

L'AUTONOMIE DES SCIENCES SPÉCIALES RECONSIDÉRÉE

Christian Sachse

Patrice Soom*

Quelle est la relation entre les sciences spéciales et la physique ? Les sciences spéciales s'attachent à décrire des entités qui sont, semble-t-il, ultimement composées d'entités purement physiques. Ceci tend à soutenir une position réductionniste à la fois sur les plans ontologique et épistémologique. Cependant, l'argument de la réalisation multiple est souvent interprété comme bloquant définitivement la possibilité de la réduction des sciences spéciales à la physique et leur confère, par là, une certaine autonomie. Cet argument conduit néanmoins au dilemme suivant : soit la réalisation multiple n'est pas compatible avec le réductionnisme ontologique et la complétude du domaine de la physique, soit elle motive une réduction de type éliminativiste.

Dans cet article, nous nous proposons de développer, en traitant des exemples concrets empruntés à la biologie et à la psychologie, une nouvelle position réductionniste, tant ontologique qu'épistémologique, qui, sans motiver l'élimination des espèces naturelles des sciences spéciales, n'entre pas en conflit avec le réductionnisme ontologique et la complétude de la physique. Bien que fondamentalement réductionniste, le caractère conservateur de cette position accorde toutefois une certaine autonomie aux sciences spéciales. Nous montrerons ainsi en quoi la biologie et la psychologie, en tant que sciences spéciales, constituent des sciences à la fois autonomes et réductibles aux sciences physiques.

I. Introduction

Le débat autour de l'autonomie des sciences spéciales est dominé par des arguments généraux. Dans cet article, nous traiterons des principaux arguments régissant le débat relatif à la question de l'autonomie des sciences spéciales, à savoir l'argument causal en faveur de la réduction des sciences spéciales et l'argument de la réalisation multiple, qui est le plus souvent interprété comme révélant l'irréductibilité des propriétés des sciences spéciales. Nous nous proposons ici de montrer, en considérant des exemples concrets empruntés à la biologie et à la psychologie, que l'argument de la réalisation multiple saisit une spécificité réelle des sciences

* La contribution de Patrice Soom à cet article a été soutenue pour le Fonds National Suisse (FNS), bourse n° PDFM1--118603/1.

spéciales qui leur confère une certaine forme d'autonomie, mais que la conclusion anti-réductionniste, qui est le plus souvent tirée sur cette base, n'est, elle, pas justifiée. Bien que débutant dans le cadre du Cercle de Vienne, le débat général concernant le réductionnisme et l'unité des sciences a pris sa forme actuelle avec Paul Oppenheim et Hilary Putnam à la fin des années 1950¹. Selon eux, la position la plus cohérente en philosophie des sciences consiste à postuler l'unité des sciences comme hypothèse de travail, ce qui suppose que les différentes sciences entretiennent d'étroites relations entre elles. Une telle approche trouve aujourd'hui sa motivation et ses fondements dans l'argument en faveur du réductionnisme ontologique (l'identité des occurrences de propriétés) et dans la complétude de la physique. La section II sera consacrée à un examen relativement détaillé de ces fondements argumentatifs dans le contexte de la biologie. Nous y examinerons en particulier le concept de survenance, le concept de complétude de la physique et, par conséquent, l'argument en faveur du réductionnisme ontologique.

La section III sera consacrée à la mise en évidence, toujours dans le contexte de la biologie, de l'obstacle que constitue l'argument de la réalisation multiple pour toute approche réductionniste – et, par là même, pour l'hypothèse de l'unité des sciences. La référence canonique concernant l'argument de la réalisation multiple date de 1974 et est due à Jerry Fodor, qui développe par là un argument fort en faveur de la pluralité des sciences et de l'autonomie des sciences spéciales. Cet argument, qui constitue une réponse directe à la position d'Oppenheim and Putnam², influence encore profondément le débat actuel. Selon Fodor, les sciences spéciales sont autonomes par rapport à la physique parce qu'elles ne sont pas réductibles à cette dernière. Cependant baser l'autonomie de la biologie sur un argument anti-réductionniste conduit au dilemme suivant : soit la réalisation multiple n'est pas compatible avec le réductionnisme ontologique et la complétude de la physique et conduit à l'épiphénoménalité des propriétés des sciences spéciales, soit elle motive une réduction de type éliminativiste. Nous allons expliquer en détail ce dilemme auquel sont confrontées les positions anti-réductionnistes dans le contexte des relations inter-théoriques à la fin de la section III.

Pour sortir de ce dilemme, nous proposons en section IV une nouvelle stratégie réductionniste capable d'incorporer la réalisation multiple. Nous exposerons un argument causal qui plaide en faveur de la construction de sous-concepts fonctionnels qui caractérisent les propriétés dont traitent les sciences spéciales et qui sont nomologiquement co-extensionnels avec des concepts physiques. A partir de là, il est possible d'argumenter en faveur de la qualité scientifique théorique de ces sous-concepts puisqu'ils ne sont pas en conflit avec le réductionnisme ontologique et la

¹ Oppenheim, Paul et Putnam, Hilary, « Unity of Sciences as a Working Hypothesis », in *Minnesota studies in the Philosophy of Sciences*, no 2, Minneapolis: University of Minnesota Press, 1958, pp. 3-36.

² Jerry A. Fodor, « Special Sciences (or: the Disunity of Science as a Working Hypothesis) », in *Synthese*, no. 28, 1974, pp. 97-115.

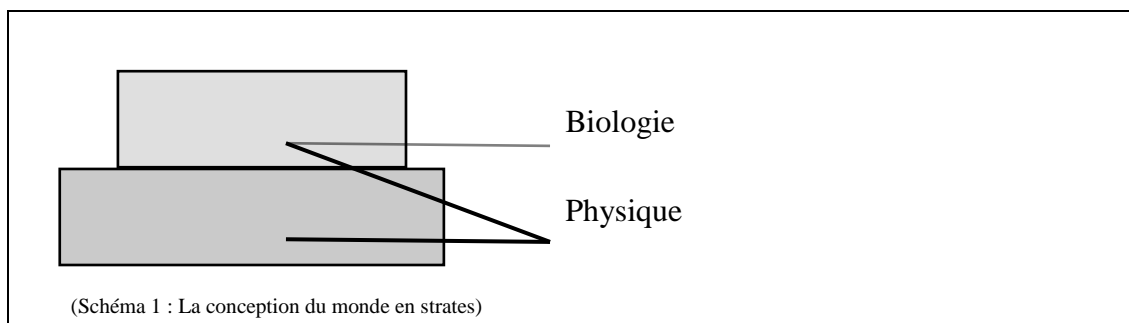
complétude de la physique. Nous allons ensuite soutenir que la différence entre un concept et ses sous-concepts est le degré d'abstraction, la qualité scientifique des concepts abstraits n'étant pas mise en question puisqu'ils sont réductibles à la physique par le biais de leurs sous-concepts. Une telle réduction n'est pas éliminativiste parce que les concepts abstraits des sciences spéciales restent indispensables dans notre système scientifique, dans la mesure où ils permettent de mettre en évidence des similarités pertinentes dans le monde – soit des espèces naturelles propres aux sciences spéciales – que la physique n'est pas capable de montrer. Le choix du titre de cet article sera alors justifié par notre conclusion finale selon laquelle les sciences spéciales possèdent une certaine autonomie dont la nature prend sens dans le cadre de la nouvelle approche réductionniste que nous proposons.

Dès la section V, nous nous tournerons ensuite vers le cas plus controversé de l'esprit en considérant la psychologie en tant que science spéciale. La section VI introduira un exemple classique de recherche en neuropsychologie, le syndrome de déconnexion, qui nous servira dans la section suivante comme point d'appui pour l'application de la stratégie discutée dans cette publication. Enfin, la section VII nous permettra d'aborder un cas plus complexe et de montrer comment les neurosciences sont elles-mêmes adossées à une série de présupposés qui ne constituent rien d'autre que les conditions nécessaires à la mise en œuvre de la stratégie de réduction fonctionnelle par sous-concepts.

II. L'argument en faveur du réductionnisme ontologique

Dans cette section, nous allons présenter les développements récents du débat concernant le rapport entre les occurrences de propriétés (en anglais *property tokens*) des sciences spéciales et les occurrences de propriétés physiques. Pour considérer en détail ce rapport, prenons ici à titre d'exemple la génétique comme un domaine de la biologie postulant un ensemble de propriétés biologiques. D'un côté, la génétique postule l'existence de gènes qui sont définis fonctionnellement par leurs effets phénotypiques, à savoir, par exemple, la production de protéines d'un certain type. D'un autre côté, du point de vue ici simplifié de la physique, une occurrence d'un gène est une configuration d'atomes, une séquence d'ADN dans un contexte moléculaire. Les gènes en question possèdent donc également des propriétés physiques. Se pose ainsi la question générale consistant à déterminer quel est le rapport entre les propriétés biologiques et les propriétés physiques qu'un objet instancie. En d'autres termes, y a-t-il une différence *ontologique* entre un gène (décrit et expliqué par la génétique) et la configuration d'atomes (décrite et expliquée par la physique) lui correspondant ?

Comme première approche, référons-nous à la conception du monde en strates³. Les objets instanciant des propriétés physiques fondamentales, comme le spin, la masse ou la charge, n’instancient pas nécessairement des propriétés biologiques. C’est la raison pour laquelle la physique est plus fondamentale que la biologie et c’est pourquoi il existe des objets physiques qui ne sont pas biologiques. Ce fait ontologique suggère l’existence d’une strate physique non-biologique. En revanche, tous les objets instanciant des propriétés biologiques instancient également des propriétés physiques. Chaque gène, par exemple, est également une configuration complexe de propriétés physiques. C’est la raison pour laquelle la physique est également plus universelle que la biologie. La physique s’applique à tous les objets instanciant des propriétés biologiques en se référant à leur structure physique. Le schéma ci-dessous illustre cette asymétrie : la biologie s’applique seulement à un ensemble particulier de propriétés existant dans le monde. La physique, par contre, s’applique aussi bien aux objets instanciant seulement des propriétés physiques qu’à ceux instanciant tant des propriétés biologiques que physiques :



Dans le contexte de la conception du monde en strates, les questions qui se posent sont les suivantes. Si un objet constituant une occurrence d’une propriété biologique quelconque est aussi systématiquement une configuration de propriétés physiques, quel est le rapport entre les propriétés biologiques et les propriétés physiques instanciées par l’objet considéré ? Par exemple, quel est le rapport entre un certain gène et la configuration d’atomes qui le compose ? Est-ce qu’il y a une *différence ontologique* entre les deux qui justifie l’existence de strates ontologiques différentes dans le monde – une strate physique et une strate biologique non-physique ? Ou, au contraire, peut-on *identifier* chaque occurrence d’une propriété biologique quelconque avec sa composition physique ? Ce qui impliquerait alors une interprétation non métaphysique de la conception du monde en strates : celle-ci reflèterait dans ce cas seulement les différentes approches épistémologiques possibles d’une seule strate ontologique qui contient des configurations plus ou moins complexes des propriétés.

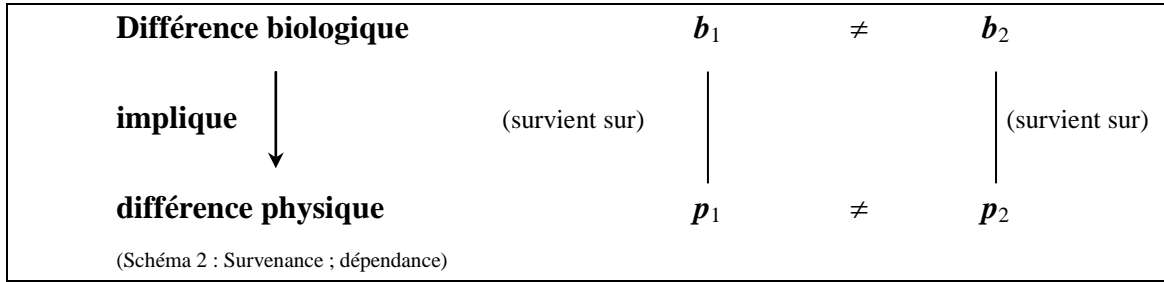
³ Voir Jaegwon Kim, « The Layered Model : Metaphysical Considerations », in *Philosophical Explorations*, no. V, 2002, pp. 2-20.

Pour donner une réponse bien argumentée à ces questions, présentons brièvement l'idée principale du concept de survenance qui précise la relation entre les occurrences des propriétés des différentes strates. Selon la thèse de la survenance, utilisée pour décrire le rapport en question, chaque monde possible qui est un double physique *minimal* du monde réel est un double *simpliciter* du monde réel⁴. En d'autres termes, si on duplique notre monde physique, on duplique par là également les occurrences de propriétés biologiques présentes dans notre monde. Dans ce double, il y a alors aussi des gènes, des plantes, des animaux, etc. Cette détermination des propriétés biologiques par des propriétés physiques, qui n'est pas une détermination causale, est soutenue pour deux raisons au moins. Premièrement, le développement de notre univers suggère que l'apparition d'une occurrence d'une propriété biologique quelconque implique qu'il y ait une base physique d'une certaine complexité. Cela montre que, dans notre monde, l'existence des propriétés physiques est nécessaire pour l'existence des propriétés biologiques. Deuxièmement, étant donné des circonstances environnementales physiques constantes, aucune différence biologique n'a jamais été observée entre deux objets physiquement identiques.

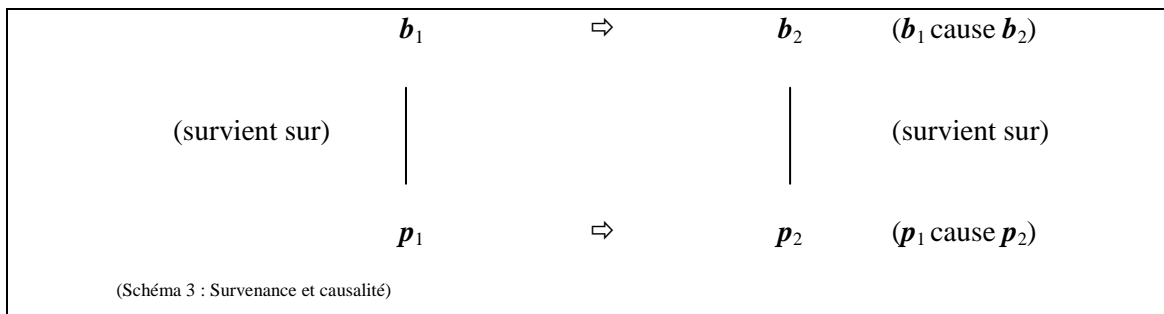
Continuons notre considération dans un contexte plus local que celui que fixe la survenance globale, puisque le critère principal permettant d'isoler une différence ontologique entre différentes occurrences d'une même propriété biologique est une différence causale et que la causalité est une relation locale⁵. Soit une occurrence de propriété biologique, b_1 , qui survient sur une configuration d'occurrences de propriétés physiques spatio-temporellement délimitée, p_1 . La relation de survenance liant les deux propriétés implique que, pour tout changement au niveau de la propriété biologique b_1 , il y a *nécessairement* un changement au niveau de la configuration physique p_1 . Par exemple, il n'est pas possible qu'une plante développe des fleurs sans qu'il y ait également un changement dans la configuration atomique qui la compose. En d'autres termes, s'il y a une différence biologique entre des occurrences de propriétés biologiques, il y a également des différences physiques entre les deux configurations physiques. Une telle relation de dépendance entre occurrences de propriétés biologiques et physiques est empiriquement suggérée par le fait qu'il n'existe aucune expérience connue dans laquelle des différences biologiques apparaissent sans différences physiques correspondantes. Le schéma suivant illustre la dépendance impliquée par le concept de survenance.

⁴ Voir Frank Jackson, *From Metaphysics to Ethics*, Oxford, Oxford University Press, 1998, pp. 9-14.

⁵ Voir Jaegwon Kim, « Making sense of emergence », in *Philosophical Studies*, no. 9, 1999, pp. 3-36, et Sydney Shoemaker, « Causality and Properties », in Van Inwagen, Peter (dir.), *Time and Cause*, Dordrecht, Reidel, 1980, pp. 109-135.



Par analogie avec la survenance globale, on peut affirmer qu'un *double* de la configuration d'occurrences de propriétés physiques p_1 est aussi un double de l'occurrence de la propriété biologique b_1 . Donnons un exemple : soit un gène (b_1) qui survient sur sa configuration d'atomes (p_1). Si on duplique cette configuration d'atomes p_1 - on obtient donc un p_2 qui est physiquement indiscernable de p_1 - alors on duplique aussi le gène b_1 , donc on produit un b_2 qui est biologiquement indiscernable de b_1 . Cette détermination non-causale et non-explicative des propriétés biologiques par les propriétés physiques est fondée par les arguments déjà mentionnés dans le contexte de la survenance globale et elle est illustrée par le schéma suivant :



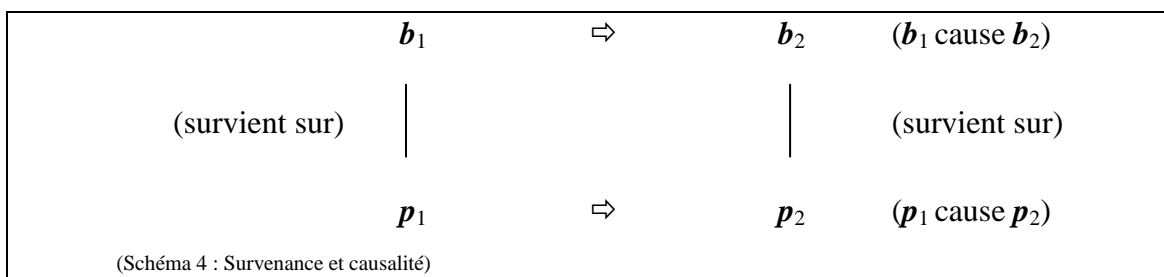
Etant donné que la survenance locale des propriétés biologiques est généralement admise, focalisons-nous sur un autre concept qui joue un rôle important dans le rapport entre les propriétés biologiques et les propriétés physiques. Il s'agit de la complétude causale, explicative et nomologique de la physique. La complétude de la physique est une hypothèse de travail qui est également largement admise, du fait de l'existence de théories physiques fondamentales et universelles. Dans le contexte qui nous intéresse, cette thèse consiste à affirmer que « *pour toutes les occurrences de propriétés physiques, dans la mesure où une occurrence d'une propriété physique possède une explication, elle possède une explication en termes physiques*⁶ ». Ainsi, pour l'explication d'un changement physique, une explication physique est toujours suffisante. De ce fait, la physique est également complète en ce qui concerne la causalité et les lois : pour tout changement physique, il y a toujours une cause physique suffisante et des lois qui décrivent cette relation causale, dans la mesure où il y a des causes et des

⁶ Michael Esfeld, *Philosophie des sciences*, Lausanne, Presses polytechniques et universitaires romandes, 2006, p. 214.

lois pour les décrire. Par exemple, pour un changement quelconque dans une configuration d'atomes, il y a une cause physique suffisante, dans la mesure bien sûr où il y a des causes. Par conséquent, pour fournir une explication suffisante de ce changement et de la relation causale associée, la physique n'a jamais besoin d'avoir recours à la biologie et à ses lois ou concepts explicatifs.

Une fois la survenance et la complétude de la physique admises, nous pouvons argumenter en faveur d'un réductionnisme ontologique de la façon suivante⁷. Considérons en premier lieu l'occurrence d'une propriété biologique b_1 survenant sur une configuration d'occurrences de propriétés physiques p_1 (à savoir, p_1 est suffisant pour b_1). L'efficacité causale de b_1 est généralement admise, c'est-à-dire que l'on considère que celle-ci n'est pas un épiphénomène et que le concept biologique correspondant à b_1 possède une portée explicative. Par conséquent, assumons que b_1 soit la cause d'une autre occurrence d'une propriété biologique, b_2 . Etant donné la prémisse de survenance, la différence biologique entre b_1 et b_2 implique qu'il y a en même temps une différence physique entre les deux. En d'autres termes, b_2 survient sur une configuration d'occurrences de propriétés physiques p_2 qui est distincte de p_1 . Par exemple, un gène (b_1) cause un certain effet phénotypique (b_2), comme, par exemple, la production de protéines d'un certain type. De plus, la survenance implique que le gène et son effet phénotypique sont physiquement différents ($p_1 \neq p_2$).

En présupposant qu'il y ait une explication biologique de l'occurrence de b_2 qui fasse référence à b_1 et à son efficacité causale, il y a aussi, du fait de la complétude de la physique, une explication physique suffisante de p_2 . La configuration d'occurrences de propriétés physiques p_2 possède une cause physique complète. Disons que p_1 est cette cause physique complète, ou, au moins, que p_1 fait partie de la cause complète de p_2 . Le schéma suivant donne un aperçu de la situation :



Du fait de la relation de survenance, la propriété biologique b_2 est déterminée par la configuration de propriétés physiques p_2 . Toutefois, en produisant p_2 , p_1 cause aussi ce qui survient sur p_2 . Nous avons donc deux causes pour b_2 : la cause biologique b_1 ($b_1 \Rightarrow b_2$) et p_1 en tant que cause de la base de survenance p_2 de b_2 ($p_1 \Rightarrow p_2$ et b_2 survient sur p_2). Dans les termes de notre exemple, la configuration d'atomes

⁷ Voir Jaegwon Kim, *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, pp. 32-69.

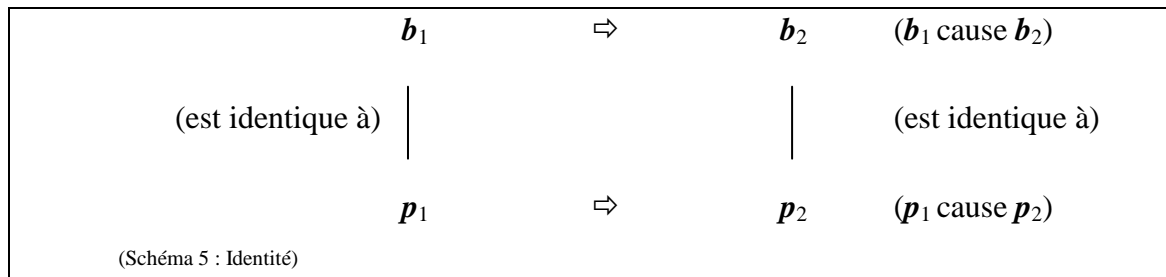
composant le gène cause les changements physiques dans la cellule qui déterminent eux-mêmes l'effet phénotypique en question (la production de protéines d'un certain type).

Dans des termes plus généraux, les relations causales entre deux occurrences de propriétés biologiques dépendent des relations causales entre les deux configurations d'occurrences de propriétés physiques qui constituent leurs bases de survenance respectives. La survenance et la complétude de la physique impliquent conjointement qu'une relation causale entre deux occurrences de propriétés biologiques ($b_1 \Rightarrow b_2$) n'est pas possible sans qu'il y ait une relation causale entre leurs bases de survenance physiques respectives ($p_1 \Rightarrow p_2$). Il n'est dès lors pas nécessaire, d'un point de vue causal, de supposer une relation causale biologique *en plus* d'une relation causale physique. La configuration d'occurrences de propriétés physiques p_1 est suffisante, en causant p_2 , pour causer b_2 . En d'autres termes, les propriétés physiques du gène en question sont suffisantes pour causer l'effet phénotypique (en causant la base de survenance de cet effet). Par conséquent, les causes physiques semblent ainsi faire déjà tout le travail que les occurrences des propriétés biologiques sont supposées accomplir.

C'est la raison pour laquelle il existe un argument causal pour *identifier* les occurrences des propriétés biologiques à leurs bases physiques sur le plan ontologique et, ce faisant, pour y identifier les causes biologiques à des causes physiques. L'impossibilité de distinguer causalement les occurrences b_1 de p_1 constitue un argument fort pour les identifier⁸. La dépendance de la biologie par rapport à la physique qui découle de la survenance et de la complétude de la physique excluent qu'il y ait une causalité biologique qui soit différente de la causalité physique ou, plus précisément, qu'il y ait quelque chose de causalement efficace en plus de la causalité physique. Pour distinguer l'occurrence d'une propriété biologique quelconque (b_1) de la configuration d'occurrences de propriétés physiques sur laquelle elle survient (p_1), il semble qu'une différence causale soit nécessaire entre les deux. S'il existait une efficacité causale des occurrences des propriétés biologiques en dehors des relations causales décrites par la physique, alors il existerait une différence ontologique entre les deux. Cependant une telle différence n'est pas compatible avec la survenance des propriétés biologiques sur des occurrences de propriétés physiques et avec la complétude de la physique. Le schéma suivant synthétise l'état actuel de ce débat, qui constitue une sorte d'accord minimal en philosophie des sciences⁹ :

⁸ Voir Michael Esfeld, *La philosophie de l'esprit. De la relation entre l'esprit et la nature*, Paris, Armand Collin, 2005, ch. 1-4 qui considère en détail l'argument et les alternatives possibles qui ne sont pas considérées ici parce qu'elles aboutissent à l'épiphénoménisme qui n'est pas une position philosophique très engageante.

⁹ Voir Jaegwon Kim, *Physicalism, or Something Near Enough*, Princeton, Princeton University Press, 2005, pp. 70-92.



En résumé, sur la base des prémisses de la survenance des occurrences de propriétés biologiques sur des configurations d'occurrences de propriétés physiques (prémisse 1) et de la complétude de la physique (prémisse 2), il est possible de formuler un argument causal en faveur du réductionnisme ontologique. Chaque occurrence des propriétés biologiques est identique à une configuration d'occurrences de propriétés physiques (token identity). Par exemple, une occurrence du gène est identique à une configuration de propriétés physiques dans un certain contexte physique. Si l'on prend ce réductionnisme ontologique comme point de départ, la conception du monde en strates ne concerne alors que le rapport *épistémologique* entre la biologie et la physique. La question à laquelle il nous faut répondre dans la suite de ce texte est donc la suivante : quel est le rapport entre la physique et la biologie *d'un point de vue épistémologique*, étant donné l'identité entre les occurrences des propriétés biologiques et ses composants physiques – c'est-à-dire, étant donné que ces deux théories se réfèrent aux mêmes occurrences de propriétés ?

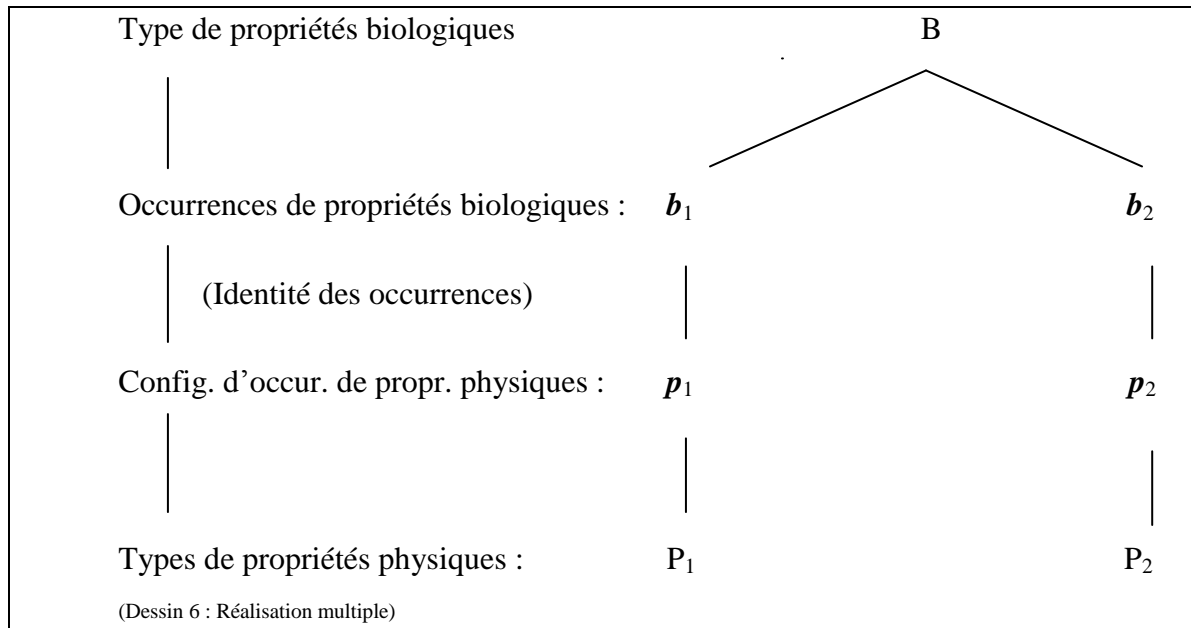
III. L'argument et le dilemme de la position anti-réductionniste

Chaque théorie décrit des occurrences de propriétés en se référant à un ensemble de *types* de propriétés qui lui sont spécifiques. Par exemple, la génétique décrit l'occurrence d'un gène, en tant que gène d'un certain type qui est lui-même fonctionnellement défini. La physique, elle, décrit l'occurrence d'une configuration d'atomes, en tant que configuration d'un certain type, lequel est en général défini par sa structure (voir aussi la contribution de Michael Esfeld dans ce numéro). Pour examiner le rapport entre des théories différentes, comme le rapport entre la biologie et la physique et les concepts qu'elles utilisent respectivement pour décrire et expliquer des occurrences de propriétés dans le monde, considérons le concept de *type* de propriétés dans le contexte de la réalisation multiple qui constitue l'argument standard contre une position réductionniste sur le plan épistémologique¹⁰.

¹⁰ Voir comme référence canonique Fodor, Jerry A., « Special Sciences (or: the Disunity of Science as a Working Hypothesis) », in *Synthese*, no. 28, 1974, pp. 97-115. Voir également Hull, David, « Reduction in Genetics-biology or Philosophy ? », in *Philosophy of Science*, no. 29, vol. 4, 1972, pp. 491-499, en particulier p. 497.

En biologie un type de propriétés est en général fonctionnellement défini. L'argument en faveur d'une telle approche fonctionnaliste est, entre autres, la présupposition de l'efficacité causale de chaque occurrence d'une propriété quelconque. Une telle efficacité causale est nécessaire pour que la biologie soit explicative, puisque la référence à des épiphénomènes n'explique rien. En conséquence, comme dans notre exemple simplifié, il y a des occurrences de gènes qui causent des effets phénotypiques, comme par exemple la production de protéines d'un certain type. Cette relation causale caractérise fonctionnellement le type de propriétés en question : chaque occurrence d'un gène causalement responsable de la production d'une protéine déterminée est un gène du même type. C'est, d'un point de vue génétique, une similarité pertinente entre différentes occurrences de propriétés.

Considérons maintenant deux occurrences de gènes appartenant au même type de propriétés, b_1 et b_2 du type B . Conséquemment à l'argument développé dans la section précédente, chaque occurrence de propriétés biologiques est considérée comme identique à une configuration d'occurrences de propriétés physiques, disons $b_1 = p_1$ et $b_2 = p_2$. Cependant, il est possible que les deux occurrences de propriétés biologiques (b_1 et b_2) appartenant au même type (B) aient pour base de survenance deux configurations de propriétés physiques (p_1 et p_2) appartenant à des types physiques différents. Par exemple, en raison de la redondance du code génétique, il est possible que les deux occurrences de gènes qui produisent des protéines du même type soient physiquement différentes. On peut résumer de la façon suivante l'idée de la réalisation multiple : il existe des occurrences de propriétés biologiques qui appartiennent à un même type de propriétés parce qu'elles satisfont à la définition fonctionnelle du type de propriétés en question. Cependant, d'un point de vue physique, ces occurrences de propriétés appartiennent à des types différents parce qu'il existe une différence dans leur composition physique respective. Le schéma suivant illustre le lien entre ce qui a été vu dans la section précédente et l'argument de la réalisation multiple :



En acceptant la possibilité de la réalisation multiple, il n'est pas possible de corréler systématiquement (bi-conditionnellement) tous les types de propriétés biologiques à des types de propriétés physiques. Il est donc généralement admis que le modèle réductionniste nagelien ne s'applique pas à la biologie parce que les corrélations bi-conditionnelles entre les types des propriétés biologiques et les types de propriétés physiques manquent systématiquement¹¹. C'est en général l'argument avancé depuis les années 1970 pour affirmer que l'unité des sciences n'existe pas et que la biologie est, et comme l'ensemble des sciences spéciales, autonome par rapport à la physique¹². Cependant, le débat autour du réductionnisme s'est poursuivi depuis les années 70 dans deux directions différentes¹³. D'un côté, il y a les approches anti-réductionnistes telles que celle de Fodor – qui accepte le réductionnisme ontologique mais réfute le réductionnisme épistémologique. D'autre côté, il y a les approches réductionnistes qui essaie de réduire aussi épistémologiquement la biologie à la physique en dépit de la réalisation multiple¹⁴. Cependant, les deux approches suggèrent directement ou aboutissent indirectement à l'élimination des sciences spéciales. Pour mieux

¹¹ Nagel, Ernest, *The Structure of Science. Problems in the Logic of Scientific Explanation*. London, Routledge, 1961, ch. 11.

¹² Voir Fodor, Jerry A., « Special Sciences (or: the disunity... », in *Synthese*, no. 28, 1974, pp. 97-115.

¹³ Voir également Brigandt, Ingo et Love, Alan, « Reductionism in Biology », in *The Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/reduction-biology/>, 2008.

¹⁴ Voir par exemple Lewis, David, « Mad Pain and Martian Pain », in Block, Ned, dir., *Readings in the Philosophy of Psychology*, vol 1, London, Methuen. 1980, Pp. 216-222. Réimprimé in Lewis, David, *Philosophical Papers*, Volume 1, Oxford, Oxford University Press. 1980, pp. 122-130. Kim, Jaegwon, *Mind in a Physical World. An Essay on the Mind-body Problem and Mental Causation*, Cambridge (Massachusetts), MIT Press, 1998, voir en particulier pp. 93-95. Kim, Jaegwon, *Physicalism, or...*, voir en particulier p. 25. Heil, John, *From an Ontological Point of View*, Oxford, Clarendon Press, 2003, pp. 61-65.

comprendre ce danger d'élimination dans le cas paradigmatique de la biologie, précisons tout d'abord la notion de type de propriétés et le concept de réalisation.

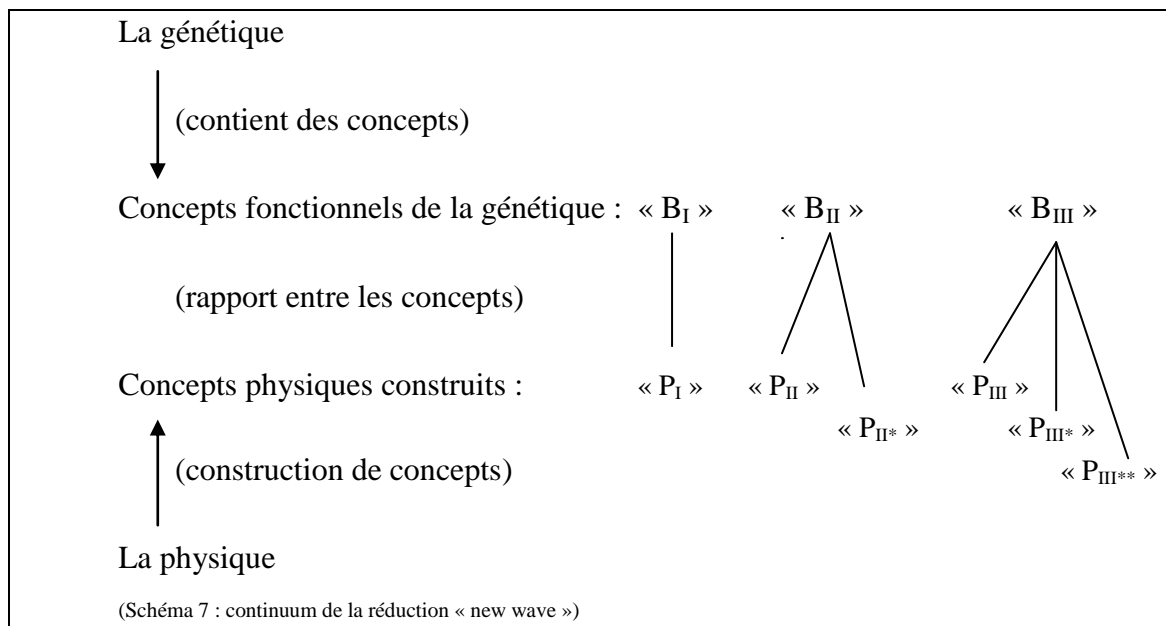
En raison de la réalisation multiple, nous l'avons vu, il ne peut pas y avoir d'identité des *types* de propriétés biologiques et physiques. Par contre, en raison du réductionnisme ontologique, il existe une identité des *occurrences* des propriétés de la biologie et de la physique. En d'autres termes, il est possible qu'il y ait une différence physique entre deux occurrences qui sont toutefois biologiquement similaires car elles satisfont la même définition fonctionnelle. Dans le contexte de la réalisation multiple et du réductionnisme ontologique, le concept de *réalisation* peut être interprété comme suit pour peu que l'on souscrive à l'approche fonctionnaliste en philosophie des sciences. Dans le monde, il existe seulement des occurrences de propriétés et chaque occurrence d'une propriété biologique est identique à quelque chose de physique (réductionnisme ontologique). Dès lors, si, par exemple, les types de propriétés génétiques ne sont pas identiques aux types de propriétés physiques, les types ne sont pas des universaux qui existent dans le monde. *Réaliser* signifie qu'une occurrence de propriété *rend vraie* l'application d'un ou de différents concepts. Un *type* de propriétés est un *concept* intégré dans une théorie. En d'autres termes, si un énoncé sur le monde est vrai, il y a quelque chose dans le monde qui le rend vrai; il y a des occurrences de propriétés dans le monde qui rendent vrai l'application des concepts de différentes théories. Dans le contexte de notre exemple génétique, il existe des occurrences de gènes dans le monde qui rendent vrais l'application du concept « gène pour la production de la protéine *PR* », certaines de ces occurrences de gènes rendent vrai le concept physique « P_1 », et d'autres rendent vrai le concept physique « P_2 ». L'argument de la réalisation multiple ou de la *référence* multiple (parce que selon notre interprétation les types des propriétés sont des concepts) est donc un argument portant sur les théories et les concepts qu'elles utilisent. Résumons l'interprétation de la conception du type de propriété et l'argument de la réalisation multiple : la réalisation multiple de type de propriétés est une référence multiple signifiant qu'il n'y a systématiquement pas de *co-extensionnalité* entre un concept biologique (comme « gène pour ... ») et *un seul* concept physique.

Cependant, étant donné la réalisation multiple, on rencontre un dilemme : soit elle n'empêche pas la réduction dans son acception éliminativiste, soit elle n'est pas cohérente avec le réductionnisme ontologique et la complétude de la physique. Considérons les deux cotés de ce dilemme l'un après l'autre.

Premièrement, considérons l'argument qui est mis en évidence par le « new wave reductionism » de John Bickle en faveur d'une approche réductionniste de type éliminativiste¹⁵. Cette position accepte la conséquence de la réalisation multiple, à savoir qu'on ne peut pas établir une relation systématique (biconditionnelle) entre un

¹⁵ Voir Bickle, John, *Psychoneural Reduction. The New Wave*, Cambridge, MIT Press, 1998, surtout les chapitres 2-4.

concept biologique comme le concept « gène pour ... » et un concept physique. Cependant, il est possible de construire des théories physiques qui s'appliquent au domaine biologique d'une manière particulière : les théories biologiques font référence à certains ensembles d'occurrences de propriétés. Par exemple, la génétique s'applique aux gènes, chromosomes, etc. Dans ce cas-là, chaque occurrence d'un gène, d'un chromosome, etc. qui est décrite et expliquée par la génétique est identique à quelque chose de physique, à certaines configurations de propriétés physiques (réductionnisme ontologique). Sur la base de la complétude de la physique, on construit une théorie physique, une théorie *image*, qui fait référence exclusivement à des configurations de propriétés physiques qui sont des occurrences des propriétés génétiques¹⁶. Ensuite, on regarde les rapports entre les concepts génétiques et les concepts physiques construits (les descriptions physiques des entités génétiques). Etant donné le réductionnisme ontologique et la complétude de la physique, les rapports entre les concepts se trouvent dans un continuum entre bi-conditionnalité (à gauche sur le schéma) et réalisation multiple de plus en plus extrême (à droite sur le schéma) :



L'approche « new wave » contient un message ontologique et un message épistémologique. Commençons avec l'ontologie. Dans le cas où le rapport entre, par exemple, le concept « gène pour ... » et le concept physique artificiellement construit qui lui est associé est bi-conditionnel, le concept « gène pour ... » montre une similarité causale dans le monde qui est une similarité physique parce qu'il existe un concept

¹⁶ Voir en ce qui concerne la construction des concepts et donc des théories, Hooker Clifford, « Towards a General Theory of Reduction. Part I : Historical and Scientific Setting. Part II : Identity in Reduction. Part III : Cross-categorical Reduction », in *Dialogue*, no. 20, 1981, surtout les pages 49-52.

physique complexe qui montre la même similarité causale. Dans ce sens il est donc possible d'être réaliste en ce qui concerne le gène en question. Par contre, dans le cas où le concept « gène pour ... » fait référence à des occurrences de propriétés qui sont décrites de manières très différentes par des concepts physiques complexes, il est suggéré que la similarité montrée par « gène pour ... » n'existe pas dans le monde. Il n'est alors pas possible d'être réaliste à l'égard des gènes réalisés multiples.

Au niveau épistémologique, cette approche « new wave » suggère le remplacement de la génétique ou une théorie biologique quelle qu'elle soit par la théorie physique construite qui lui correspond. Chaque théorie physique construite est causalement, nomologiquement et explicativement complète et, donc, préférable à la théorie biologique qui lui correspond. Cette approche éliminativiste est indépendante du rapport entre les concepts biologiques et les concepts physiques construits – la réalisation multiple n'empêche pas cette approche. En termes généraux, chaque théorie biologique peut être remplacée par une théorie physique construite artificiellement (ou un ensemble de théories construites) qui est co-extensionnelle avec la théorie biologique en question. Cette conclusion éliminativiste constitue la première partie du dilemme auquel notre approche cherche à proposer une solution.

Si on admet la possibilité d'adopter la stratégie de Bickle, tout en voulant néanmoins empêcher l'élimination de la biologie, on doit argumenter en faveur de la valeur scientifique de la biologie. Cela signifie que les concepts abstraits de la biologie confèrent une valeur scientifique à la théorie qui les produit car la théorie permet par leur usage de mettre en évidence des similarités dans le monde que la physique ne pourrait pas isoler. Allant dans ce sens, Fodor prend l'argument de la réalisation multiple pour un argument *anti-réductionniste*. Il soutient que la biologie est scientifiquement indispensable et, donc, autonome, parce qu'il n'existe pas de concepts physiques co-extensionnels avec les concepts biologiques. Pourtant, la valeur scientifique d'un concept (et donc de la théorie à laquelle il appartient) est mise en question car, en l'absence de lien systématique avec les concepts physiques, il semble que la similarité retenue grâce au concept biologique en question n'existe pas dans le monde. Pour justifier la qualité scientifique des concepts abstraits biologiques sans établir un tel lien systématique avec les concepts physiques, il faut admettre que les concepts biologiques (comme « gène pour ... », par exemple) captent des aspects ontologiques qui ne sont pas saisissables par des concepts physiques ; c'est pour cette raison qu'on ne peut pas éliminer les concepts biologiques. Une telle position implique cependant de rejeter le réductionnisme ontologique et le principe de complétude de la physique : il y aurait dans ce cas des aspects ontologiques du monde qui ne seraient pas physiques et qui n'admettraient donc pas d'explication en termes physiques. C'est ici la conséquence de la seconde partie du dilemme issu de l'argument de la réalisation multiple. Or, c'est précisément cette même conséquence que la nouvelle approche réductionniste que nous proposons s'efforce d'éviter. Si l'on veut esquiver la conclusion

du « new wave reductionism » de Bickle, il faut interpréter l'argument de la réalisation multiple sans devoir en supporter la conclusion anti-réductionniste, de manière à éviter un conflit avec la complétude de la physique et le réductionnisme ontologique¹⁷.

IV. L'autonomie des sciences spéciales dans un nouveau cadre réductionniste

Pour résoudre le dilemme que nous venons de présenter, nous proposons ici une stratégie réductionniste *conservative* (non-éliminativiste). En utilisant toujours la biologie comme cas paradigmatique de science spéciale, nous nous efforcerons de montrer dans cette section que la réalisation multiple n'empêche pas l'établissement de liens systématiques entre des concepts fonctionnels abstraits et des concepts physiques. Notre point de départ est la thèse selon laquelle toute différence du point de vue physique entre différents réalisateurs d'une propriété biologique implique une différence fonctionnelle qui est observable du point de vue biologique. Cela présuppose une théorie causale des propriétés (voir aussi la contribution de Michael Esfeld dans ce numéro). Plus précisément, la manière dont les occurrences d'une propriété (décrite par un seul concept) biologique satisfont la définition fonctionnelle de cette propriété est différente s'il existe des différences physiques entre les occurrences de la propriété en question. Cette différence fonctionnelle est détectable en termes biologiques, et, par conséquent, il est possible de construire des sous-concepts fonctionnels des concepts abstraits qui sont co-extensionnels avec des concepts physiques. Ce lien systématique avec la physique, laquelle est complète, garantit la qualité scientifique de ces sous-concepts. La différence entre un concept abstrait et ses sous-concepts est donc seulement son degré d'abstraction, puisque la définition du type fonctionnel saisit seulement les aspects fonctionnels communs à ses différents réalisateurs. Une telle différence dans le degré d'abstraction ne pose cependant pas de problème pour la qualité scientifique des concepts abstraits parce que, même si les concepts biologiques ne sont pas co-extensionnels avec les concepts physiques, on peut tout de même les y réduire au moyen de leurs sous-concepts. Une telle réduction n'est enfin pas éliminativiste, parce que le caractère indispensable de la biologie au sein de notre système scientifique est revendiqué. L'argument est qu'il n'y a pas de concepts physiques qui pourraient mettre en évidence des similarités ontologiques entre des occurrences de propriétés de la même manière que le font les concepts abstraits biologiques. Considérons chaque étape de cette stratégie en détail¹⁸.

¹⁷ Voir Sachse, Christian, *Reductionism in the Philosophy of Science*, Frankfurt, Ontos Verlag, 2007, chapitre III, sections I-III.

¹⁸ Voir Esfeld, Michael et Sachse, Christian, « Reduction by Means of Functional Sub-types », in *International Studies in the Philosophy of Science*, no. 21, vol. 1, 2007, pp. 1-17, et Sachse, Christian, "Reductionism in the Philosophy of Science", Frankfurt, Ontos Verlag, 2007, chapitre II, sections XVIII-XXI.

Premièrement, pour examiner la thèse selon laquelle chaque différence physique implique une différence fonctionnelle, prenons l'exemple des occurrences de gènes appartenant à un même type mais qui sont physiquement différentes en raison de la redondance du code génétique (réalisation multiple). Supposons qu'il existe des occurrences de gènes qui sont décrites par le concept « gène pour la production des protéines de type *PR* ». De manière à prendre en compte la possibilité de la réalisation multiple, disons que la moitié de ces gènes est décrite par le concept physique « P_1 » et que l'autre moitié de ces gènes est décrite par le concept physique « P_2 ». Cette différence physique entre les gènes du type physique « P_1 » et du type physique « P_2 » implique une différence causale. En d'autres termes, il y a une différence dans la manière dont les protéines de type *PR* sont produites. Ce qui nous permet d'affirmer cela, c'est le fait qu'il n'est pas possible de distinguer des occurrences par des concepts (types) différents s'il n'existe pas de différence causale entre les occurrences. Ensuite, pour distinguer les occurrences de gènes réalisées par « P_1 » des occurrences de gènes réalisées par « P_2 », il est nécessaire qu'il existe une différence physique entre les occurrences de gènes réalisées par « P_1 » et celles réalisées par « P_2 ». Il y a donc une différence causale (si on présuppose la théorie causale des propriétés – voir Esfeld dans ce numéro) et cette différence causale est, comme mentionné, une différence dans la manière dont les effets caractéristiques du type de gènes en question (selon sa définition fonctionnelle) sont produits par les occurrences de gènes réalisés respectivement par « P_1 » et « P_2 »¹⁹. Les occurrences de gènes réalisées par « P_1 » interagissent différemment avec leur environnement physique que les occurrences de gènes réalisées par « P_2 ». Un exemple génétique qui se base sur des expériences de laboratoire est le suivant : une différence dans la base d'ADN entre des gènes d'un seul type implique une différence dans le temps et dans la précision de la production des protéines²⁰ ou même une différence dans la fonctionnalité des protéines²¹. En d'autres termes, les différences physiques entre les occurrences de gènes peuvent avoir des effets collatéraux. Par exemple, les occurrences de gènes réalisés par « P_1 » produisent des protéines de type « *PR* » plus rapidement et plus précisément que les gènes réalisés par « P_2 ».

Deuxièmement, on peut facilement concevoir un environnement possible dans lequel une telle différence physique, donc, une différence affectant la manière dont la définition fonctionnelle est satisfaite, mène à une différence qui est détectable du point de vue biologique²². Par exemple, les biologistes peuvent observer une différence

¹⁹ Voir Kim, Jaegwon, « Making Sense of Emergence », in *Philosophical Studies*, no. 95, 1999, pp. 3-36, voir en particulier pp. 17-18.

²⁰ Voir par exemple Bulmer, Michael, « The Selection-mutation-drift Theory of Synonymous Codon Usage », in *Genetics*, 129, 1991, pp. 897-907.

²¹ Voir par exemple Kimchi-Sarfaty, Chava, et al., « A 'Silent' Polymorphism in the MDR1 Gene changes Substrate Specificity », in *Science*, no. 315, 2007, pp. 525-528.

²² Voir Rosenberg, Alexander, *Instrumental Biology or the Disunity of Science*, Chicago, Chicago University Press, 1994, p. 32.

statistique dans la valeur de « fitness » des organismes qui possèdent des gènes réalisés par des configurations d'occurrences de propriétés physiques différentes : un organisme porteur des gènes qui sont décrits par le concept physique « P₁ » possède un avantage sélectif dans un environnement où la vitesse ou la précision de production des protéines en question sont essentielles pour sa survie par rapport à un organisme porteur des gènes du même type mais décrits par le concept physique « P₂ ». C'est une différence fonctionnelle dans le contexte de la sélection naturelle, laquelle constitue le cadre général pour la définition fonctionnelle des propriétés en biologie.

Troisièmement, s'il existe une telle différence fonctionnelle, de manière simplifiée, il existe une différence de « fitness » entre des occurrences de gènes, décrites par le concept « gène pour la production des protéines PR » (« B ») et réalisées par des configurations d'occurrences de propriétés physiques différentes, et on peut, en théorie, concevoir (construire) des sous-concepts différents en termes biologiques ou génétiques : il y a les occurrences du sous-type « B₁ » et les occurrences du sous-type « B₂ ». Par conséquent, on peut établir une connexion systématique (biconditionnelle) entre les sous-concepts « B₁ » et « B₂ » et les concepts physiques « P₁ » et « P₂ » (« B₁ » ⇔ « P₁ », « B₂ » ⇔ « P₂ »). En d'autres termes, il est, en théorie, possible d'établir une co-extension nomologique entre des sous-concepts génétiques et des concepts physiques. L'argument pour la co-extension *nomologique* est le suivant : pour chaque différence physique, qui implique une différence causale-dispositionnelle et donc une différence dans la manière dont la fonction est mise en oeuvre, on peut concevoir un environnement dans lequel cette différence causale-dispositionnelle se manifeste de manière fonctionnelle, à savoir en tant que différence dans la valeur de « fitness » de l'organisme étudié. Pour conclure, la réduction des sous-concepts biologiques à la physique s'avère dès lors possible. A la différence des approches de Lewis et Kim, l'approche proposée se base sur la possibilité que les sous-concepts soient construits *purement* en termes biologiques²³. Dans le cadre de la réduction de Lewis et Kim – réduction dite *species-* ou *structure-specific* – les sous-concepts contiennent des critères physiques. C'est la raison pour laquelle leur modèle suggère finalement l'élimination de la biologie²⁴. Nous allons montrer dans la suite de cette section pourquoi le danger de l'élimination est évité par le modèle que nous proposons.

L'extension d'un concept abstrait est identique à l'extension cumulée de celles de ses sous-concepts. La réduction épistémologique de cet ensemble de sous-concepts à la physique s'avère possible parce que tous les sous-concepts sont nomologiquement

²³ Lewis, David, « Mad Pain and Martian Pain », in Block, Ned, dir., *Readings in the Philosophy of Psychology*. Volume 1, London, Methuen, 1980, pp. 216-222. Réimprimé in Lewis, David, *Philosophical Papers*, Volume 1, Oxford, Oxford University Press, 1983, pp. 122-130. Kim, Jaegwon, *Mind in a Physical world. An essay...*, voir en particulier pp. 93-95. Kim, Jaegwon, *Physicalism, or...*, voir en particulier p. 25.

²⁴ Kim, Jaegwon, « Making sense of Emergence », in *Philosophical Studies*, no. 95, 1999, pp. 3-36, voir en particulier pp. 17-18. Voir également Kim, Jaegwon, *Physicalism, or...*, voir en particulier p. 26 et p. 58.

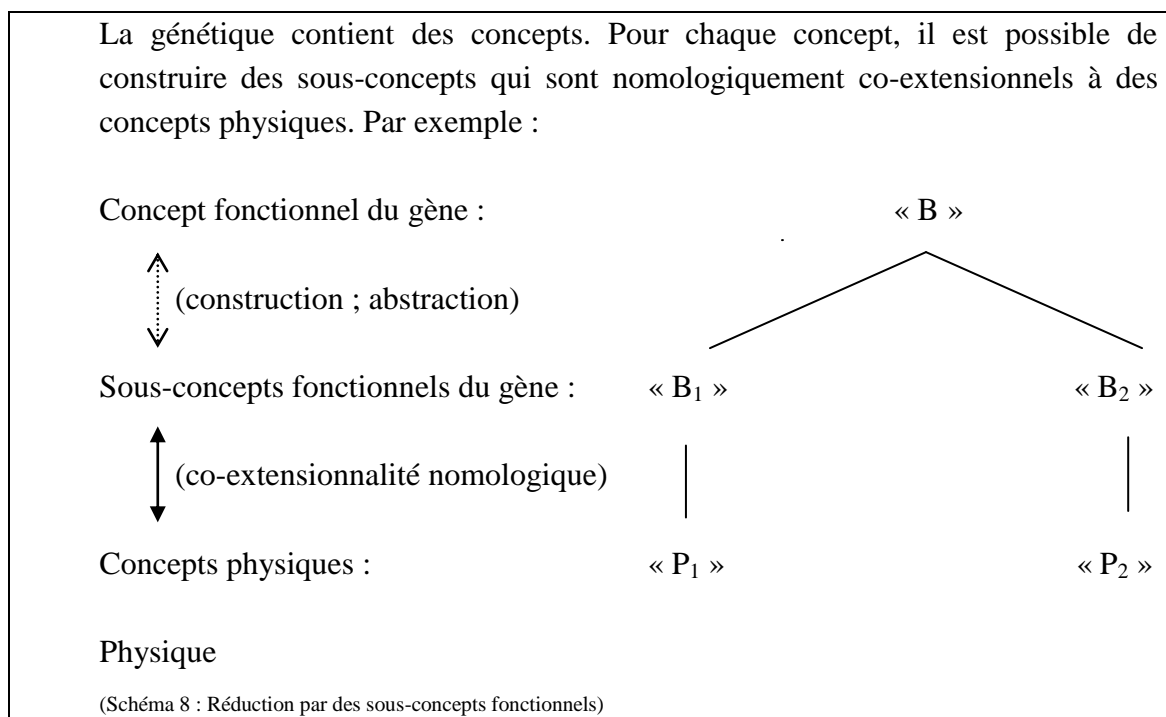
co-extensionnels avec les concepts physiques. En d'autres termes, une génétique construite – qui considère seulement des sous-concepts qui sont nomologiquement co-extensionnels avec les concepts physiques – est réductible à la physique.

Ce résultat intermédiaire, la co-extensionnalité des sous-concepts biologiques et des concepts physiques, est à considérer un peu plus en détail. Il nous faut aborder à présent la question de la qualité scientifique des sous-concepts en question. Les sous-concepts sont liés systématiquement à la physique qui, elle, est complète. Les sous-concepts montrent des similarités causales entre des occurrences de propriétés qui sont des similarités physiques. Le fait d'affirmer la qualité scientifique de ces sous-concepts n'est donc pas en conflit avec le réductionnisme ontologique et la complétude de la physique. Par contre, la co-extensionnalité des sous-concepts biologiques et des concepts physiques suggère l'élimination au niveau épistémologique de cette biologie construite. En effet, dès que deux concepts sont liés bi-conditionnellement, le concept physique est à préférer parce qu'il appartient à une théorie complète. Cependant, notre stratégie consistant à construire des sous-concepts n'a pas de prétention normative, c'est-à-dire que nous ne soutenons pas qu'elle doit être appliquée dans la recherche en biologie ou qu'il faille construire une biologie ne contenant que des sous-concepts (nous ne proposons pas de réductionnisme méthodologique). La question qui nous intéresse concerne le statut des concepts abstraits biologiques dans ce modèle réductionniste recourant à des sous-concepts. Est-ce qu'il est maintenant possible de revendiquer la qualité scientifique pour les concepts abstraits biologiques sans entrer en conflit avec le réductionnisme ontologique et la complétude de la physique ?

Pour répondre à cette question, considérons la différence entre le concept abstrait « gène pour la production des protéines *PR* » et ses sous-concepts. La seule différence entre « gène pour ... » (« *B* ») et ses sous-concepts « *B*₁ » et « *B*₂ » est le degré d'abstraction. En d'autres termes, chaque sous-concept du concept « gène pour la production des protéines *PR* » contient ce concept et, en plus, une spécification fonctionnelle additionnelle telle que, ici de manière simplifiée mais empiriquement bien fondée, « production rapide des protéines ». C'est pour cette raison que l'on peut déduire le concept « gène pour la production des protéines *PR* » de chacun de ses sous-concepts. Par conséquent, la qualité scientifique du concept abstrait comme « gène pour la production des protéines *PR* » est bien fondée : le concept abstrait ne montre pas quelque chose *en dehors* de ses sous-concepts mais il montre une similarité pertinente qui est seulement implicite dans les sous-concepts. Ainsi, attribuer une qualité scientifique fondée aux concepts abstraits biologiques n'est pas, comme nous l'avons montré, en conflit avec le réductionnisme ontologique et la complétude de la physique : puisque seuls les sous-concepts sont co-extensionnels avec des concepts physiques, le concept abstrait « gène pour la production des protéines *PR* » reste indispensable d'un point de vue scientifique parce qu'il n'y a pas de concept physique qui soit co-

extensionnel avec ce concept abstrait. Par conséquent, la stratégie présentée évite le tenant éliminativiste du dilemme de la réalisation multiple.

En d’autres termes, les concepts abstraits de la génétique ou de la biologie en général montrent des similarités pertinentes entre des occurrences de propriétés dans le monde que la physique n’est pas capable de mettre en évidence par ses capacités conceptuelles. Cet avantage épistémologique de la génétique (et de ses concepts abstraits) est revendiqué seulement dans le cadre d’une position réductionniste – chaque concept abstrait est systématiquement lié, via ses sous-concepts, à des concepts physiques. Le schéma suivant donne un survol de cette dernière partie de la stratégie que nous proposons :



En résumé, l’argument en faveur du réductionnisme ontologique est un argument causal qui se base sur les deux prémisses philosophiques que sont, d’une part, le concept de survenance et, d’autre part, celui de complétude de la physique. Si on prend ce point de départ, une position anti-réductionniste qui se base sur l’argument de la réalisation multiple aboutit à un dilemme – épiphénoménisme ou éliminativisme – à l’égard de la biologie et de ses propriétés. Pour éviter ces deux conséquences, nous proposons une stratégie réductionniste non-éliminativiste qui considère la possibilité de la réalisation multiple comme un argument en faveur du caractère indispensable des concepts abstraits biologiques, tels que « gène pour la production des protéines *PR* ». Notre approche est également basée sur un argument causal, à savoir que des différences physiques impliquent des différences fonctionnelles. En d’autres termes, si on peut concevoir pour chaque différence physique un monde dans lequel cette

différence physique implique une différence fonctionnelle, une approche réductionniste conservative est possible. Dans ce cadre réductionniste, la biologie est d'une certaine façon indépendante : une différence physique n'implique pas dans toutes les circonstances, ni de la même façon, une différence biologique. Notre monde est, de façon contingente, dans un état dans lequel les différences physiques entre des occurrences de propriétés biologiques n'impliquent pas *toujours*, ni de la *même manière*, une différence biologique. Pour le montrer avec notre exemple, dans la plupart des conditions environnementales, il n'y a pas de pression sélective qui s'exerce en particulier sur la rapidité de la production des protéines. Ainsi, ce n'est pas seulement pour des raisons pragmatiques qu'en biologie on peut faire abstraction de certaines différences physiques au niveau des réalisateurs physiques des occurrences de propriétés biologiques. En ce sens, la biologie est indépendante des détails physiques sans être en conflit avec le réductionnisme.

La première partie de cet article s'est efforcée de présenter une position générale à l'égard de la relation entre physique et sciences spéciales à la fois réductionniste et conservative. Si nous avons jusqu'ici adressé notre argument au cas de la biologie en particulier, il faut souligner qu'il s'adresse à l'ensemble des sciences dites spéciales. Pour autant que les prémisses de survenance, liant propriétés physiques et propriétés des sciences spéciales, et de complétude de la physique soient assumées dans le cadre d'une théorie causale des propriétés, l'ensemble de la démarche demeure valide. Nous nous proposons dans le reste de cet article de montrer comment appliquer ce cadre théorique générale au cas plus contesté de la psychologie et de la relation que celle-ci entretient avec les neurosciences.

V. La psychologie en tant que science spéciale

La relation entre les propriétés mentales et le monde physique constitue le cœur du problème cartésien de l'insertion de l'esprit dans le monde de la nature. La version contemporaine de ce problème trouve sa formulation canonique dans le problème de l'exclusion causale tel que l'a formulé Jaegwon Kim²⁵. Or, cette formulation repose sur les mêmes prémisses que celles que nous avons considérées dans le cadre de la biologie.

Du fait de l'utilisation de descriptions psychologiques se référant à des propriétés mentales pour expliquer la genèse du comportement, ces dernières sont premièrement supposées avoir une efficacité causale dans la production du comportement. Si ces dernières étaient épiphénoménales, le problème de l'interaction entre esprit et nature ne se poserait simplement pas. Deuxièmement, les propriétés mentales dépendent étroitement du corps pour leur existence en ce qu'elles surviennent sur celui-ci. De surcroît, la causalité étant un processus local, les propriétés mentales surviennent localement sur la base de l'ensemble d'occurrences de propriétés physiques

²⁵ J. Kim, *Mind in a Physical World. An Essay...*

constituant l'individu les instanciant. Enfin et troisièmement, comme expliqué plus haut, la physique est causalement complète. Ainsi, le même argument que celui que nous avons développé plus haut à l'égard des propriétés biologiques s'applique. Si l'on désire sauvegarder la qualité explicative de la psychologie à l'égard de la production du comportement sans pour autant enfreindre le principe de complétude causale du domaine physique, alors il faut considérer les occurrences de propriétés mentales comme identiques à des occurrences de propriétés physiques²⁶. Les descriptions psychologiques se réfèrent alors aux mêmes entités que les descriptions physiques ou neurobiologiques d'un individu.

L'argument de la réalisation multiple s'applique ici également. La thèse de l'identité des occurrences n'implique pas que types psychologiques et physiques soient identiques (que leurs concepts soient co-extensionnels). Il est tout à fait possible que deux individus physiquement ou neurobiologiquement différents instancient la même propriété mentale (tombent sous le même concept psychologique). Par conséquent, les types psychologiques et les types physiques ont une extension différente. Intervient alors à nouveau le dilemme de la réalisation multiple. Soit on considère que les descriptions psychologiques mettent en évidence des similarités ontologiques que les neurosciences ne peuvent saisir, ce qui entre en conflit avec la complétude de la physique et le réductionnisme ontologique, soit on conteste la réalité ontologique des similarités en question, ce qui suggère l'élimination de la psychologie.

Pour éviter ce dilemme, nous nous proposons donc de présenter les modalités d'application de la stratégie discutée précédemment au cas de la relation entre psychologie et neurosciences, ces dernières étant ici interprétées comme cette partie des sciences physiques dédiée à l'étude des mécanismes cérébraux responsables de la production du comportement. Pour ce faire, nous recourons à deux exemples empiriques assurant le bien fondé scientifique de notre démarche. Le premier exemple, le cas du syndrome de déconnexion, vise à illustrer clairement la démarche, alors que le second, la schizophrénie schneidérienne, montrera que cette même démarche peut s'appliquer à des cas nettement plus complexes.

VI. Syndrome de déconnexion inter-hémisphérique

Du point de vue neuropsychologique, les hémisphères du cerveau humain constituent deux entités travaillant de manière relativement autonome. Ainsi, la plupart des informations sensorielles sont traitées en premier lieu par l'hémisphère situé contralatéralement par rapport à l'origine des signaux nerveux. Il en va de même pour la

²⁶ Il existe dans la littérature contemporaine un débat concernant l'acceptabilité de la surdétermination des propriétés mentales par des causes mentales additionnelles. Sans vouloir entrer ici dans ce débat, nous excluons ici cette solution car elle tend à réduire les propriétés mentales à des épiphénomènes. Pour une défense de la surdétermination, voir Harbecke Jens (2008), *Mental Causation – Investigating the Mind's Powers in a Natural World*, Frankfurt (Main), Ontos.

plupart des signaux moteurs efférents. La main droite est par exemple sous le contrôle du cortex moteur de l'hémisphère gauche et vice-versa pour la main gauche.

L'échange d'informations entre les deux hémisphères est toutefois crucial dans le processus conduisant à la production de comportements courants qui nécessitent la coordination des moitiés droite et gauche du corps, tels que marcher, se laver les mains ou encore assembler des pièces d'un puzzle²⁷. Le principal élément permettant l'échange d'information entre les deux hémisphères est le corps calleux. Celui-ci est une large bande médullaire reliant les deux hémisphères du cerveau des mammifères et est composée d'environ 200 à 800 millions de fibres commissurales pouvant être regroupées en trois grandes catégories²⁸. La plupart de ces fibres constituent des connexions topographiques en ce sens qu'elles connectent les aires cérébrales à leur contrepartie contralatérale. Un second groupe de connexions projettent d'une aire donnée vers les zones corticales vers lesquelles l'aire homologue contralatérale projette elle-même. L'aire visuelle V1 du cortex occipital droit est ainsi directement connectée avec les aires visuelles V2 des hémisphères gauche et droit. Enfin, un dernier groupe de connexions inter-hémisphériques projette de manière non systématique vers un ensemble diffus de destinations.

Le syndrome de déconnexion inter-hémisphérique rassemble un ensemble d'anormalités comportementales ayant pour origine neurologique commune des lésions ou une absence totale de corps calleux. De telles lésions peuvent apparaître accidentellement tout au long de la vie de l'individu, notamment en cas de traumatisme ou de tumeur localisés dans la région de la commissure centrale. L'agénésie, partielle ou totale, du corps calleux est l'une des malformations du cerveau les plus répandues, puisque celle-ci touche chez l'homme environ 3 à 7 individus sur 1000²⁹. Elle peut apparaître tant sous la forme d'une pathologie spécifique qu'en tant que partie d'un ensemble plus vaste de malformations. L'absence de corps calleux a en effet été reliée avec de nombreux réarrangements chromosomiques consistants dans plus d'une vingtaine de syndromes de malformation de type autosomique et x-lié³⁰. Enfin, la commissurotomie, l'ablation chirurgicale du corps calleux, a été réintroduite dans les années 1960 par les chirurgiens Philip Vogel and Joseph Bogen comme traitement électif des cas graves d'épilepsie. La réintroduction d'une opération aussi invasive répond à l'interprétation des graves crises d'épilepsie comme étant le résultat de la propagation d'une tempête électromagnétique dans tout le cerveau à partir de zones de départ isolées. L'ablation du corps calleux permet « *de confiner la crise à un côté [du*

²⁷ Pour une description d'un tel cas clinique, voire Kolb, B. and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*. New York, Worth, 2003, pp. 426-427.

²⁸ Kolb, B. and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*. New York, Worth, 2003, p.433

²⁹ M. F. Bedeschi, et al., "Agenesis of the Corpus Callosum: Clinical and Genetic Study in 63 Young Patients", in *Pediatric Neurology*, 34/3, 2006, pp. 186-193.

³⁰ W. B. Dobyns, "Absence Makes the Search Grow Longer", in *American Journal of Human Genetic*, 58, 1996, pp. 7-16.

cerveau] et tend à préserver la conscience durant une attaque [...] et permet au patient de prendre des mesures de précaution ou de contrôle à l'approche de la crise »³¹.

Les conséquences comportementales de la commissurotomie ont été largement étudiées depuis la réintroduction de cette opération. Un des résultats les plus surprenants de ces études est le constat que cette intervention chirurgicale n'a finalement que peu de conséquences quant à la vie quotidienne des patients³². Ceux-ci mettent en règle générale une année pour se remettre de l'opération et, après un à deux ans, un examen médical conventionnel ne permet plus de les distinguer comportementalement de sujets normaux. Ils sont alors à mêmes de reprendre leur activité scolaire ou professionnelle³³. Des tests cognitifs spécifiques sont ainsi nécessaires pour détecter des différences comportementales entre sujets sains et patients commissurotomisés afin de montrer que chez ces derniers, chaque hémisphère traite individuellement la moitié des informations traitées par le cerveau dans son entier, ce qui laisse supposer que chez ces patients, « *chaque hémisphère [semble] avoir ses propres sensations, percepts, pensées et souvenirs qui ne sont pas accessibles pour l'autre hémisphère* »³⁴.

Les tests classiques permettant de discriminer patients et sujets sains visent à déterminer si une information présentée dans un premier temps à un seul hémisphère est à disposition d'un seul ou des deux hémisphères pour une action motrice. Un test très simple consiste par exemple à demander à un individu dont les yeux ont été bandés de toucher un objet avec la main gauche et de trouver un objet similaire avec la main droite³⁵. Étant donné que la main gauche est sous le contrôle somesthésique et moteur du seul hémisphère droit, et vice-versa pour la main droite, en l'absence de communication inter-hémisphérique, les patients commissurotomisés sont incapables d'appareiller les objets correctement.

De nombreuses procédures expérimentales ont été développées dans le but d'isoler les spécificités comportementales de ce type de patients dans l'exécution de différentes tâches motrices et modalités sensorielles. Sans vouloir ici considérer l'ensemble ces données, un exemple plus complet portant sur les capacités visuelles de ces patients nous permettra dans la section suivante de présenter les lignes générales de l'application de la stratégie de réduction fonctionnelle discutée ici.

La connectivité nerveuse du système visuel humain est croisée de manière très spécifique. L'information visuelle en provenance de chaque hémichamp visuel est traitée par le cortex visuel occipital contralatéral. Puisque ce croisement concerne les hémichamps visuels et non les yeux eux-mêmes, les structures nerveuses connectant

³¹ Traduction par les auteurs de R. W. Sperry, M. S. Gazzaniga, and J. E. Bogen, "Interhemispheric Relationship: The Neocortical Commissures; Syndrome of Hemisphere Disconnection", in P. J. Vinken and G. W. Bruyn (Ed.), *Handbook of Clinical Neurology*, Amsterdam, North-Holland Publishing Company, 1969, p. 274.

³² *Ibidem*, p. 275.

³³ Kolb, B. and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*, p. 433.

³⁴ *Ibidem*, p. 433.

³⁵ *Ibidem*, p. 433.

réтины et cortex visuel sont relativement complexes. Les signaux optiques provenant de l'hémichamp visuel gauche sont dans un premier temps convertis en signaux nerveux par la moitié droite de la rétine de chaque œil et sont ensuite acheminées au corps géniculé droit *via* le chiasme et les nerfs optiques. A partir de là, les informations sont transmises par les radiations optiques à la couche IV de l'air V1 du cortex visuel primaire.

Dans le cas d'individus normaux, les informations provenant de l'un des deux hémichamps visuels sont toutefois à disposition des deux hémisphères. Deux mécanismes contribuent à ce mélange d'informations en dépit du cloisonnement des deux moitiés du système visuel. Premièrement, le corps calleux permet l'échange d'informations entre les deux hémisphères en connectant, comme décrit plus haut, des aires cérébrales homologues situées de chaque côté du cerveau. Deuxièmement, au contraire de ce que laisse supposer l'apparente stabilité de notre perception visuelle, les globes oculaires entretiennent constamment des micro-mouvements de balayage. Ces mouvements ont pour fonction principale d'augmenter la zone d'observation détaillée couverte par la fovéa, mais aussi de maintenir la perception visuelle dans le temps, puisque les champs récepteurs qui recouvrent l'ensemble de la rétine sont avant tout sensibles aux variations de luminance. Ce faisant, ils permettent également de répartir l'information visuelle provenant de chaque hémichamp dans les deux moitiés du système visuel. Ainsi, une partie de l'information provenant de l'un des hémichamps visuels se retrouve à disposition de l'hémisphère ipsilatéral simplement parce que les yeux ont modifié leur position durant une très courte période.

L'existence de ces micro-mouvements de balayage explique pourquoi les patients souffrant du syndrome de déconnexion manifestent, au niveau de la perception visuelle, des performances relativement normales dans leur vie quotidienne. En dépit de l'absence de communication inter-hémisphérique, une partie de l'information visuelle relative à chaque hémichamp reste néanmoins à disposition de l'hémisphère ipsilatéral, et peut ainsi être utilisée par l'hémisphère en question pour l'exécution de différentes tâches cognitives. Les micro-mouvements de balayage constituent ainsi un puissant mécanisme de compensation des conséquences du syndrome de déconnexion qui tend, lui, à isoler les deux hémisphères.

Comme brièvement mentionné plus haut, le diagnostic de dommages ou d'absence de corps calleux à partir de critères strictement comportementaux requiert l'utilisation de tests cognitifs spécifiques. Dans le cas de la perception visuelle, les tests classiques en question fonctionnent de la manière suivante³⁶. Le sujet est assis à table et fait face à un écran ajustable disposé de manière à ce que le sujet ne puisse pas voir ses propres mains. Des images sont ensuite brièvement flashées sur l'une ou l'autre moitié

³⁶ R. W. Sperry, M. S. Gazzaniga, and J. E. Bogen, "Interhemispheric Relationship: The Neocortical Commissures; Syndrome of Hemisphere Disconnection", in P. J. Vinken and G. W. Bruyn (Ed.), *Handbook of Clinical Neurology*, pp. 275-277.

de l'écran alors que le sujet doit s'efforcer de conserver son attention visuelle sur un point rouge au centre de l'écran. Les images sont flashées durant moins d'un dixième de seconde, de manière à ce que les micro-mouvements de balayage ne puissent pas répartir l'information visuelle dans les deux hémisphères. Ainsi des stimuli visuels peuvent être présentés sélectivement à un hémisphère spécifique ou aux deux. L'expérimentateur tente ensuite de déterminer si l'information visuelle présentée au sujet est à disposition de l'un ou des deux hémisphères pour l'exécution d'une action motrice pouvant consister en la formulation d'un rapport verbal ou l'appréhension à l'aveugle avec les mains d'un objet en rapport avec le stimulus.

Dans l'exécution de telles tâches, les patients souffrant du syndrome de déconnexion se comportent de manière générale normalement si l'information visuelle est présentée dans l'hémichamp visuel droit. Ils peuvent sans problèmes décrire les stimuli présentés. En revanche, si les stimuli sont présentés dans l'hémichamp visuel gauche au sein d'une séquence aléatoire de stimulations, les réponses verbales des patients suggèrent qu'ils ne sont pas perçus. Toutefois, si le temps de présentation est augmenté, les patients sont capables de décrire le stimulus, grâce au mécanisme de compensation discuté ci-dessus. Ce type de conséquences de la commissurotomie est tout à fait définitif.

De tels résultats pourraient être interprétés à première vue comme indiquant un déficit propre au système visuel droit. Des tests supplémentaires excluent toutefois cette possibilité. Lorsqu'une réponse non verbale est demandée pour évaluer les perceptions des patients, il est évident que les patients comprennent la tâche demandée et qu'ils perçoivent et identifient le stimulus, puisqu'ils peuvent par exemple saisir de la main gauche un objet correspondant à l'image présentée dans l'hémichamp visuel gauche sans pour autant pouvoir le nommer. L'impossibilité de nommer les images en question provient ici du fait que les principales aires responsables de la production du langage – les aires de Broca et de Wernicke – sont situées dans l'hémisphère gauche qui, en l'absence de communications inter-hémisphériques, n'a pas accès au stimulus. Ceci démontre que, d'une part, le système visuel des patients est parfaitement fonctionnel. D'autre part, comme les mêmes patients sont capables de décrire des stimulations visuelles présentées à droite, leurs capacités purement linguistiques sont, elles aussi, intactes.

De manière plus générale, les handicaps comportementaux des patients souffrant de déconnexion inter-hémisphérique ne peuvent pas être interprétés comme résultant de déficits purement perceptifs ou moteurs. Les déficits sensoriels peuvent être systématiquement exclus, car les patients peuvent toujours utiliser un côté de leur corps pour manifester leur perception. Le même argument vaut pour la question des déficits moteurs. Si les patients ne peuvent répondre comme demandé par la tâche expérimentale à une stimulation donnée, ils peuvent néanmoins toujours effectuer l'action demandée sur la base d'une autre stimulation. L'origine neurale du syndrome

de déconnexion apparaît ainsi clairement comme résultant non d'un défaut dans les aires motrices ou sensorielles primaires mais bien d'un déficit dans le traitement global de l'information entre celles-ci et plus spécifiquement comme d'une absence de communication entre les deux hémisphères.

Permettez-nous de résumer cette section en soulignant qu'une large variété de tests visuels de nature analogue confirme la conclusion que, dans le cas des patients commissurotomisés, l'information visuelle ayant son origine dans l'un des deux champs visuels est traitée en isolation par l'un des deux hémisphères. Les interactions entre les éléments présents à l'intérieur de l'un des hémichamps visuels sont préservées, mais l'intégration des informations provenant de chaque hémichamp fait défaut. Les patients souffrant du syndrome de déconnexion sont toutefois capables de vivre une vie quotidienne relativement normale, notamment grâce au puissant mécanisme de compensation que constituent les micro-mouvements de balayage des yeux. Ceci n'exclut pas pour autant de discrètes spécificités comportementales, sur lesquelles nous allons nous appuyer dans la section suivante, où il s'agira de présenter l'application de la stratégie de réduction fonctionnelle par introduction de sous-types.

VII. Application de la stratégie de réduction fonctionnelle par sous-types au cas du syndrome de déconnexion

Afin de développer l'application de la stratégie de réduction fonctionnelle à la relation entre psychologie et neuropsychologie, introduisons l'exemple suivant. Soit deux paires d'individus regardant la télévision et discutant le contenu d'un très bon film. Chaque individu est par là susceptible d'être décrit par la psychologie en tant que percevant l'écran de la télévision et son contenu. Les quatre individusinstancient ainsi la même propriété mentale perceptuelle. Comme dans le cas présenté plus haut, la première étape de la réduction fonctionnelle consiste en la définition de cette propriété mentale particulière en termes purement fonctionnels. Celle-ci peut être fonctionnalisée de la manière suivante :

(A) Un sujet S perçoit visuellement un objet x ssi, lorsqu'il regarde attentivement x et lorsqu'il désire formuler une description verbale de x , S décrit verbalement x avec succès.

Il y a bien sûr énormément d'autres cas de perception puisque le sens commun attribue ce type d'état dans beaucoup d'autres scénarios. Par exemple, la capacité de former une description verbale n'est pas une condition nécessaire pour l'attribution d'une propriété perceptuelle. Mais notre propos n'est pas ici de fournir une définition fonctionnelle exhaustive de la perception. Une formulation générique de telles

définitions fonctionnelles est fournie par la formalisation des termes théoriques de Ramsey-Lewis :

(F) $(\exists x)(\exists y)(\exists z)((x \text{ est causé par les conditions environnementales } c_1 \vee c_2 \vee \dots \vee c_n) \ \& \ (x \text{ tend à causer l'instanciation des propriétés mentales } y \vee z) \ \& \ (x \text{ tend à causer les comportements } b_1 \vee b_2 \vee \dots \vee b_n))$

La seconde étape de la réduction fonctionnelle consiste en l'identification de réalisateurs physiques – neuropsychologiques dans notre cas – de la description (A). Assumons à présent qu'à l'intérieur de chacune des paires d'individus présentes dans notre exemple, l'un des deux individus est un patient souffrant du syndrome de déconnexion. Ainsi formulée, la description (A) s'applique néanmoins à l'ensemble des individus considérés. Nous sommes donc dans un cas clair de réalisation multiple, puisque une seule et même description fonctionnelle s'applique à des individus qui tombent néanmoins sous deux descriptions neuropsychologiques différentes.

A ce stade, introduisons finalement dans notre exemple le fait que la première paire d'individus regarde un film parfaitement normal. La seconde paire d'individus regarde, elle, exactement le même film, à ceci près que cette version du film contient de très brèves apparitions d'un objet particulier sur la gauche de l'écran.

Puisque les patients souffrant du syndrome de déconnexion diffèrent de sujets normaux du point de vue neuropsychologique, leurs cerveaux respectifs ne traitent pas de la même manière l'information visuelle. Comme expliqué dans la section précédente, de telles différences cognitives peuvent être détectées comportementalement dans des circonstances particulières. Dans le cadre de notre exemple, les circonstances en question peuvent tout à fait être les suivantes. Dans la première paire, l'individu normal et le patient se comportent de manière tout à fait comparable. La situation est cependant différente dans le cas de la seconde paire. Le sujet normal perçoit et commente le film en étant tout à fait capable de se référer aux brèves apparitions caractérisant cette version du film. En revanche, le patient atteint du syndrome de déconnexion est certes capable de percevoir et de commenter le film, mais sans toutefois être à même de se référer aux brèves apparitions en question. De telles spécificités comportementales ne constituent dans ce contexte rien d'autre que les effets collatéraux discutés dans la section IV et leur prise en compte permet de construire les sous-types fonctionnels au travers de l'affinement du grain des descriptions des propriétés fonctionnelles. Nous obtenons ainsi deux nouvelles descriptions fonctionnelles (A') et (A'') étroitement liées à notre définition (A) de départ :

(A') Un sujet *S* perçoit visuellement un objet *x* ssi, lorsqu'il regarde attentivement *x* et lorsqu'il désire formuler une description verbale de *x*,

S décrit verbalement *x* avec succès en faisant possiblement référence aux brèves apparitions sur la gauche de l'écran.

(A'') Un sujet *S* perçoit visuellement un objet *x* ssi, lorsqu'il regarde attentivement *x* et lorsqu'il désire formuler une description verbale de *x*, *S* décrit verbalement *x* avec succès en excluant systématiquement toute référence aux brèves apparitions sur la gauche de l'écran.

La description générale (A) s'appliquait à l'ensemble des individus considérés dans notre exemple. La description (A') ne s'applique, elle, en revanche qu'au sujet normal de la seconde paire. Enfin, la description (A'') ne s'applique par contraste, dans la seconde paire, qu'au patient. Ainsi, l'extension respective des descriptions (A') et (A'') est plus petite que celle de la description originelle (A). En d'autres termes, la prise en compte purement fonctionnelle des spécificités comportementales permet d'affiner le grain des descriptions psychologiques sans pour autant recourir à des critères physiques pour individuer les sous-descriptions fonctionnelles³⁷.

Soulignons ici que ce résultat est précisément celui qui est requis par la stratégie de réduction fonctionnelle présentée dans cet article. L'adjonction de spécifications fonctionnelles aux descriptions psychologiques standards permet d'en réduire l'extension. Il est clair que la description (A'') ne s'applique pas seulement au patient souffrant du syndrome de déconnexion puisqu'il existe de nombreuses autres raisons pour lesquelles un individu pourrait se trouver dans l'incapacité de formuler une description faisant référence aux apparitions distinguant les deux situations de notre exemple. Néanmoins il est possible d'introduire dans la description (A'') des spécifications fonctionnelles supplémentaires caractérisant les dispositions comportementales des patients déconnectés de manière à affiner davantage le grain de description et d'en réduire ainsi encore l'extension. Par exemple, puisque les individus souffrant de déconnexion ont des souvenirs et des dispositions motrices particulières³⁸, le patient de la seconde paire d'individu est parfaitement capable de dessiner de la main gauche l'objet apparaissant brièvement ou d'en sélectionner un similaire dans une collection grâce au seul toucher. En affinant de la sorte le grain de nos descriptions fonctionnelles, il est possible de saisir précisément les dispositions comportementales spécifiques à de tels patients. En d'autres termes, l'introduction de sous-types fonctionnels permet d'atteindre la condition nécessaire à la réduction inter-théorique considérée comme la plus problématique, à savoir la co-extensionnalité entre les descriptions de haut et de bas niveaux. A ceci s'ajoute le fait que les neurosciences expliquent pourquoi les individus tombent sous ces descriptions fonctionnelles

³⁷ C'est là principalement ce qui distingue notre position de celle développée par Kim (voir Kim, Jaegwon, *Physicalism, or...*). Individuant les sous-descriptions fonctionnelles sur la base de critères physiques, Kim ne sort pas du dilemme de la réalisation multiple.

³⁸ B. Kolb and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*, p. 433.

présentent les spécificités comportementales qui les caractérisent. Ainsi, les deux conditions permettant la réduction sont remplies : co-extensionnalité et explication.

Il convient ici de souligner un point important. Le réductionnisme épistémologique est souvent compris comme impliquant le remplacement ou l'élimination de la théorie réduite par des généralisations de niveau inférieur considérées comme plus générales, plus fiables et plus précises. L'argument soutenant une telle interprétation consiste en ceci qu'en cas de réduction, les similarités dégagées par les descriptions de niveau supérieur sont perdues car les différents réalisateurs possibles d'un même état fonctionnel n'ont possiblement rien en commun du point de vue des catégories de la théorie à laquelle on réduit la théorie de départ.

Dans le cadre du modèle de réduction fonctionnel présenté, ce n'est absolument pas le cas. Celui-ci n'entraîne ni remplacement ni élimination. Sur le plan ontologique, ce modèle de réduction est associé à la théorie de l'identité des occurrences. Or, l'identité est symétrique et donc conservative. En soutenant que tout ce qui existe est ou est identique avec des occurrences de propriétés physiques, le réductionnisme ontologique n'élimine rien. Si un objet tombe sous une description physique et une description mentale, alors cet objet constitue tant une occurrence de propriétés mentales que de propriétés physiques, sans privilège particulier de l'une ou l'autre des deux modalités descriptives. Sur le plan épistémologique, la stratégie de réduction fonctionnelle n'a également aucune conséquence éliminativiste. La description générale de départ (A) n'est pas éliminée. Elle est préservée dans toute sa qualité descriptive en tant que description de haut niveau ne retenant que les spécifications fonctionnelles communes aux définitions (A') et (A''). En d'autres termes, la description (A) entretient une relation d'abstraction avec les descriptions (A') et (A'') puisqu'elle abandonne les spécifications fonctionnelles relatives aux conditions critiques de manifestations permettant de discriminer fonctionnellement les différents réalisateurs pour ne conserver que les spécifications portant sur les cas généraux. De cette manière, l'introduction de sous-descriptions fonctionnelles permet d'établir une position réductionniste mais conservative à l'égard des généralisations de haut niveau de la psychologie, puisque les traits communs aux différents réalisateurs d'une propriété psychologique sont caractérisés de manière fonctionnelle.

VIII. Les conditions critiques de manifestation et la quête de la localisation des fonctions

La possibilité de trouver des conditions de test adéquates permettant de discriminer fonctionnellement les différents réalisateurs d'une même propriété psychologique au travers des effets collatéraux les caractérisant respectivement est cruciale pour l'approche présentée. De manière plus formelle, il est ici requis que pour chaque réalisateur p_n tombant sous une description fonctionnelle F , il y ait au moins un

ensemble de conditions environnementales possibles dans lesquelles les particularités caractérisant p_n dans la manière dont il produit l'effet caractéristique de F soient observables au travers du comportement. La question demeurant à ce stade de notre développement est de savoir s'il est possible de trouver de telles conditions environnementales pour toute spécificité relevant du point de vue de la neuropsychologie, à savoir pour tout type de réalisateur neuropsychologique d'une propriété mentale F .

Une manière relativement triviale de traiter cette question consiste à revenir sur ce qui caractérise la neuropsychologie elle-même. Kolb and Whishaw définissent celle-ci comme étant « *l'étude de la relation entre le cerveau humain et le comportement* » dont « *le but central est le développement d'une science du comportement humain basée sur le fonctionnement du cerveau humain* »³⁹. De ce point de vue, la neuropsychologie est elle-même adossée à cette prémisse fondamentale selon laquelle il existe une relation systématique entre spécificités cérébrales et comportementales. Ceci a pour conséquence directe que les types neuropsychologiques sont construits de manière à regrouper sous une description commune différents types de spécificités comportementales qui partagent une cause particulière au niveau neurologique. Il est alors évident que pour tout type neuropsychologique réalisant une propriété mentale donnée, il existe alors des spécificités comportementales dans l'exécution de tâches comportementales particulières.

Une manière plus sophistiquée de traiter cette même question consiste à s'intéresser de manière détaillée à la pratique scientifique. Une tâche tout à fait représentative du travail des neuropsychologues est la localisation de fonctions cognitives dans le cerveau. Dans le but de dresser une carte fonctionnelle du cerveau, leurs expérimentations visent à créer un contraste statistiquement significatif entre des sujets de contrôle et des sujets pathologiques ou médicamenteux de manière à illustrer l'impact causal de particularités neurologiques sur la production du comportement. En d'autres termes, ces chercheurs s'efforcent de créer un contraste entre les deux paires suivantes de corrélation : d'un côté une corrélation entre le niveau de performances présenté par des sujets de contrôle dans l'exécution d'une tâche cognitive donnée et les structures cérébrales de ceux-ci ; de l'autre côté une corrélation entre la production d'un niveau de performance comportementale différent par des sujets testés dans l'exécution de la même tâche et leurs structures neurologiques particulières. Si une telle corrélation peut être isolée, les particularités neurologiques des sujets testés nous renseignent quant au lieu de l'exécution de la tâche cognitive testée.

La localisation des fonctions cognitives a été poursuivie durant une longue période en étudiant *post mortem* les particularités de cerveau appartenant à des sujets ayant connu de leur vivant diverses pathologies cognitives. L'histoire des neurosciences contient de nombreux exemples d'application de cette méthode, dont le plus classique

³⁹ B. Kolb and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*, p. 2.

est certainement la double dissociation entre la fonction lexicale de l'aire de Wernicke et la fonction syntactique de l'aire de Broca. Les lésions de l'aire de Wernicke sont responsables de handicaps dans la compréhension du langage et dans la production de séquences complexes et signifiantes de mots. Des lésions de l'aire de Broca entraînent, elles, des difficultés au niveau de la syntaxe et de la fluidité de la parole ainsi que dans la compréhension de phrases grammaticalement complexes. La mise en évidence de telles dissociations demeure la démarche fondamentale dans la vaste entreprise qu'est l'étude de la localisation des fonctions cognitives, même si les récents développements de l'imagerie moderne multiplient les possibilités d'investigation. L'imagerie permet aux scientifiques d'observer immédiatement l'activité neurale générée par l'exécution d'une tâche cognitive particulière. Il est ainsi possible d'établir plus directement une corrélation entre les résultats comportementaux d'un groupe de patients dans l'exécution d'une tâche cognitive et les particularités neurales qui les caractérisent. Mais ceci ne change en rien la démarche de fond consistant à corrélérer spécificités neurales et comportementales.

L'établissement de telles corrélations rencontre deux difficultés fondamentales. La première consiste en ceci que les scientifiques doivent s'efforcer de développer un modèle correct de la manière dont le cerveau traite l'information de manière à hypothétiser des fonctions cognitives existantes⁴⁰ qu'ils peuvent ensuite étudier. La phrénologie du XIX^{ème} siècle illustre tout à fait ce qui peut constituer une taxonomie erronée de fonctions cognitives. Celle-ci postulait en effet que certaines zones cérébrales étaient spécialisées dans des tâches telles que par exemple l'amitié ou le calcul mathématique, d'où vient la célèbre expression « avoir la bosse des maths ». La phrénologie partage avec la neuropsychologie contemporaine cette idée que chaque zone du cortex a une fonction déterminée mais en diffère non seulement en considérant que ceci donnait lieu à des protubérances crâniennes mais surtout au travers de la taxonomie de fonctions prise en compte. La seconde difficulté consiste ni plus ni moins à concevoir un dispositif expérimental permettant de tester de tels modèles.

Prenons ici à titre d'exemple de neuropsychopathologie complexe le cas de la schizophrénie schneidérienne et les récents développements dans l'identification des mécanismes neuraux de la reconnaissance des actions auto-générées par l'agent. Du point de vue cognitif, l'hypothèse du monitoring central constitue le cadre théorique le plus courant dans l'analyse des processus de reconnaissance de ce type d'action. Sans entrer dans les détails, l'idée générale est que « *chaque fois que les centres moteurs génèrent un signal sortant pour la production d'un mouvement, une copie de cette commande (« la commande efférente ») est stockée. Les signaux ré-afférents entrant sont comparés avec la copie. Si une différence est observée entre les deux types de signaux, une nouvelle commande est générée, jusqu'à ce que le produit du mouvement*

⁴⁰ C. Craver, *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience*, Oxford, Clarendon Press, 2007, p. 128.

corresponde au mouvement désiré »⁴¹. La reconnaissance des actions auto-générées est donc basée selon ce modèle sur la concordance entre le résultat attendu d'une intention d'action et son résultat effectif.

La déficience de ces mécanismes conduit aux anormalités cognitives caractérisant la schizophrénie, telles que les hallucinations auditives, les insertions de pensées, le sentiment d'influence d'autrui sur ses pensées, actions et émotions⁴², mais aussi à des perceptions altérées, délusions et croyances qui déforment la réalité⁴³. Par exemple, les hallucinations auditives consistant dans le fait d'entendre des voix sont expliquées dans le cadre de l'hypothèse de monitoring central comme une incapacité pathologique du patient à reconnaître son propre discours intérieur comme étant le sien⁴⁴. Les aires motrices sont engagées hors-ligne dans la production d'une représentation linguistique intérieure, mais le cortex auditif interprète cette activité comme si elle était le produit d'une cause externe⁴⁵. La schizophrénie apparaît ainsi comme une pathologie affectant la communication et la comparaison d'informations entre les aires motrices et sensorielles.

De nombreux programmes de recherche ont tenté de localiser les mécanismes responsables de la reconnaissance des actions auto-générées⁴⁶. Farrer et al.⁴⁷ ont récemment mis sur pied un protocole expérimental dont l'idée centrale est de demander au sujet l'exécution d'une tâche qui soit aussi proche que possible de la fonction cognitive utilisée, à savoir la comparaison entre résultat attendu de la commande motrice et résultat effectif. La tâche en question était pour le sujet de juger si oui ou non les images affichées sur un écran étaient les effets de ses actions sur un joystick placé en dehors de son champ de vision. L'expérimentateur pouvait, lui, introduire un degré variable de discrédance entre les actions du patient et les mouvements affichés sur l'écran.

Sans entrer dans davantage de détails, nous voudrions souligner les deux points suivants. Premièrement, l'imagerie cérébrale permet aux neuroscientifiques de corrélérer l'exécution d'une tâche expérimentale avec une activité neurale particulière. Dans le cas des études de l'équipe de Farrer, une corrélation a été dégagée chez des sujets sains entre le degré de contrôle du retour visuel par les sujets et l'activation du gyrus

⁴¹ M. Jeannerod, "The Mechanism of Self-Recognition in Humans", in *Behavioural Brain Research*, 143, 2003, p. 8.

⁴² C. Farrer, et al., "Neural Correlates of Action Attribution in Schizophrenia", in *Psychiatry Research: Neuroimaging*, 131, 2004, pp. 31.

⁴³ B. Kolb and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*, p. 724.

⁴⁴ N. Franck, et al., "Defective Recognition of One's Own Actions in Patients with Schizophrenia", in *American Journal of Psychiatry*, 158, 2001, pp. 454-459.

⁴⁵ M. Jeannerod, "The Mechanism of Self-Recognition in Humans", in *Behavioural Brain Research*, 143, 2003, p. 14.

⁴⁶ Pour un historique des recherches, voir M. Jeannerod, "The Mechanism of Self-Recognition in Humans", in *Behavioural Brain Research*, 143, 2003, pp. 1-15.

⁴⁷ C. Farrer, et al., "Neural Correlates of Action Attribution in Schizophrenia", in *Psychiatry Research: Neuroimaging*, 131, 2004, pp. 31-44.

angulaire droit, en particulier les aires Brodman 39 et 40 à droite⁴⁸. L'activation maximale des aires en question intervient lorsque les mouvements affichés sont totalement déconnectés des manipulations du joystick effectuées par le sujet.

Cependant, et c'est là notre second point, une telle corrélation n'établit par fermement que l'aire considérée est responsable de la bonne exécution de la tâche consistant à juger si oui ou non les résultats affichés sont le résultat de ses propres actions. C'est seulement lorsque que l'on peut comparer les résultats obtenus avec des sujets exécutant correctement la tâche avec ceux de patients présentant un déficit clair dans l'exécution de la tâche et que la relation entre l'activité neurale et le comportement prend la forme d'une dépendance contrefactuelle qu'il est possible de considérer l'activité neurale localisée comme une cause authentique du comportement observé. Or, c'est précisément là le résultat obtenu par le groupe de recherche de Farrer. Les patients schizophrènes effectuent avec un taux d'erreur considérable la tâche expérimentale, ce qui témoigne clairement des difficultés qu'ils rencontrent dans la distinction entre des mouvements auto-générés et des mouvements influencés, pour tout ou partie, par une cause externe. Ces mêmes patients montrent, au contraire des sujets de contrôle, une activité aberrante du gyrus angulaire droit dans ces mêmes conditions expérimentales⁴⁹. Ce résultat confirme ainsi l'hypothèse selon laquelle le gyrus angulaire droit serait le lieu de la comparaison entre copie efférente et signal ré-afférent⁵⁰. Qui plus est, ce résultat est consistant avec les conséquences de graves lésions du lobe pariétal droit, puisque de telles lésions sont notamment responsables du syndrome de négligence unilatéral, dont les symptômes comprennent entre autres des difficultés au niveau de la discrimination soi/monde. De tels patients dénieient fréquemment la propriété de la partie gauche de leur propre corps⁵¹ et éprouvent des difficultés dans la reconnaissance d'action auto-générées⁵².

Ces résultats confirment le modèle courant d'explication de la schizophrénie. Dans le contexte de la théorie du monitoring central, hallucinations et délusions sont interprétées au niveau cognitif en termes d'échec dans le processus d'intégration de la perception et l'action. Au niveau physiologique, ce déficit cognitif est expliqué par une connectivité corticale anormale, qui « *interrompt la modulation par les régions frontales [de l'activité] des régions postérieures du cerveau impliquées dans le*

⁴⁸ M. Jeannerod, "The Mechanism of Self-Recognition in Humans", in *Behavioural Brain Research*, 143, 2003, p. 8.

⁴⁹ C. Farrer, et al., "Neural Correlates of Action Attribution in Schizophrenia", in *Psychiatry Research: Neuroimaging*, 131, 2004, pp. 37 et 42.

⁵⁰ M. Jeannerod, "The Mechanism of Self-Recognition in Humans", in *Behavioural Brain Research*, 143, 2003, p. 9.

⁵¹ E. Deprati, et al., "Recognition of Self Produced Movement in a Case of Severe Neglect", in *Neurocase*, 6, 2000, pp. 183-201.

⁵² A. Sirigu, et al., "Perception of Self-Generated Movement Following Left Parietal Lesion.", in *Brain*, 122, 1999, pp. 1867-1874.

traitement des conséquences sensorielles d'une action »⁵³, ce qui rend problématique l'identification de la source de la perception. L'absence incriminée de modulations provoque ainsi une interprétation erronée des conséquences d'une action auto-générée : les aires sensorielles primaires interprètent ces conséquences comme si celles-ci étaient l'effet d'une cause externe⁵⁴.

Si nous revenons maintenant à la stratégie de réduction fonctionnelle présentée dans cet article, nos développements relatifs au cas nettement plus complexe de la schizophrénie schneidérienne nous permettent de formuler les commentaires suivants. Premièrement, un tel programme de recherche présuppose et s'appuie simplement sur la possibilité de trouver des conditions critiques de test qui permettent de valider nos hypothèses relatives aux sous-fonctions cognitives présupposées par le modèle du monitoring central en matière de reconnaissance des actions auto-générées. Comme expliqué plus haut, cette possibilité est *ipso facto* ouverte pour la schizophrénie et les autres neuropsychopathologies, étant donné que tous ces cas se rapportent à des handicaps observables du point de vue comportemental. Les difficultés soulevées par de telles recherches sont toutefois multiples. Il s'agit, d'une part, de formuler des hypothèses exactes concernant les sous-fonctions qui sous-tendent l'exécution de la capacité comportementale étudiée et, d'autre part, de mettre sur pied un protocole expérimental qui saisisse et teste précisément la partie incriminée du comportement, dans le but de corrélérer les résultats comportementaux avec une sous-fonction cognitive isolée aussi précisément que possible et dont le corrélat neural est observable.

Or, et c'est là notre seconde remarque, les conditions critiques nécessaires à l'établissement de telles corrélations ne sont rien d'autre que les conditions saillantes de manifestation des effets collatéraux requis par la stratégie de réduction fonctionnelle discutée ici. Dans le cas de la schizophrénie, sujets normaux et patients sont clairement distinguables au travers de leurs performances dans les tests utilisés par l'équipe de Farrer. Or, comme les résultats de ces tests sont saillants comportementalement, ils peuvent être employés pour rendre les définitions fonctionnelles des propriétés mentales plus précises, comme dans le cas du syndrome de déconnexion. Les définitions fonctionnelles issues de l'intégration des effets collatéraux propres à des individus schizophrènes seront bien entendu nettement plus complexes que celles que nous avons présentées pour le cas du syndrome de déconnexion. Mais c'est là cependant un problème pratique qui ne saurait se transformer en problème général de principe.

Il semble dès lors que les conditions présupposées par la stratégie de réduction fonctionnelle par sous-types de la psychologie à la neuropsychologie sont remplies.

⁵³ C. Farrer, et al., "Neural Correlates of Action Attribution in Schizophrenia", in *Psychiatry research: Neuroimaging*, 131, 2004, p. 41.

⁵⁴ S. J. Blackmore, C. Frith, and D. Wolpert, "Spatio-Temporal Prediction Modulates the Perception of Self-Produced Stimuli", in *Journal of cognitive neurosciences*, 11, 1999, pp. 551-559. Voir aussi C. D. Frith and R. Dolan, "The Role of the Prefrontal Cortex in Higher Cognitive Functions", in *Brain Research: Cognitive Brain Research*, 7, 1-2, 1996, pp. 172-161.

D'une part, la possibilité de construire des sous-descriptions fonctionnelles existe pour des cas complexes de neuropsychopathologies, avec ce résultat qu'il est possible d'obtenir des descriptions psychologiques au grain affiné qui sont co-extensionnelles avec les descriptions neuropsychologiques. D'autre part, les neurosciences font actuellement des progrès considérables dans le but de produire des explications réductives des causes de pathologies cognitives telles que la schizophrénie. Les conditions de co-extensionnalité et d'explication réductive sont donc remplies, ou du moins atteignables *en principe*. Face au dilemme de la réalisation multiple, notre stratégie de réduction fonctionnelle permet d'aborder ces progrès dans un esprit conservatif plutôt qu'éliminativiste : la qualité scientifique des descriptions psychologiques est assurée au travers de leurs relations systématiques d'abstraction avec les classifications de grain inférieur.

A ce point, il devrait être clair qu'il n'y a pas d'objection de principe à l'application de cette procédure à toute autre pathologie neuropsychologique, à savoir à toute pathologie dont les effets sont observables comportementalement. Il est ainsi en principe possible d'introduire un sous-concept fonctionnel pour chaque type de réalisateur neuropsychologique d'une propriété mentale : pour les patients souffrant du syndrome de déconnexion en contraste de sujet normaux, mais aussi pour les patients atteints du syndrome de négligence, de vision aveugle, etc. Il suffit ainsi simplement d'adapter de la manière suivante les descriptions fonctionnelles, en mettant à profit les outils de test que nous procure la neuropsychologie :

(F') $(\exists x)(\exists y)(\exists z)((x \text{ est causé par les conditions environnementales } c_1 \vee c_2 \vee \dots \vee c_n) \ \& \ (x \text{ tend à causer l'instanciation des propriétés mentales } y \vee z) \ \& \ (x \text{ tend à causer les comportements } b_1 \vee b_2 \vee \dots \vee b_n)) +$
 descriptions supplémentaires des causes actuelles + descriptions supplémentaires des autres propriétés mentales impliquées + descriptions supplémentaires du comportement causé par x .

Cet affinement des descriptions fonctionnelles est possible car toute spécificité neurologique pertinente du point de vue de la neuropsychologie induit des spécificités comportementales qu'il est possible de mettre à profit afin de construire une théorie générale de la production du comportement par le cerveau. En d'autres termes, une spécificité structurale neurologique modifie la manière dont l'information est traitée par le cerveau, ce qui implique certaines différences dans la manière dont l'effet comportemental est produit, ce qui n'est rien d'autre qu'une modification dans la manière dont les propriétés mentales produisent leur effet. Or, ceci s'exprime nécessairement comportementalement dans certaines circonstances.

Si la présente contribution s'est intéressée à la relation entre psychologie et neuropsychologie, il faut souligner que la stratégie discutée ici peut s'appliquer de

manière à distinguer fonctionnellement des réalisateurs non pathologiques d'une même propriété mentale, pour autant que les systèmes nerveux de ceux-ci ne soient pas totalement identiques d'un point de vue physique. En d'autre terme, il n'y a pas d'objection de principe à l'extension de la stratégie présentée ici au cas de la relation entre la psychologie et la neurobiologie. Les différences distinguant les réalisateurs requerront la finesse de description propre à la neurobiologie, par exemple en prenant en compte de très fines variations de connectivité entre deux individus jugés comme normaux du point de vue neuropsychologique. Puisque, comme souligné plutôt, l'existence de différences dans la composition physique implique des différences causales⁵⁵, nous devrions être capables de trouver des conditions critiques dans lesquelles les différenciations pertinentes du point de vue neurobiologique sont liées avec des particularités comportementales⁵⁶. Ainsi cette procédure doit être applicable jusqu'au niveau de la physique fondamentale. Bien entendu, les descriptions fonctionnelles prenant en compte ce type de différenciations seront complexes à tel point qu'il sera probablement impossible de les manier en pratique. Mais c'est là à nouveau un problème pratique et en tout cas pas une objection de principe allant à l'encontre de la possibilité théorique de réduire conservativement la psychologie et de quitter ainsi le dilemme posé par la réalisation multiple en philosophie des sciences.

IX. Conclusion

Au terme de ce survol, il se dégage une position générale et cohérente en philosophie des sciences articulée autour d'une conception fonctionnaliste des propriétés, que celles-ci appartiennent aux sciences spéciales ou à la physique. L'argument causal en faveur du réductionnisme ontologique se base sur les prémisses de survenance et de complétude du domaine des états physiques. Ces deux prémisses s'appliquant de la même manière à la caractérisation des rapports entre la physique et les différentes sciences spéciales, la stratégie de réduction fonctionnelle présentée ici relativement aux cas de la biologie et de la psychologie apparaît comme l'esquisse d'un cadre fonctionnaliste général permettant de rendre justice à l'hypothèse d'Oppenheim et de Putnam concernant l'unité des sciences sans toutefois conduire à l'élimination des sciences spéciales. Cette position se démarque ainsi clairement du débat actuel sur le réductionnisme, au sein duquel l'argument de la réalisation est le plus souvent interprété comme montrant l'incompatibilité entre le réalisme à l'égard des sciences spéciales et la réductibilité de ces dernières, donnant lieu au dilemme présenté dans la section III.

Pour peu que l'on accepte qu'il existe un lien systématique entre différences physiques et différences fonctionnelles, il est possible de quitter ce dilemme pour fonder

⁵⁵ Voir aussi la contribution de Michael Esfeld dans ce numéro.

⁵⁶ Pour un exemple reliant si l'on peut dire molécules et comportement, voir J. Bickle, *Philosophy and Neuroscience. A Ruthlessly Reductive Account*, Dordrecht, Kluwer, 2003.

une position réductionniste mais non-éliminativiste. La clef de voûte consiste donc en l'articulation de cette stratégie avec une théorie causale des propriétés de manière à établir un lien systématique entre différenciations physiques et fonctionnelles⁵⁷ : les spécificités physiques distinguant les différents réalisateurs d'une même propriété fonctionnelle donnent lieu à des spécificités causales susceptibles de se manifester de manière saillante dans certaines conditions environnementales. La position présentée ici prend la forme d'un cadre théorique général fonctionnaliste unifiant en philosophie des sciences. Les propriétés des sciences spéciales comme les propriétés physiques sont traitées fonctionnellement, ce qui permet de résorber le fossé donnant lieu au débat contemporain entre ces deux types de disciplines tout en reconnaissant une autonomie relative aux sciences spéciales. Leur qualité scientifique est fondée sur leur capacité de mettre en évidence des similarités causales entre différentes entités physiques, qui ne sauraient apparaître à l'échelle de la physique. En ce sens, les sciences spéciales sont indépendantes de la physique, sans que ne s'ensuive un conflit avec le réductionnisme.

⁵⁷ Voir encore la contribution de Michael Esfeld dans ce numéro, notamment à l'égard de la réalisation multiple quidditiste si l'on se situe dans un cadre métaphysique humien.