifs :

a. déficit absolu : peu importe la quantité d'information *objective* que nous pourrions accumuler sur l'activité du cerveau, nous ne saurions toujours pas expliquer *pourquoi* tel ou tel phénomène objectif produit ou génère tel ou tel phénomène subjectif.

b. déficit intermodal comparatif : peu importe la quantité d'information *objective* que nous pourrions accumuler sur l'activité du cerveau, nous ne saurions toujours pas expliquer *pourquoi* telle ou telle activité neuronale dans le système visuel produit ou génère une expérience visuelle plutôt qu'auditive ?

c. déficit intramodal comparatif : peu importe la quantité d'information *objective* que nous pourrions accumuler sur l'activité du cerveau, nous ne saurions toujours pas expliquer *pourquoi* telle ou telle activité neuronale dans le système visuel produit ou génère une expérience visuelle de rouge plutôt qu'une expérience visuelle de vert ?

La question de la distinction des différentes modalités sensorielles (et des différences au sein de celles-ci) est aussi importante pour les énoncivistes, dans la mesure où la majorité des arguments de nature empirique mis de l'avant en faveur de l'approche énoncive portent sur des cas *substitution sensorielle*, c'est-à-dire des situations où l'expérience propre à une modalité sensorielle, e.g. l'expérience visuelle, est réalisée par un autre organe (biologique ou prosthétique) ou par une autre aire du cerveau que lors de situations normales. En mettant l'accent sur ces cas précis, les énoncivistes tentent de faire valoir que leur approche permet de combler les deux déficits comparatifs, dans la mesure où les différences intermodales et intramodales sont dues aux diverses contingences sensorimotrices qui régissent notre expérience.

Ce type d'argument vise aussi à répondre aux critiques de *chauvinisme sensorimoteur* formulées par Clark et Toribio (2001, voir aussi Clark 2008 p. 167–195 pour une discussion plus approfondie). Selon ces derniers, les énaïvistes ont raison d'insister sur la relation étroite entre contenu perceptuel et action incarnée. Cependant, ils affirment que les énaïvistes vont trop loin en formulant les contingences sensorimotrices à un niveau d'abstraction relativement bas. En effet, à en croire O'Regan et Noë, même les traits les plus caractéristiques des modalités sensorielles, comme les saccades visuelles, sont comptées au nombre des contingences sensorimotrices. Par cette remarque, les auteurs s'inquiètent que les changements les plus mineurs dans notre appareillage perceptuel feraient en sorte que nous avons une expérience différente.

En guise de réponse, les énaïvistes précisent que la perception n'est pas caractérisée par le système perceptuel physique en tant que tel (i.e. l'organe concerné, par exemple, l'oeil dans le cas de la vision), mais par les contingences sensorimotrices c'est-à-dire par des *lois* qui peuvent être exprimées en termes mathématiques. En ce sens, les critères distinguant les modalités sensorielles les unes des autres sont formulées en termes fonctionnels qui sont relativement abstrait de l'organe physique devant donner lieu à la perception réelle. Les partisans de l'approche énaïve défendent ce point en faisant référence à la vaste littérature sur la substitution sensorielle.

Pour mieux illustrer la plasticité dont il est question, considérons le cas des prothèses visuelles élaborées par Paul Bach-y-Rita et ses collaborateurs depuis les années soixantes-dix, prothèses connues sous le nom de *tactile-vision substitution system* (TVSS) (pour un aperçu, c.f. Bach-y-Rita, Tyler & Kaczmarek 2003). On observe aujourd'hui plusieurs cas de substitution sensorielle grâce au progrès technologique, mais des phénomènes similaires à la substitution sensorielle peuvent se manifester

dans des situations aussi banales que celle d'une personne non voyante qui trouve son chemin à l'aide d'une canne et qui, grâce à celle-ci « perçoit une marche, un pied, une flaque d'eau » alors que « pendant ces tâches perceptuelles elle est inconsciente [*unaware*] des sensations dans sa main (là où se trouvent les senseurs biologiques). Plutôt, elle perçoit les éléments de son environnement comme étant des images mentales qui sont dérivées d'informations tactiles dont l'origine est l'extrémité de sa canne. » (ibid., p. 286).

Le TVSS fonctionne à partir du même principe, et a pour objectif, à la base, de permettre à des aveugles de voir à l'aide d'une caméra qui est placée sur des lunettes. La stimulation visuelle reçue par la caméra est ensuite transformée et envoyée sur une grille de pixel qui est localisée soit sur le dos, soit sur la langue du sujet. Lorsque celui-ci est libre dans ses déplacements, c'est-à-dire libre de contrôler l'influx sensoriel à l'aide de ses mouvements, il arrive après un certain temps d'adaptation à faire l'expérience du monde en trois dimensions.

Les énaïvististes insistent : la perception par TVSS n'est pas une forme de toucher, mais serait plutôt une forme de perception visuelle (amoindrie certes), et ce, même si la stimulation sensorielle se fait sur la peau plutôt que sur la rétine. Ils défendent, en effet, que cette forme de perception distale est une forme de vision à proprement parler, puisqu'elle est régie par les mêmes régularités sensorimotrices que la vision normale. En ce sens, O'Regan et Noë affirment que les travaux de Bach-y-Rita « constitue une vérification empirique » de la principale proposition de l'approche énaïvative (2001, p. 958).

Mentionnons cependant que cette interprétation des travaux de Bach-y-Rita est sujette à controverses. Jesse Prinz, par exemple, remarque que les prothèses visuelles élaborées par Bach-y-Rita permettent effectivement aux sujets de « faire des inférences » sur l'emplacement des objets dans l'espace à l'aide d'information tactile,

mais cela ne veut pas nécessairement dire qu'elle donne lieu à de la perception visuelle à proprement parler : peut-être que les sujets ne font que « deviner » où se trouvent les objets perçus dans l'espace (2006). Cette opposition paraît cependant peu convaincante. En effet, comme on peut le voir dans de nombreux vidéos disponibles sur Internet, ceux qui essaient ce type de prothèses visuelles sont étonnés et décrivent vraiment leur expérience en termes visuels. Cela concorde d'ailleurs avec les témoignages qu'ils donnent de leur expérience.

Par ailleurs, le type de fonctionnalisme adopté par les énaïvistes pose aussi problème, puisqu'ils ne précisent pas à quel niveau d'abstraction les modalités sensorielles doivent être définies. Est-ce que les saccades visuelles comptent comme un aspect essentiel de la vision humaine ? Le cas échéant, la perception par TVSS n'est pas une forme de vision. Faute de circularité, la théorie énaïve ne peut individuer les modalités sensorielles, ni par l'organe impliqué, ni par la zone cérébrale activée, ni par le type d'expérience *subjective* encourue (un qualia visuel versus un qualia auditif, par exemple). On constate ici un point d'ombre sur lequel les énaïvistes devront fournir davantage de détail.

Dire qu'un système est une forme de vision s'il est contraint par certaines contingences sensorimotrices, c'est une chose, ils doivent donner un critère pour déterminer quelles contingences sont pertinentes (et combien) pour qu'un processus soit considéré comme étant visuel. La perception par TVSS, par exemple, ne permet pas de percevoir la couleur : peut-on considérer que c'est une forme de vision ? Il est vrai que plusieurs animaux sont bichromatiques et ne perçoivent pas la couleur de la même manière que nous; cependant, à première vue, la perception de la couleur est un élément important de la vision humaine normale. Il est donc légitime de se demander si la perception par TVSS est une forme de perception visuelle. Ce problème indique que les énaïvistes doivent encore nous donner un critère indépendant pour distinguer les modalités sensorielles les unes des autres.

4.6 L'approche énaactive et la conscience

Dans la section précédente, j'ai introduit la distinction entre les déficits d'explication comparatifs et le déficit d'explication absolu, de même que l'hypothèse énaactive s'y rapportant. On peut cependant douter que l'approche énaactive puisse être considérée comme une théorie de la conscience en bonne et due forme, et ce, pour deux raisons.

D'une part, si l'approche énaactive permet d'expliquer l'expérience *visuelle*, il n'est pas certain qu'elle puisse expliquer l'expérience d'autres modalités sensorielles, et il est encore plus douteux qu'elle puisse rendre compte de la douleur, un cas paradigmatique d'expérience consciente. S'il est vrai, tel que je l'ai dit dans le chapitre précédent, qu'il ne faut pas refuser l'énaactivisme parce qu'il entre en conflit avec nos intuitions (pré)théoriques, il n'en demeure pas moins que sur cet enjeu, le fardeau de la preuve repose du côté des énaactivistes. Si l'approche énaactive est une théorie qui doit couvrir la totalité des expériences conscientes, alors ses défenseurs doivent expliquer comment des savoir-faire sensorimoteurs peuvent jouer un rôle dans notre expérience de la douleur.

Ici, on pourrait certes faire valoir que l'énaactivisme est une théorie de l'expérience *perceptuelle* et non pas de la conscience en général; le fait qu'elle ne puisse rendre compte de la douleur ne serait donc pas une opposition concluante. En un sens, cela est vrai. Cependant, un des principaux « arguments de vente » en faveur de l'énaactivisme consiste à dire que cette approche permet de résoudre les mystères entourant la conscience. Dans cette perspective, il est difficile de croire qu'il n'existe pas de qualia visuel, mais qu'il existe par contre des qualia pour la douleur. Autrement dit, à moins de fournir un argument convaincant en ce sens, il semblerait qu'une théorie satisfaisante *de la conscience* doive rendre compte à la fois de l'aspect subjectif de la perception et de l'aspect subjectif de la douleur.

D'autre part, dans plusieurs passages⁴⁹ O'Regan et Noë suggèrent que pour être conscient de quelque chose, il ne suffit pas à l'organisme d'exercer sa maîtrise des contingences sensorimotrices. Les attentes générées par ces savoir-faire sensorimoteurs doivent ensuite être recrutées par les systèmes cognitifs responsables de la pensée et de la planification. O'Regan et Noë font cette affirmation dans un passage où les auteurs discutent de la perception de la couleur d'une voiture :

But if you should turn your attention to the color of the car ahead of you, and think about it, or discuss it with your friend, or use the knowledge of the car's color to influence decisions you are making, then, we would say, you are aware of it. For a creature (or a machine for that matter) to possess visual awareness, what is required is that, in addition to exercising the mastery of the relevant sensorimotor contingencies, it must make use of this exercise for the purposes of thought and planning. (O'Regan et Noë 2001, p. 944, c.f. aussi p. 1011–3).

Des passages comme celui-là suggèrent que les savoir-faire sensorimoteurs auxquels les énéactives font si souvent référence, ne suffisent pas à expliquer la conscience : d'autres processus cognitifs ont aussi un rôle à jouer. O'Regan et Noë reconnaissent ainsi le rôle de l'*attention* lorsqu'il est question de déterminer pourquoi nous sommes conscients de certaines choses, mais pas de d'autres. Dans leur publication datant de 2001, les auteurs adhèrent d'ailleurs à la majorité des chercheurs oeuvrant dans le domaine et qui expliquent la cécité au changement en termes attentionnels (p. 954).

Dans une perspective similaire, on peut par contre déplorer le fait que O'Regan, Noë et Hurley ne soient pas plus explicites sur la relation entre la conscience et l'attention,

⁴⁹ Un passage où cela est particulièrement évident est celui-ci, où les auteurs discutent la perception de la couleur d'une voiture : « But if you should turn your attention to the color of the car ahead of you, and think about it, or discuss it with your friend, or use the knowledge of the car's color to influence decisions you are making, then, we would say, you are aware of it. For a creature (or a machine for that matter) to possess visual awareness, what is required is that, in addition to exercising the mastery of the relevant sensorimotor contingencies, it must make use of this exercise for the purposes of thought and planning. » (O'Regan et Noë 2001, p. 944). Voir aussi p. 1011–3.

question qui a suscité une littérature massive en sciences cognitives et en neurosciences. Tallon-Beaudry a par exemple défendu que si nous concevons l'attention et la conscience comme intrinsèquement liées, alors le rôle fonctionnel de cette dernière dans l'architecture cognitive, de même que son avantage évolutionnaire, concernera d'avantages les fonctions exécutives; au contraire, si elles sont distinctes, alors la conscience ne devra pas être définie pour son rôle cognitif, mais plutôt par son caractère phénoménal (2012, p. 397). De plus, toujours selon Tallon-Baudry, nous pouvons retrouver dans la littérature au moins trois manières différentes de concevoir la relation entre attention et conscience. D'abord, selon la *gateway hypothesis*, l'attention facilite, voire est nécessaire, à la conscience; ensuite, selon la *reverse dependence hypothesis* le fait qu'un stimulus a été détecté, ou pas, peut mobiliser différents processus attentionnels; enfin, selon la *cumulative influence hypothesis* est l'hypothèse inverse, la conscience et l'attention sont deux mécanismes distincts, mais influenceraient tous les deux (peut-être à des degrés différents) la prise de décision et les rapports verbaux. Nous avons ici affaire à bref aperçu, mais on comprend tout de même que la relation entre l'attention et la conscience peut prendre une multitude de formes différentes.

Une autre critique adressée à l'approche éactive concerne le domaine d'application de cette théorie. Clark et Prinz, par exemple, soutiennent que l'approche éactive n'est pas une théorie visant à expliquer *pourquoi certains états mentaux deviennent conscients*, mais plutôt une théorie visant à expliquer pourquoi nos perceptions conscientes ont le contenu qu'elles ont plutôt qu'un autre (2012, p. 171). Andy Clark soutient aussi cette interprétation (2006). J'adhère ici à cette interprétation, dans la mesure où elle concorde avec l'accent que mettent les Hurley et Noë (2003) sur les déficits d'explication relatifs. En effet, comme je l'ai déjà mentionné, une des forces de cette approche est de fournir une hypothèse intéressante expliquant pourquoi nous sommes conscients de rouge plutôt que de vert (déficit intramodal), ou encore

pourquoi nous sommes conscients auditivement plutôt que visuellement (déficit extramodal).

Les deux problèmes que je viens de soulever indiquent des points qui devront être précisés, si les partisans de l'approche énative espèrent que leur théorie soit considérée comme une alternative sérieuse aux autres théories de la conscience que l'on retrouve sur le marché.

4.7 Conclusion

Dans l'introduction de ce chapitre, j'ai affirmé que la stratégie énative consistait à adopter et radicaliser la thèse de Dennett selon laquelle les représentations subpersonnelles ne deviennent pas conscientes en parvenant à un point précis dans le cerveau, point où elles seraient ensuite représentées (c'est ce à quoi Dennett fait référence en parlant de « *Théâtre Cartésien* »). Dans les faits, on constate que cette stratégie consiste à abaisser notre fardeau explicatif. Comme je l'ai mentionné à de multiples reprises, les énavivistes font de même en niant que la conscience soit l'occurrence de quelque chose dans le cerveau. En effet, pour les énavivistes, la conscience est un processus continu impliquant le cerveau, le corps et l'environnement de l'organisme percevant. Il n'y a donc pas d'endroit ou de moment précis où nous pouvons espérer trouver un corrélat neuronal à la conscience.

Plusieurs choses ont été dites dans ce chapitre. D'abord, nous avons exposé les grandes lignes de la conception traditionnelle en science de la vision et du programme de recherche visant à identifier les corrélats neuronaux. Nous avons ensuite présenté les divers aspects de la théorie énative de l'expérience perceptuelle, de même que les arguments en faveur de celle-ci. Enfin, nous avons tenté d'évaluer le réel domaine d'application de cette théorie.

Cela dit, l'antireprésentationnalisme propre à la théorie énaïve était en partie déjà présent chez Dennett, Ballard et Brooks; alors que les principaux arguments pour l'externalisme véhiculaire avaient déjà été proposés dans le cadre des théories de l'esprit incarné. En effet, Dennett défendait déjà l'idée d'après laquelle le détail de notre expérience perceptuelle n'est ni dans notre esprit, ni dans notre cerveau, mais se trouve plutôt *dans le monde*⁵⁰. À ce point, il est en effet légitime de se demander en quoi consiste l'originalité de l'approche énaïve telle qu'exposée dans les pages qui ont précédées, puisqu'elle arrive à des propositions similaires. Sur ce point, j'ai fait valoir que l'originalité de l'approche énaïve ne réside pas dans sa conception active de la vision, ni dans son antireprésentationnalisme, mais consiste plutôt en l'utilisation de la notion de connaissance sensorimotrice afin *d'articuler autrement* des problèmes philosophiques classiques, notamment celui du déficit d'explication touchant la conscience. En effet, l'idée principale de l'approche énaïve est que nous pouvons éviter le déficit d'explication si nous évitons de concevoir la conscience comme étant l'occurrence de quelque chose dans l'esprit ou dans le cerveau.

Enfin, ce chapitre a soulevé le fait qu'une part non négligeable des difficultés de l'approche énaïve est dû au fait qu'il est ardu de déterminer la place qu'elle occupe vis-à-vis de la conscience que l'on retrouve actuellement sur le marché, parce que les énaïvistes sont réticents lorsqu'il est question de parler en termes d'états mentaux ou d'activité neurologique. Cela est d'autant plus problématique, que la conception énaïve de la conscience donne un rôle central à l'attention, qui est un processus cognitif se déroulant dans le cerveau.

⁵⁰ C.f. « and no matter how vivid your impression is that you see all that detail, the detail is in the world, not in your head. » (Dennett 1991, p. 355).

CONCLUSION

L'approche énative, une théorie de la conscience ?

Nous pouvons répartir les diverses approches au déficit d'explication de la conscience en trois catégories. Dans la première catégorie, nous retrouvons les approches selon lesquelles nous ne pourrions jamais donner une explication en termes naturalistes à la conscience : le déficit d'explication est tout simplement *insoluble* (McGinn 1991, Levine 1983, 1993, Nagel 1974). Dans la deuxième catégorie, nous retrouvons des approches selon lesquelles nous pouvons expliquer la conscience en la réduisant à quelque chose que nous comprenons mieux, que ce soit un *corrélat cognitif* (e.g. l'attention, Prinz 2012), un *corrélat neuronal* (e.g. une oscillation particulière dans le cortex cérébral, Crick et Koch 1990), ou la *relation entre certains états mentaux* (e.g. la relation entre un jugement assertif d'ordre supérieur et un autre état mental, Rosenthal 1995). Comme je l'ai mentionné à plusieurs reprises, ces approches adhèrent toutes à une forme particulière d'*internalisme* selon lequel il existe un substrat minimal (dans le cerveau) suffisant pour donner lieu à la conscience. Enfin, parmi la troisième catégorie d'approches se trouve l'énaturalisme, théorie selon laquelle la conscience ne peut être expliquée en faisant uniquement référence à l'occurrence de quelque chose dans l'esprit ou dans le cerveau, puisque celle-ci est *constituée* par l'organisme et l'environnement avec lequel il interagit.

Dans le troisième chapitre, j'ai soutenu que l'externalisme énatif n'est pas nécessairement incompatible avec le modèle d'explication qui prédomine en sciences cognitives et en neurosciences. Même si cette forme d'externalisme est contre-intuitive, il n'en demeure pas moins que nous ne pouvons rejeter cette thèse sans

fournir un argument en bonne et due forme. Autrement dit, puisque nous n'avons aucune idée de la manière dont l'activité neuronale peut donner lieu à notre expérience, nous devons avoir de bonnes raisons pour rejeter la thèse selon laquelle le substrat de la conscience se trouve partiellement dans notre environnement.

Cela dit, certaines difficultés subsistent et doivent recevoir une réponse si nous souhaitons que l'énactivisme soit pris au sérieux par la communauté scientifique, notamment par les chercheurs oeuvrant en sciences cognitives et en neurosciences. En effet, si l'explication mécaniste n'exclut en rien que la cognition puisse être expliquée par un mécanisme qui s'étend en dehors du cerveau, la *localisation* et la *décomposition* restent les deux heuristiques de recherche principales de ce modèle d'explication (Bechtel & Richardson 2010). En ce sens, les partisans de l'énactivisme doivent nous indiquer comment le mécanisme donnant lieu à la conscience peut être *décomposé* et comment ses composantes peuvent être *localisées* physiquement (dans notre cerveau ou hors de celui-ci). Le fait que les énéactivistes ne nous donnent aucune piste pour identifier de telles composantes suggère qu'en fait, ils défendent une thèse plus forte que celle selon laquelle la conscience est le résultat du système complexe composé de l'organisme et de son environnement. Par là, j'entends la thèse selon laquelle la conscience serait une propriété se situant au niveau personnel et qui ne peut être expliqué complètement en termes de processus cognitifs ou neuronaux.

Dans le quatrième chapitre, j'ai exposé les arguments principaux en faveur de la théorie énéactive de l'expérience perceptuelle et de la conscience, de même que les nombreux problèmes auxquels fait face cette théorie. D'après cette théorie, la perception est constituée de deux éléments : d'une part, de l'organisme qui *explore* et *sonde* les structures d'information de son environnement à l'aide d'*habiletés* sensorimotrices bien spécifiques; d'autre part, d'un ensemble d'*attentes* à l'égard de ce que l'organisme devrait percevoir s'il se déplace d'une façon donnée, attentes qui sont

elles-mêmes possibles parce que l'organisme dispose de certaines connaissances sensorimotrices.

Depuis le début des années quatre-vingt-dix, la conception active de la perception n'est plus, en soi, aussi controversée. Cela dit, je me suis efforcé de faire montrer que l'approche énative ne se réduit pas à une théorie de la vision. Ce serait plutôt comme théorie de l'expérience perceptuelle qu'elle démontre toute son originalité et son potentiel.

Malgré cela, cette approche fait face à de nombreuses limitations : d'une part, si la conception énative de l'expérience perceptuelle visuelle est convaincante, elle l'est beaucoup moins lorsqu'on l'applique aux autres modalités sensorielles de même qu'à la douleur; d'autre part, en tant que théorie de la conscience, elle accorde une place centrale à l'attention sans nous en dire davantage sur le rôle que l'attention doit jouer vis-à-vis de la conscience. Ces problèmes nous montrent clairement les limites que cette approche devra éventuellement surmonter.

Cela étant dit, où en sommes-nous maintenant ? L'objectif de ce mémoire était de mettre en contexte philosophique plus large la théorie énative de l'expérience perceptuelle. En ce sens, dans ce mémoire j'espère avoir contribué à l'avancement des connaissances sur deux questions différentes.

D'une part, j'ai situé la théorie énative dans le contexte plus large « du problème » de la conscience et des autres théories qui ont été proposées pour le résoudre. À bien des égards, il fut jusqu'à présent difficile d'évaluer la plausibilité de l'approche énative parce que les auteurs qui défendent cette approche tendent à ne pas confronter directement et dans le détail les approches « internalistes ».

D'autre part, en ayant proposé une analyse plus détaillée de ce qu'on entend par « problème de la conscience », j'ai pu établir clairement que l'approche énaïve n'est pas une théorie de la conscience phénoménale en tant que telle, mais qu'elle est plutôt une hypothèse visant à combler les déficits d'explication relatifs. En ce sens, bien qu'intéressante, cette théorie ne permet pas, pour l'instant, à dissoudre le mystère entourant les qualia.

BIBLIOGRAPHIE

- Adams, F., & Aizawa, K. 2010. « Defending the bounds of cognition ». Dans Menary 2010, 67-80.
- Akins, K. 1996. *Perception : Vancouver Studies in Cognitive Science*, Oxford University Press.
- Akins, K, et S. Winger. 1996. « Ships in the Night: Churchland and Ramachandran on Dennett's Theory of Consciousness. ». Dans Akins 1996.
- Baars, B. J. 1988. *A Cognitive Theory of Consciousness*. Cambridge University Press.
- Baars, B. J., & Gage, N. M. 2010. *Cognition, brain, and consciousness: Introduction to cognitive neuroscience*. Academic Press.
- Bach-y-Rita, P., Tyler, M. E., & Kaczmarek, K. A. 2003. « Seeing with the brain » *International journal of human-computer interaction*, 15(2), 285-295.
- Ballard, D. 1991. « Animate vision ». *Artificial intelligence*, 48(1), 57-86.
- Ballard, D. 1996. « On the Function of Visual Representation ». Dans Akins 1996.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. 1997. « Deictic codes for the embodiment of cognition ». *Behavioral and Brain Sciences*, 20(04), 723-742.
- Beer, R. D. 2000. « Dynamical approaches to cognitive science. » *Trends in Cognitive Sciences*, 4, 91-99.
- Bechtel, W., et A. A. Abrahamsen. 1990. *Connectionism and the Mind: Introduction to Parallel Processing in Networks*. Blackwell Publishers.
- Bechtel, W. P. 1994. « Levels of Description and Explanation in Cognitive Science ». *Minds and Machines* 4(1), 1-25.
- Bechtel, W. 1998. « Representations and cognitive explanations: Assessing the dynamicist's challenge in cognitive science ». *Cognitive Science* 22(3), 295-317.
- Bechtel, W. 2002. « Decomposing the Brain: A Long Term Pursuit ». *Brain and Mind* 3(1), 229-242.
- Bechtel, W. 2008 « Mental Mechanisms: Philosophical Perspectives on the Sciences of Cognition and the Brain ».
- Bechtel, W. 2009. « Looking down, around, and up: Mechanistic explanation in psychology ». *Philosophical Psychology* 22 (5): 543-564.
- Bechtel, W., & Richardson, R. C. 2010. *Discovering complexity*. The MIT Press.

- Bechtel, W. P., et R. S. Stufflebeam. 2001. « Epistemic Issues in Procuring Evidence About the Brain: The Importance of Research Instruments and Techniques ». *Philosophy and the Neurosciences: A Reader*. Blackwell.
- Bennett, M. R., Hacker, P. M. S., & Bennett, M. R. 2003. *Philosophical foundations of neuroscience*. Oxford: Blackwell.
- Bennett, M., Dennett, D., Hacker, P., & Searle, J. 2007. *Neuroscience and Philosophy: Brain, Mind, and Language*. Columbia University Press, New York.
- Bickle, J. 2003. *Philosophy and Neuroscience: a Ruthlessly Reductive Account*. Dordrecht; Boston: Kluwer Academic Publishers.
- Block, N. 1995. « On a Confusion About a Function of Consciousness ». *Brain and Behavioral Sciences* vol. 18, 227-247.
- Block, N. « Action in Perception by Alva Noë ». *The Journal of Philosophy* 102(5), 259-272.
- Block, N., Flanagan, O. J., & Güzeldere, G.. 1997. *The nature of consciousness: Philosophical debates*. Cambridge, MA: MIT Press.
- Brook, A., & Akins, K. 2005. *Cognition and the brain: The philosophy and neuroscience movement*. Cambridge University Press.
- Brooks, R. 1991a. « Intelligence Without Representation ». *Artificial Intelligence* 47: 139-159.
- Brooks, R. 1991b. « Intelligence Without Reason ».
- Burge, T. 1979. « Individualism and the Mental. » Réédité dans Burge 2007.
- Burge, T. 1986. « Individualism and psychology. » *The Philosophical Review*, 95(1), 3-45.
- Burge, T. 2006. Postscript to "Individualism and the Mental", Réédité dans Burge 2007.
- Burge, T. 2007. *Foundations of mind*. Oxford University Press.
- Carruthers, P.. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge University Press.
- Chalmers, D. 1995/1999. « Facing Up to the Problem of Consciousness ». *Explaining Consciousness: The Hard Problem*, édité par Jonathan Shear, MIT Press, p. 9-30. Cambridge: A Bradford Book.
- Chalmers, D. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. 1999. « First-Person Methods in the Science of Consciousness ». *Consciousness Bulletin*.

- Chalmers, D. 2000. « What is a Neural Correlate of Consciousness? » *Neural Correlates of Consciousness*. MIT Press.
- Chalmers, D. 2002. « Does Conceivability Entail Possibility? » *Conceivability and Possibility*, 145–200. Oxford University Press.
- Chalmers, D. 2004. « How Can We Construct a Science of Consciousness? » *The Cognitive Neurosciences Iii*. Mit Press.
- Chomsky, N. 1966. *Cartesian linguistics: A chapter in the history of rationalist thought*. Cambridge University Press.
- Churchland, P. S. 1986. *Neurophilosophy: Toward a unified science of the mind-brain*. MIT press.
- Churchland, P. M., et P. S. Churchland. 1994. « Intertheoretic Reduction: A Neuroscientist's Field Guide ». *The Mind-Body Problem: A Guide to the Current Debate*. Cambridge: Blackwell.
- Churchland, P. S., V. S. Ramachandran, et Terrence J. Sejnowski. 1993. « A Critique of Pure Vision ». In *Large-scale neuronal theories of the brain*. MIT Press.
- Churchland, P. S., et V. S. Ramachandran. 1993. « Filling In: Why Dennett is Wrong ». *Dennett and His Critics*. Blackwell.
- Churchland, P. M., and P. S. Churchland. 1990. « Intertheoretic Reduction: A Neuroscientist's Field Guide ». In *Seminars in the Neurosciences 2*:249–56.
- Clark, A. 1997. *Being there: Putting brain, body, and world together again*. MIT press.
- Clark, A. 1998. « Twisted tales: Causal complexity and cognitive scientific explanation. » *Minds and Machines*, 8(1), 79-99.
- Clark, A. « Memento's revenge: The extended mind, extended. ». In Menary 2010.
- Clark, A. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension: Embodiment, Action, and Cognitive Extension*. Oxford University Press.
- Clark, A., & Chalmers, D. 1998. *The extended mind*. Réédité dans Menary 2010.
- Clark, A. and Toribio, J. 2001. « Sensorimotor chauvinism?: Commentary on O'Regan, J.K., and Noë, A. A sensorimotor approach to vision and visual consciousness, » *Behavioral and Brain Sciences*, 24(5), 979-980.
- Cohen, M. A., & Dennett, D. C. 2011. « Consciousness cannot be separated from function ». *Trends in cognitive sciences*, 15(8), 358-364.
- Crane, T. 1992. *The Contents of Experience: Essays on Perception*. Cambridge University Press.
- Crane, T. *Introduction*, Dans Crane 1992.

- Crane, T. 1998. *Intentionality as the mark of the mental*. Royal Institute of Philosophy Supplement, 229-252.
- Crane, T. *The Origin of Qualia*, in Crane et Patterson 2000.
- Crane, T. 2011. « The Problem of Perception », *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Edward N. Zalta (ed.),
URL = <<http://plato.stanford.edu/archives/spr2011/entries/perception-problem/>>.
- Crane, T. & Patterson, S. 2000. *History of the mind-body problem*. Routledge.
- Crane, T. S. 2000. *Introduction*. Dans Crane & Patterson 2000.
- Craver, C. 2003. « The Making of a Memory Mechanism ». *Journal of the History of Biology* 36(1): 153–95.
- Craver, C. 2006. « When Mechanistic Models Explain ». *Synthese* 153 (3): 355–376.
- Craver, C. 2007. *Explaining the Brain*. Oxford University Press, Oxford.
- Craver, C. 2009. « Mechanisms and Natural Kinds ». *Philosophical Psychology* 22 (5): 575–594.
- Craver, C. et W. Bechtel. 2007. « Top-Down Causation Without Top-Down Causes ». *Biology and Philosophy* 22 (4): 547–563.
- Craver, C. et L. Darden. 2001. « Discovering Mechanisms in Neurobiology: The Case of Spatial Memory ». *Theory and Method in Neuroscience*. Pittsburgh: University of Pitt Press.
- Crick, F. 1995. *Astonishing hypothesis: The scientific search for the soul*. Scribner Book Company.
- Cummins, R. 2000. « How does it work?" versus" what are the laws?": Two conceptions of psychological explanation ». *Explanation and cognition*, 117-144.
- Cussins, A. 2003. « Content, conceptual content, and nonconceptual content. ». Dans Gunther (Éd.) (2003). *Essays on nonconceptual content*. MIT press.
- Davidson, D. 1973. « On the Very Idea of a Conceptual Scheme ». *Proceedings and Addresses of the American Philosophical Association* 47, 5-20.
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford University Press, trad. fr. *Actions et événements*, P.U.F. (Épiméthée), 1993.
- Davis, G., & Driver, J. (2003). *Effects of modal versus amodal completion upon visual attention: a function for filling-in*. Dans Pessoa et De Weerd 2003, 128-150.
- Dehaene, S., & Naccache, L. (2001). « Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. » *Cognition*, 79(1), 1-37.

- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). « Conscious, preconscious, and subliminal processing: a testable taxonomy. » *Trends in cognitive sciences*, 10(5), 204-211.
- Dennett, D. C. 1969. *Content and Consciousness*. Routledge.
- Dennett, D. C. 1971. « Intentional Systems ». *Journal of Philosophy* 68 (February): 87–106.
- Dennett, D. C. 1976. « Conditions of personhood. In What Is a Person? » . Réédité dans Dennett 1981.
- Dennett, D. C. 1978. « Toward a Cognitive Theory of Consciousness ». *Minnesota Studies in the Philosophy of Science* 9.
- Dennett, D. C. 1981. *Brainstorms: Philosophical essays on mind and psychology*. The MIT Press.
- Dennett, D. C. 1982. « How to Study Human Consciousness Empirically, or, Nothing Comes to Mind ». *Synthese* 53(2): 159-180.
- Dennett, D. C. 1983. « Styles of Mental Representation ». *Proceedings of the Aristotelian Society* 83: 213–226.
- Dennett, D. C. 1987. *The Intentional Stance*. MIT Press.
- Dennett, D. C. 1991a. *Consciousness Explained*. Penguin.
- Dennett, D. C. 1991b. « Real Patterns ». *Journal of Philosophy* 88(1): 27–51.
- Dennett, D. C. 1992a. « Filling in Versus Finding Out: A Ubiquitous Confusion in Cognitive Science ».
- Dennett, D. C. 1995. « Do Animals Have Beliefs? » *Comparative Approaches to Cognitive Science*. MIT Press.
- Dennett, D. C. 1996. « Seeing is Believing - Or is it ? ». Dans Akins 1996.
- Dennett, D. C. 2001. « The Fantasy of First-Person Science ».
- Dennett, D. C. 2003. « Who's on First? Heterophenomenology Explained ». *Journal of Consciousness Studies* 10 (9): 19–30.
- Dennett, D. C. 2007a. « Heterophenomenology Reconsidered ». *Phenomenology and the Cognitive Sciences*, 6.
- Dennett, D. C. 2007b. *Philosophy as Naive Anthropology: Comment on Bennett and Hacker*. Dans Bennett, Dennett, Hackett et Searle 2007.
- Dennett, D. C., et M. Kinsbourne. 1992. « Time and the Observer: The Where and When of Consciousness in the Brain ». *Behavioral and Brain Sciences* 15: 183–201.

- Depraz, N., F. V., et P. Vermersch. 2003. *On Becoming Aware: A Pragmatics of Experiencing*. John Benjamins.
- Dretske, F. 1995. *Naturalizing the mind*. Bradford Books.
- Dretkse, F. 1996. « Phenomenal Externalism or If Meanings Ain't in the Head, Where Are Qualia?. » *Philosophical Issues*, 7, 143-158.
- Dreyfus, H. 1972. *What Computers Can't Do*. Harper and Row.
- Dreyfus, H., et Sean D. Kelly. 2007. « Heterophenomenology: Heavy-Handed Sleight-of-Hand ». *Phenomenology and the Cognitive Sciences* 6 (1-2): 45–55.
- Edelman, G. M. 1991. *Bright Air, Brilliant Fire*. New York: Basic.
- Evans, G. (1982). *The Varieties of Reference*. Oxford: Clarendon Press.
- Fodor, J., et Z. Pylyshyn. 1988. « Connectionism and Cognitive Architecture: a Critical Analysis ». *Cognition* 28 (1-2): 3-71.
- Fodor, J. 1974. « Special Sciences ». *Synthese* 28(2): 97–115.
- Fodor, J. 1983. *Modularity of Mind*. Mit Press.
- Fodor, J. 2000. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. MIT Press.
- Fodor, J. A., & Lepore, E. 1992. *Holism: A shopper's guide*. Oxford: Blackwell.
- Gallagher, S.. 2002. « Experimenting with Introspection ». *Trends in Cognitive Sciences* 6 (9): 374–375.
- Gallagher, S. 2003. « Phenomenology and Experimental Design: Toward a Phenomenologically Enlightened Experimental Science ». *Journal of Consciousness Studies* 10 (9-10): 85–99.
- Gallagher, S, et M. Overgaard. 2005. « Introspections Without Introspeculations ». In *Pain: New Essays on its Nature and the Methodology of its Study*. Cambridge Ma: Bradford Book/Mit Press.
- Gallagher, S. et J. B. Sorensen. 2006. « Experimenting with Phenomenology ». *Consciousness and Cognition* 15 (1): 119–134.
- Gallagher, S. et F. Varela. 2003. « Redrawing the Map and Resetting the Time: Phenomenology and the Cognitive Sciences ». *Canadian Journal of Philosophy* 29, 93-132.
- Gallagher, S., et D. Zahavi. 2008. *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. Routledge.
- Gazzaniga, M. S. 1988. *Mind matters*. Houghton Mifflin.
- Gibson, J. 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin.

- Glennan, S. S. (1996). « Mechanisms and the nature of causation ». *Erkenntnis*, 44(1), 49-71.
- Greenwood, J. D. (2009). *A Conceptual History of Psychology*. McGraw-Hill Higher Education.
- Grimes, J. 1996. « On the Failure to Detect Changes in Scene across Saccades ». In Akins 1996.
- Harman, G. 1990. *The Intrinsic Quality of Experience*. Dans Block et al. 1997.
- Haugeland, J. 2000. *Having thought: Essays in the metaphysics of mind*. Harvard University Press.
- Held, R., and A. Hein. 1963. « Movement produced stimulation in the development of visually guided behavior. » *Journal of Comparative and Physiological Psychology* 56: 873-876.
- Hempel, C.. 1965. « Aspects of Scientific Explanation. » *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.
- Hurley. 1998. *Consciousness in Action*. Harvard University Press.
- Hurley, S. 2010. « Varieties of Externalism ». Dans Menary 2010.
- Hurley, S., et A. Noë. 2003. « Neural Plasticity and Consciousness: Reply to Block ». *Trends in Cognitive Sciences* 7 (1): 342.
- Husserl, E. 1989. *La philosophie comme science rigoureuse*. Trad. Marc B. de Launay, Paris, Presses universitaires de France.
- Hutto, D. D., & Myin, E. 2012. *Radicalising Enactivism*. The MIT Press.
- Jack, A., et A. Roepstorff. 2002. « Introspection and Cognitive Brain Mapping: From Stimulus-response to Script-report ». *Trends in Cognitive Sciences* 6 (8): 333-339.
- Jack, A., et A. Roepstorff. 2004. « Trust or Interaction? Editorial Introduction ». *Journal of Consciousness Studies* 11 (7-8).
- Jack, A. I., & Prinz, J. J. 2004. « Searching for a scientific experience ». *Journal of Consciousness Studies*, 11(1), 51-55.
- Jackson, F. 1982. « Epiphenomenal qualia ». *The Philosophical Quarterly*, 32(127), 127-136.
- Kirchhoff, M. D. 2012. « Extended cognition and fixed properties: steps to a third-wave version of extended cognition. » *Phenomenology and the Cognitive Sciences*, 11(2), 287-308.
- Kriegel, U. 2009. *Subjective Consciousness: A Self-Representational Theory*. Oxford University Press.

- Leopold, D. A., and Logothetis, N. K. 1996. « Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry ». *Nature* 379: 549–553.
- Logothetis, N., and Schall, J. 1989. « Neuronal correlates of subjective visual perception ». *Science* 245: 761–763.
- Leopold, David A., Alexander Maier, et Nikos K. Logothetis. 2003. « Measuring Subjective Visual Perception in the Nonhuman Primate ». *Journal of Consciousness Studies* 10 (9-10): 115–130.
- Levine, J. 1983. « Materialism and qualia: the explanatory gap ». *Pacific Philosophical Quarterly*, 64: 354-361.
- Levine, J. 1993. « On leaving out what it's like ». In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Levine, J. 2001. *Purple Haze: The Puzzle of Conscious Experience*. Cambridge, Mass: The MIT Press.
- Loar, B. 1990. « Phenomenal properties ». Dans J. Tomberlin (Éd.), *Philosophical Perspectives: Action theory and philosophy of mind*. Atascadero, CA: Ridgeview Publishing.
- Lowe, E. J. 1992. *Experience and its objects*. Dans Crane 1992, 79-104.
- Lutz, A. 2002. « Toward a Neurophenomenology as an Account of Generative Passages: A First Empirical Case Study ». *Phenomenology and the Cognitive Sciences* 1 (2): 133–67.
- Lutz, A, J.-P. Lachaux, J. Martinerie, et F. J. Varela. 2002. « Guiding the Study of Brain Dynamics by Using First-person Data: Synchrony Patterns Correlate with Ongoing Conscious States During a Simple Visual Task ». *Proceedings of the National Academy of Sciences* 99 (3): 1586-1591. doi:10.1073/pnas.032658199.
- Lycan, W. G. 1987. *Consciousness*. MIT Press.
- Lycan, W. G. 1996. *Consciousness and Experience*. MIT Press.
- Lycan, W. G. 2001. « The Case for Phenomenal Externalism ». *Noûs*, 35(s15), 17-35.
- Lycan, W. G. 2004. « The Superiority of HOP to HOT ». Dans Gennaro, R. J. (Éd.). (2004). *Higher-order theories of consciousness: An anthology*. John Benjamins Publishing.
- Marr, D. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. The MIT Press.
- Martin, M. 1992. *Sight and touch*. Dans Crane 1992.
- Maturana, H., & Varela, F. 1980. *Autopoiesis and cognition*, D. Reidel, Dordrecht, Holland.

- Maudlin, T. 2003. « Distilling metaphysics from quantum physics ». MJ Loux, DW Zimmerman, *The Oxford Handbook of Metaphysics*, Oxford University Press, Oxford, 461-487.
- McCauley, R. N., & Bechtel, W. (2001). « Explanatory pluralism and heuristic identity theory. » *Theory & Psychology*, 11(6), 736-760.
- McGinn, C. 1989. « Can we solve the mind-body problem? » *Mind*, 98: 349-66
- McGinn, C. 1991. *The Problem of Consciousness*. Oxford: Blackwell.
- McGinn, C. 1993. *Problems in Philosophy: The Limits of Inquiry*, Blackwell.
- McGovern, K., & Baars, B. J. (2007). Cognitive theories of consciousness. The Cambridge handbook of consciousness, 177-205.
- MacKay, D. M. 1962. « Theoretical models of space perception. » In *Aspects of the Theory of Artificial Intelligence*, ed. C. A. Muses, 83-104. New York: Plenum
- Mackay D. 1967. « Ways of looking at perception. » In Wathen-Dunn W (ed) *Models for the perception of speech and visual form*. MIT Press, Cambridge, MA.
- Mack, A., & Rock, I. 1998. *Inattention blindness*. The MIT Press.
- Macpherson, F. 2011. *The senses: Classic and contemporary philosophical perspectives*. Oxford University Press.
- McDowell, J. (1994). *Mind and world*. Harvard University Press.
- Menary, R. 2006a. « Attacking the Bounds of Cognition ». *Philosophical Psychology* 19 (3): 329-344.
- Menary, R. 2006b. *Radical Enactivism: Intentionality, Phenomenology and Narrative: Focus on the Philosophy of Daniel D. Hutto*. Amsterdam: J Benjamins.
- Menary, R. 2007. *Cognitive integration: mind and cognition unbounded*, Basingstoke : Palgrave Macmillan
- Menary, R. 2010. *The Extended Mind*. Mit Press.
- Merleau-Ponty, M.. 1945. *Phénoménologie de la perception*. Gallimard.
- Metzinger, T.. 2000. *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. MIT Press.
- Nagel, E. 1961, *The Structure of Science: Problems in the Logic of Scientific Explanation*, New York: Harcourt Brace and World.
- Nagel, T. 1974. « What Is It Like to Be a Bat? » *The Philosophical Review* 83 (4): 435-450.
- Nagel, T. 1979a. *Mortal Questions*. Cambridge University Press.

- Nagel, T. 1979b. « Panpsychism ». In *Mortal Questions*. Cambridge University Press.
- Nagel, T. 1979c. « Subjective and Objective ». In *Mortal Questions*, 207–222. Cambridge University Press.
- Nagel, T. 1986. *The View From Nowhere*. Oxford University Press.
- Neisser, U.. 1999. *Ecological Psychology*. In Wilson & Keil 1999.
- Noë, A., et S. L. Hurley. 2003. « The Differential Brain in Action ». *Trends in Cognitive Sciences* 7 (5): 195–196.
- Noë, A.. 1997a. « Cartesian Psychology and Physical Minds ». *Philosophical Review* 106 (3): 434–436.
- Noë, A. 1997b. « Perception and Content ». *Behavioral and Brain Sciences* 20(1), 154–155.
- Noë, A. 1999. « Thought and Experience ». *American Philosophical Quarterly* 36 (3) (juillet 1): 257-265.
- Noë, A. 2001. « Experience and the Active Mind ». *Synthese* 129 (1) (octobre 1): 41-60.
- Noë, A. 2002a. « Is Perspectival Self-Consciousness Non-Conceptual? » *Philosophical Quarterly* 52 (207): 185–194.
- Noë, A. 2002b. « Is the Visual World a Grand Illusion? » *Journal of Consciousness Studies* 9 (5-6): 1–12.
- Noë, A. 2002c. « Is Perspectival Self-Consciousness Non-Conceptual? » *The Philosophical Quarterly* 52 207 : 185-194.
- Noë, A. 2003. « Causation and Perception: The Puzzle Unravelling ». *Analysis* 63 (2): 93–100.
- Noë, A. 2005. « Against Intellectualism ». *Analysis* 65 (4) (octobre 1): 278-290.
- Noë, A. 2006. *Action in Perception*. A Bradford Book.
- Noë, A. 2007a. « Inattentional Blindness, Change Blindness and Consciousness ». In *The Blackwell Companion to Consciousness*. Blackwell.
- Noë, A. 2007b. « The Critique of Pure Phenomenology ». *Phenomenology and the Cognitive Sciences* 6 (1-2).
- Noë, A. 2007c. « The Problem with the Picture Picture of Visual Experience: A Comment on Jacob and Jeannerod's Ways of Seeing ». *Dialogue* 46 (2): 347–351.
- Noë, A. 2008. « Précis of Action in Perception: Philosophy and Phenomenological Research ». *Philosophy and Phenomenological Research* 76 (3): 660–8211.
- Noë, A. 2009. *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons From the Biology of Consciousness*. Hill and Wang.

- Noë, A. 2012. *Varieties of Presence*. Harvard University Press.
- Noë, A., et S. Hurley. 2003. « The Deferential Brain in Action ». *Trends in Cognitive Sciences* 7 (5): 195–196.
- Noë, A., et K. J. O'Regan. 2000. « Perception, Attention, and the Grand Illusion ». *Psyche* 6 (15).
- Noë, A., et K. J. O'Regan. 2002. « On the Brain-Basis of Visual Consciousness: A Sensorimotor Account ». In *Vision and Mind: Selected Readings in the Philosophy of Perception*. MIT Press.
- Noë, A., et E. Thompson. 2002. *Vision and Mind: Selected Readings in the Philosophy of Perception*. MIT Press.
- Noë, A., et E. Thompson. 2004. « Are there Neural Correlates of Consciousness? ». *Journal of Consciousness studies*, 11(1), 3-28.
- O'Regan, J. K., et A. Noë. 2001a. « A Sensorimotor Account of Vision and Visual Consciousness ». In *Behavioral and Brain Sciences* 24 (5): 883–917.
- O'Regan, J. K., et A. Noë. 2001b. « Acting Out Our Sensory Experience ». *Behavioral and Brain Sciences* 24 (5): 1011–1021.
- O'Regan, J. K., et A. Noë. 2001c. « What It Is like to See: A Sensorimotor Theory of Perceptual Experience ». *Synthese* 129 (1) (octobre 1): 79-103.
- O'Shaughnessy, B. 1989. « The sense of touch. » *Australasian Journal of Philosophy*, 67(1), 37-58.
- Overgaard, M. 2001. « The Role of Phenomenological Reports in Experiments on Consciousness ». *Psychology* 12 (29): 1–10.
- Overgaard, M. 2004. « On the Naturalizing of Phenomenology ». *Phenomenology and the Cognitive Sciences* 3 (4): 365–79.
- Overgaard, Morten, et T. A. Sorenson. 2004. « Introspection Distinct From First-Order Experiences ». *Journal of Consciousness Studies* 11.
- Pessoa, L., E. Thompson, et A. Noë. 1998a. « Filling-in is for Finding Out ». *Behavioral and Brain Sciences* 21 (6): 781–796.
- Pessoa, L., E. Thompson, et A. Noë. 1998b. « Finding Out About Filling-In: A Guide to Perceptual Completion for Visual Science and the Philosophy of Perception ». *Behavioral and Brain Sciences* 21 (21): 723–802.
- Pessoa, L., E. Thompson, et A. Noë. 2001. « Filling-In: One or Many? » *Behavioral and Brain Sciences* 24 (6): 1137–1139.
- Pessoa, L. et P. De Weerd. 2003. *Filling-In : From Perceptual Completion to Cortical Reorganization: From Perceptual Completion to Cortical Reorganization*. Oxford University Press.

- Petitot, J., F. Varela, B. Pachoud, et J.-M. Roy, éd. 1999. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. 1^{re} éd. Stanford University Press.
- Prinz, J. 2006. « Putting the brakes on enactive perception ». *Psyche*, 12(1), 1-19.
- Prinz, J. 2009. « Is Consciousness Embodied? » In *Cambridge Handbook of Situated Cognition*. Éds. Aydede et Robbins, Cambridge University Press.
- Prinz, J. 2012. *The Conscious Brain: How Attention Engenders Experience*. Oxford University Press, USA.
- Putnam, H. 1967. « The Nature of Mental States », réimprimé dans *Mind, Language, and Reality* (1975), 429-440.
- Putnam, H. 1975. « The Meaning of Meaning, ». *Philosophical Papers, Vol. II : Mind, Language, and Reality*, Cambridge University Press Cambridge University Press.
- Quine, van O. W.. 1951 « Two Dogmas of Empiricism », *Philosophical Review*, 60: 20-43;
- Robinson, H. 1994. *Perception*, London: Routledge.
- Rosenthal, D. M. 1997. *A Theory of Consciousness*, in Block *et al.* 1997.
- Rosenthal, D. M. 2005. *Consciousness and Mind*. Oxford: Clarendon Press.
- Rowlands, M. 1999. *The body in mind: Understanding cognitive processes*. Cambridge University Press.
- Rowlands, M. 2003. *Externalism: Putting mind and world back together again*. McGill-Queen's Press.
- Searle, J. R. 1992. *The Rediscovery of the Mind*. MIT Press.
- Siewert, C. 1998. *The Significance of Consciousness*. Princeton University Press.
- Siewert, C. 2007. « In Favor of (Plain) Phenomenology ». *Phenomenology and the Cognitive Sciences* 6 (1-2).
- Siewert, C. 2008. « Review of Evan Thompson, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind* ». *Notre Dame Philosophical Reviews* 2008 (1).
- Simons, D. J., & Chabris, C. F. 1999. *Gorillas in our midst: Sustained inattentional blindness for dynamic events*. *Perception*, 28(9), 1059-1074.
- Stewart, J., Gapenne, O., & Di Paolo, E. A. (Eds.). 2010. *Enaction: toward a new paradigm for cognitive science*. The MIT Press.
- Stich, S. P. 1978. « Autonomous Psychology and the Belief-Desire Thesis. » Réédité dans Alvin Goldman *Readings in Philosophy and Cognitive Science*, MIT Press.

- Stich S. P. 1983. *From folk psychology to cognitive science: The case against belief*. The MIT Press.
- Stoerig, Petra. 2007. « Hunting the Ghost: Toward a Neuroscience of Consciousness ». In Zelazo *et al.* 2007.
- Sutton, J. 2010. « Exograms and interdisciplinarity: history, the extended mind, and the civilizing process. » *The extended mind*, 189-225.
- Sytsma, J., et E. Machery. 2009. « How to Study Folk Intuitions About Phenomenal Consciousness ». *Philosophical Psychology* 22 (1): 21–35.
- Sytsma, J., et E. Machery. 2010. « Two Conceptions of Subjective Experience ». *Philosophical Studies* 151 (2): 299–327.
- Tallon-Baudry, C. 2011. « On the neural mechanisms subserving consciousness and attention. » *Frontiers in psychology*, 2.
- Thagard, P. 2012. « Cognitive Science », *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2012/entries/cognitive-science/>.
- Thelen, E., Smith, L. B. 1994. *A dynamic systems approach to the development of cognition and action*. The MIT press.
- Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. 2001. « The dynamics of embodiment: A field theory of infant perseverative reaching. » *Behavioral and brain sciences*, 24(1), 1-34.
- Thompson, E.. 2007. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard University Press.
- Thompson, E., A. Lutz, et D. Cosmelli. 2005. « Neurophenomenology: An Introduction for Neurophilosophers ». In *Cognition and the Brain: The Philosophy and Neuroscience Movement*. Cambridge University Press.
- Thompson, E. 2004. Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela. *Phenomenology and the cognitive Sciences*, 3(4), 381-398.
- Thompson, E., Alva N., et Luiz P.. 1999. « Perceptual Completion: A Case Study in Phenomenology and Cognitive Science ». In *Naturalizing Phenomenology*. Stanford University Press.
- Thompson, E., Lutz, A. et Cosmelli D.. 2005. « Neurophenomenology: An Introduction for Neurophilosophers. » Dans Brook et Akins 2005.
- Thompson, E., & Stapleton, M. 2009. Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi*, 28(1), 23-30.

- Torrance, S. 2005. « In search of the enactive: Introduction to special issue on enactive experience ». *Phenomenology and the Cognitive Sciences*, 4(4), 357-368.
- Tribble, E. 2005. « Distributing cognition in the globe. » *Shakespeare Quarterly*, 56, 135-155.
- Tye, M. 1992. *Visual Qualia and Visual Experience*, in Crane 1992.
- Tye, M. 1995. *Ten problems of consciousness: A representational theory of the phenomenal mind*. The MIT Press.
- van Gelder, T. 1995. « What Might Cognition Be If Not Computation? » *Journal of Philosophy* 92(7), 345-81.
- van Gelder, T. 1998. « The Dynamical Hypothesis in Cognitive Science ». *Behavioral and Brain Sciences* 21(5), 615-28.
- van Gelder, Tim , et Robert Port. 1995. *Mind As Motion: Explorations in the Dynamics of Cognition*. MIT Press.
- van Gulick, Robert. 2011. « Consciousness », *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2011/entries/consciousness/>.
- Varela, F. J., E. Rosch, et E. Thompson. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. 1^{re} éd. The MIT Press.
- Varela, F. J. 1996. « Neurophenomenology: A Methodological Remedy for the Hard Problem. » In *Explaining Consciousness: The Hard Problem*, édité par Jonathan Shear, MIT Press, p. 337-357. Cambridge: A Bradford Book.
- Velmans, M. 2007a. « An Epistemology for the Study of Consciousness ». *The Blackwell Companion to Consciousness* éd. Velman et Schneider. Blackwell.
- Velmans, M. 2007b. « Heterophenomenology Vs. Critical Phenomenology ». *Phenomenology and the Cognitive Sciences* 6 (1-2).
- Weiskrantz, L. 1986. *Blindsight : A case study and implications*. Oxford University Press.
- Wheeler, M. 2005. *Reconstructing the cognitive world: The next step*. MIT Press.
- Wheeler, M. *In defense of extended functionalism*. In Menary 2010.
- Wilson, M. 2002. « Six views of embodied cognition ». *Psychonomic bulletin & review*, 9(4), 625-636.
- Wilson, R. A., & Keil, F. C. (Eds.). 1999. *The MIT encyclopedia of the cognitive sciences*. MIT press.
- Zahavi, D. 2005. *Subjectivity and selfhood: Investigating the first-person perspective*. The MIT press.

- Zahavi, D. 2004a. « Back to Brentano? » *Journal of Consciousness Studies* 11 (10-11): 66–87.
- Zahavi, D. 2004b. « Phenomenology and the Project of Naturalization ». *Phenomenology and the Cognitive Sciences* 3 (4): 331–47.
- Zahavi, D. 2013. « Naturalized Phenomenology: A Desideratum or a Category Mistake? » *Royal Institute of Philosophy Supplement* 72: 23–42.
- Zelazo, P. D., Moscovitch, M., & Thompson, E. (Eds.). (2007). *The Cambridge handbook of consciousness*. Cambridge University Press.