

RÉPHA
Revue Étudiante
de Philosophie Analytique

Numéro 2
Printemps 2010

REMERCIEMENTS

Nous voudrions exprimer notre reconnaissance aux membres du comité de lecture qui ont accepté de se consacrer sans relâche à cette entreprise et qui n'ont cessé de nous encourager. Premier et second numéros confondus, nous avons eu le plaisir de recevoir les conseils et relectures de:

Bruno Ambroise, Adrien Barton, Laure Blanc-Benon, Denis Bonnay, Christine Clavien, Fabrice Correia, Luc Faucher, Charles Girard, Jean-Baptiste Joinet, Xavier Kieft, Stéphane Lemaire, Stéphane Leyens, Alberto Masala, Anne Meylan, Ruwen Ogien, Fabrice Pataut, Jérôme Ravat, Xavier Sabatier, Christian Sachse, Fabrice Teroni, Hugo Viciana, et encore, tant d'autres personnes.

Que tous sachent combien l'intérêt qu'ils portent à Répha lui est essentiel.

Aussi, nous souhaiterions à nouveau rendre hommage à nos soutiens académiques (Daniel Andler, Mélika Ouelbani, Pascal Ludwig) qui ont vu naître cette revue.

Pour en savoir plus: www.repha.fr
66, rue Didot, 75014 PARIS

Association Loi du 1901, présidente: Marie Robert, vice-président: Arturs Logins, trésorier: Nicolas Liabeuf, directeur de la publication: Marie Robert, responsable de la rédaction: Arturs Logins.

Imprimeur: Poly Flash SARL & Lalande Digital Press,
15-17 rue Labrouste, 75015 Paris,
tél. : +33 (0)1.53.68.16.10,
fax : +33(0)1.68.16.27,
polyflash@mac.com

Dépôt à parution: 26 juillet 2010.
Le numéro ISSN 3 782871 107008.

ÉDITORIAL

Vous tenez entre les mains le second numéro de RÉPHA. La revue a pour vocation, – est-il nécessaire de le rappeler ? – de publier des articles de philosophie analytique. Son objectif à long terme est de s'insérer dans le paysage philosophique francophone. Sa fierté est de respecter minutieusement les normes académiques de la publication (méthode de relecture en double aveugle, rapporteurs compétents, comité éditorial). Sa force est d'être orientée vers les jeunes chercheurs, les étudiants avancés, les post-doctorants, qu'ils se présentent comme auteurs ou comme lecteurs.

Les entreprises éditoriales, on le sait, connaissent de nombreuses difficultés pouvant concerner tout autant l'élaboration intellectuelle de leur projet [puisqu'il faut que tous les contributeurs se mettent en accord textuel] que son élaboration matérielle [allant de l'épreuve parfois difficile de la communication électronique à la conception graphique de la maquette en passant par toutes autres formes d'aléas qui disparaîtront dès la publication]. Tout cela est encore plus vrai d'un jeune projet. Les méandres sont nombreux dans lesquels ces entreprises sinuent et contre lesquels elles doivent toujours lutter pour élargir leur horizon, faire connaître leurs aspirations, enfin et surtout, espérer s'implanter dans les esprits. Nous voudrions simplement réitérer notre volonté de permettre à tout auteur d'exprimer sa pensée dès lors qu'elle manifeste clarté et argumentation. Et le fait même que vous teniez en mains le second numéro est la preuve de notre détermination à surmonter les difficultés et d'arriver, même modestement, à *faire* de la philosophie. Nous sommes convaincus qu'une revue est le lieu par excellence du progrès de l'activité philosophique. Cette conviction est à son tour la raison qui motive l'existence de RÉPHA.

Nous souhaitons faire parler une certaine tradition d'origine anglo-saxonne. Mais la tradition n'est pas essentielle. Avant toutes choses, et en particulier, avant d'être anglo-saxonne, la langue source de la philosophie exposée dans RÉPHA est une langue dite « de recherche internationale », et même « langue scientifique de... ». Cette locution a son étrangeté et est critiquable quant au domaine qu'elle vise. D'une part, il est nécessaire qu'une langue visant la recherche manifeste une universalité car, autrement, les communications et les débats, si indispensables, seraient impossibles. C'est pourquoi il est requis d'une langue de recherche qu'elle *transcende* toute idiosyncrasie pour pouvoir énoncer des thèses, des arguments, des preuves, des lois, qui soient valables universellement. Mais, d'autre part, nous n'avons d'autre choix que de nous en remettre à une langue particulière : l'anglais, en l'état. Or, par définition, une langue comme l'anglais, vernaculaire, n'est pas universelle.

La volonté de RÉPHA est de montrer que la recherche, au moins en philosophie, ne tient pas à la singularité d'une langue naturelle quelconque, (anglaise, française, espagnole, italienne, allemande, etc.), mais plutôt à l'habileté des philosophes pour la manier à la seule fin de l'argumentation

pertinente et tranchante. Il semble évident que la langue universelle en philosophie n'est pas *une* langue singulière, mais *toute* langue respectant la rigueur qu'exige la recherche philosophique. En cela, RÉPHA donne la possibilité de se conformer à ce que Pascal Engel a baptisé le « troisième commandement du philosophe analytique gallique »¹.

Aussi, avant de présenter l'ensemble des articles que nous publions dans ce numéro, signalons une nouveauté qui ravive et surtout précise encore la nature de la revue. L'avenir verra sans aucun doute de nouvelles traductions prendre place à côté de celle, inaugurale, de l'article de Joshua Knobe, « Qu'est-ce que la philosophie expérimentale ? ». En faisant écho à l'injonction de Xavier Kieft, dans la présentation de ce même article qui appelle de ses vœux une philosophie expérimentale *à la française*, il faudrait complètement se tourner vers *une philosophie analytique à la française*. Les premiers pas en ce sens sont sans doute ceux qui vont vers la lisibilité et la visibilité. La visibilité s'incarne : SOPHA, RÉPHA, Igitur, Institut Jean Nicod, IHPST, Archives Poincaré, etc. La lisibilité revient à la clarté, et cette dernière étant notre fer de lance, force est de reconnaître qu'une « philosophie de la clarté », qui définirait son objet de concert et sans fausses notes, n'est jamais qu'une affaire sans cesse recommencée.

Comme à notre habitude, désormais, l'article introducteur revient à un chercheur qui présente une discussion pertinente en lien avec la pratique de la philosophie analytique. M. Gayon (Paris I, IHPST), spécialiste de la philosophie de la biologie, a accepté de clôturer un cycle de réflexions initié avec l'année de Darwin. Conjoint à l'article de Florent Franchette (doctorant, Paris I) proposant une perspective programmatique pour la computation physique - qui pourrait prendre une implémentation quantique -, RÉPHA fait ainsi place à la philosophie des sciences qui était jusqu'ici absente.

Valentine Reynaud (doctorante, Lyon III) défend l'idée selon laquelle la notion d'inné peut conserver un usage scientifique si elle comprise de façon non équivoque par la psychologie du développement, et plus précisément, si elle est comprise comme modalité privilégiée de l'explication développementale d'un trait.

Sarah Troubé (ENS Ulm), en interrogeant l'« épistémologie » propre aux délires, déplace l'argumentation habituellement empruntée par la philosophie cognitive et centre l'objet de cette dernière sur un cas qu'on considère traditionnellement comme « cas -limite ».

Enfin, la traduction par Raphaël Verchère et Luc-Étienne de Boyer des Roches (Lyon III) d'un article de J. Knobe sur la philosophie expérimentale, et replacée dans son cadre par Xavier Kieft (Paris IV), illustre la diversité des contributions qui auront droit de cité dans la revue.

Sur ce, il ne nous reste plus qu'à vous souhaiter une bonne lecture !

L'équipe RÉPHA

¹ « Tu n'écriras pas seulement en anglais, mais aussi en français, et pour tes collègues et étudiants, et tu ne citeras pas seulement des travaux en anglais dans tes écrits. » ("La philosophie analytique en France: un bilan institutionnel", in Cahiers de philosophie de l'université de Caen, Actes du 3ème colloque de la SOPHA, La normativité, PUC 2002.)

SOMMAIRE

VITALISME ET PHILOSOPHIE DE LA BIOLOGIE Jean GAYON	7
LA THÈSE DE TURING PHYSIQUE ET L'INFORMATIQUE QUANTIQUE Florent FRANCHETTE	19
FAUT-IL RENONCER À LA NOTION D'«INNÉ» ? Valentine REYNAUD	25
L'ÉTRANGETÉ DÉLIRANTE: LA PHILOSOPHIE À L'ÉPREUVE DE LA PSYCHOPATHOLOGIE CLINIQUE Sarah TROUBÉ	35
LES ÉPOUVANTAILS TRANSPARENTS N'EXISTENT PAS. NOTE POUR UNE PHILOSOPHIE EXPÉRIMENTALE À LA FRANÇAISE Xavier KIEFT	43
QU'EST-CE QUE LA PHILOSOPHIE EXPÉRIMENTALE ? Joshua KNOBE Traduction par Raphaël VERCHÈRE et Luc-Étienne DE BOYER	49

VITALISME ET PHILOSOPHIE DE LA BIOLOGIE

Jean GAYON

Université Paris 1-Panthéon Sorbonne

Institut d'histoire et de philosophie des sciences et des techniques
(UMR 8590 CNRS/Paris 1/ENS)

1. INTRODUCTION

Nous soulevons ici la question du sort fait au vitalisme dans la philosophie de la biologie contemporaine. Nous entendons 'philosophie de la biologie' au sens conventionnel que ce terme a pris dans la littérature internationale de philosophie des sciences depuis les années 1970. Ceci ne veut pas dire qu'il n'y ait pas eu de philosophie de la biologie au sens d'une réflexion philosophique sur la biologie antérieurement (sur ce point voir l'intéressante étude de Byron 2007), mais que dans les années 1970 le terme *philosophy of biology* (celui-là et pas un autre) est devenu le nom d'une sous-discipline au sein de la philosophie des sciences, avec ses sociétés, ses congrès, ses périodiques, au service d'une certaine manière pour les philosophes de parler de la biologie (Hull 1969, 2008 ; Gayon 2009). Bien que cela ne soit jamais dit explicitement, cette tradition contemporaine se reconnaît à quelques maximes méthodologiques caractéristiques : — distinction franche entre histoire de la biologie et philosophie de la biologie ; — parti pris corrélatif de traquer et résoudre des énigmes conceptuelles associées à des recherches contemporaines en biologie ; — conviction qu'il n'y a pas de frontière bien définie entre la philosophie des sciences et la science elle-même (la philosophie est plus 'conceptuelle' et la science est plus 'empirique', mais ce n'est qu'une question de degré) ; — valorisation de collaborations étroites entre philosophes et scientifiques pour résoudre les énigmes ; — conception résolument régionaliste de la philosophie des sciences ; — évitement des questions de philosophie pratique liées à la médecine, ou à d'autres champs de biologie appliquée. Le terme de 'philosophie de la biologie' fonctionne aujourd'hui comme une sorte de label. Bien que l'expression, créée par William Whewell en 1840 ait été sporadiquement utilisé depuis cette date, en concurrence avec des expressions telles que 'philosophie biologique', 'philosophie des sciences de la vie', 'épistémologie de la biologie', 'épistémologie biologique', 'épistémologie des sciences de la vie', il ne l'a été jusque dans les années 1970 que de manière générique, sans connoter une orientation intellectuelle spéciale. Il s'est répandu de manière fulgurante au début des 1970, d'abord en Amérique du Nord, Australie et Nouvelle-Zélande, puis, dans les années 1990 et surtout 2000 dans le reste du monde. Que ce terme-là plutôt qu'un autre se soit imposé a résulté d'une histoire largement contingente. Mais c'est celui-là qui s'est imposé, à la manière de ces marques de fabrique auxquelles la culture nord-américaine nous a habitués. Ceci ne va pas sans créer des frictions et des équivoques : certains philosophes qui

s'intéressent aux sciences de la vie en ignorant les canons non-écrits de cette école de pensée sont parfois surpris, voire outragés, de s'entendre dire qu'ils ne font pas de la 'philosophie de la biologie', et plaident pour une acception générique du terme, selon laquelle tout philosophe qui parle d'une manière ou d'une autre des sciences de la vie est un 'philosophe de la biologie'. Mais le collège invisible des réseaux institutionnels fonctionne, avec ses congrès, comités de lecture, collections éditoriales.

Qu'en est-il donc du vitalisme dans cette école de pensée ? Comme on le verra, c'est un terme qui est en est quasi totalement absent, ou plus exactement qui en est devenu absent. Il nous faudra comprendre cet état de choses, et pour cela remonter à la période où la philosophie néo-positiviste des sciences était à son acmé, dans les années 1950 et 1960. Dans cette période, la question du vitalisme a été discutée avec soin.

2. CONSTAT DE DISPARITION

Établissons d'abord le constat de disparition. J'utiliserai quelques données chiffrées. Nous ne visons ici ni à l'exhaustivité ni à une enquête bibliométrique digne de ce nom. Nous nous sommes appuyé sur des données aisément accessibles, qui ont le mérite de pointer clairement le phénomène.

J'ai d'abord réalisé une enquête sur le périodique *Biology and Philosophy*. Fondé en 1986, ce journal s'est vite imposé comme une revue internationale majeure, reconnue par les biologistes autant que par les philosophes de sciences, et représentant de manière paradigmatique l'esprit de la philosophie de la biologie. Sur vingt-trois années (soit près de cent numéros), aucun article n'a comporté le mot 'vitalisme' dans son titre. Seul un compte rendu de livre le contient (Spicker 1987). De deux choses l'une : ou bien aucun auteur n'a jugé opportun de proposer un article portant sur le vitalisme, ou bien le comité éditorial a écarté tous les articles de cette sorte. Pour trancher cette question, nous avons posé la question à Michael Ruse, qui l'a fondé et en a été le rédacteur en chef de 1986, date de la fondation du journal, à 2000, où il s'est retransché sur la position d'*emeritus editor*. Voici ce qu'il nous a répondu :

« Good question – I cannot remember any pieces on vitalism although cannot swear that nothing came in pushing the idea – certainly nothing from any professional philosopher as it were. However, don't forget that I did edit and publish the correspondence between John Greene and Theodosius Dobzhansky – although Dobzhansky was no vitalist (I am inclined to think that Julian Huxley got pretty close) he was an enthusiast for Teilhard de Chardin, who strikes me as vitalistic in respects. From a historical point of view I was happy to publish the exchange although I doubt that today it would be published in B and P. » (correspondance personnelle, 16 mai 2010).

Pour comprendre cette absence massive d'intérêt pour le vitalisme, j'ai effectué une recherche plus large avec *Google Scholar* sur plusieurs revues de philosophie des sciences et d'histoire des sciences. *Google scholar* comptabilise en effet les occurrences de mots dans le corps des textes dans ceux des volumes qui ont été numérisés. Voici le résultat brut de cette recherche, effectuée début juin 2010.

Nom de la revue	Nombre d'occurrences du mot 'vitalism'	Type de la revue
<i>Biology and philosophy</i>	50	Revue internationale de philosophie des sciences
<i>Philosophy of Science (PSA)</i>	15	
<i>British Journal for the Philosophy of Science</i>	49	
<i>British Journal for the History of Science</i>	276	Revue internationale d'histoire des sciences
<i>Journal of History of Biology</i>	123	

Ces données ne permettent pas de faire des pourcentages, en rapportant le nombre d'occurrences au nombre des articles, mais elles suggèrent que le mot 'vitalisme' est plus important pour les historiens de la biologie que pour les philosophes de la biologie. La comparaison BJPS/BJHS est frappante à cet égard. Ces deux journaux sont issus d'une bipartition à partir d'un même périodique, les deux séries ayant aujourd'hui un poids similaire. Or il y a presque six fois plus d'occurrences du terme 'vitalisme' dans la série historique qu'il n'y en a dans la série philosophique (276/49).

Voici une autre comparaison instructive, réalisée au sein même de la série entière des numéros de *Philosophy and Biology* :

Terme	<i>Vitalism</i>	<i>Reductionnism</i>	<i>Mechanism</i>
Nombre d'occurrences du terme dans le texte	50	250	720

Ce tableau est plus surprenant que les précédents. Le poids plus important du mot 'réductionnisme', relativement à celui de 'vitalisme', se comprend assez bien si l'on songe à l'importance du sujet du réductionnisme dans la littérature contemporaine de philosophie des sciences. Le réductionnisme est, de fait, assez fréquemment présent dans les titres et surtout les mots-clés des articles, ce qui veut dire que c'est un sujet de réflexion considéré comme légitime.

Les 720 occurrences du mot 'mécanisme' sont plus

surprenantes. Nous avons effectué des sondages qualitatifs. Ils montrent que seule une très petite fraction des articles utilisant le mot 'mécanisme' ont pour objet d'élucider la notion de ce qu'est *un* mécanisme [e.g. revue de Craver 2007 par Levy 2009]. Ce questionnement est de surcroît récent, et typique d'auteurs insatisfaits par les analyses des théories biologiques fondées sur le concept de loi. En fait, la majorité des articles utilisant le mot 'mécanisme' prennent ce mot au sens d'une vision particulière du monde vivant (*le* mécanisme, c'est-à-dire ce qu'on devrait appeler en fait le 'mécanicisme'). Ceci ne laisse pas d'étonner pour un journal qui a la réputation de s'intéresser de manière trop exclusive à l'évolution, en négligeant la biologie du fonctionnement. Les auteurs qui emploient le mot 'mécanisme' ne se disent pas pour autant 'mécanistes'.

Ces chiffres, rapprochés des précédents, suggèrent donc deux inférences. (1) les auteurs de *B&P* préfèrent réfléchir sur le 'réductionnisme' que sur le 'vitalisme' ; le premier mot est non seulement davantage utilisé, mais aussi davantage thématiqué. (2) L'usage relativement abondant du mot 'mécanisme' (14 fois plus que 'vitalisme') n'est pas lié à une réflexion moderne sur les mérites respectifs du vitalisme et du mécanisme. Quoique ce dernier soit assez peu interrogé pour lui-même, le terme semble faire partie du lexique ordinaire et acceptable, tandis que celui de 'vitalisme' a massivement régressé.

Voici une dernière donnée quantitative, qui vise à illustrer un éventuel biais culturel et linguistique. Par curiosité, j'ai relevé début juin 2009 le nombre de sites auxquels renvoie *Google* (tout court) pour les versions anglaises et françaises de 'vitalisme' et 'réductionnisme' :

Ces chiffres totalisent tous les usages du mot dans n'importe

	Français	Anglais
Vitalism(e)	388 000	182 000
Réductionnism(e)	44 000	607 000

quel contexte sur internet. La plupart du temps, le contexte n'est pas celui de la philosophie ou de l'histoire des sciences. Ils montrent un biais culturel et linguistique massif. Les francophones utilisent huit fois plus le mot 'vitalisme' que le mot 'réductionnisme', tandis que les anglophones utilisent 3,5 fois plus le mot 'reductionnism' que le mot 'vitalism'.

J'ai enfin réalisé un test sur quelques livres collectifs de langue anglaise, largement diffusés, ayant pour objet explicite la philosophie de la biologie, et comportant un index. Les index donnent les résultats suivants :

Éditeur scientifique	Titre	Année	Occurrences du mot 'vitalisme' dans l'index	Remarques
M. Ruse	<i>What Philosophy of Biology Is. Essays Dedicated to David Hull</i>	1989	2	p. 6 : le mot renvoie à une page consacré au problème de la réduction mais n'apparaît pas dans le texte. Apparemment, le 'vitalisme' serait l'opposé de la « réduction ontologique » p. 55. Usage accidentel, au sens de « explication par les miracles »
D. Hull & M. Ruse	<i>The Philosophy of Biology</i>	1998	0	
M. Ruse & D. Hull	<i>Oxford Handbook of Philosophy of Biology</i>	2008	0	

Un examen des collections de philosophie de la biologie existant en langue anglaise (par exemple la prestigieuse collection de Cambridge University Press) conduirait au même constat. La plupart du temps, le mot 'vitalisme' est totalement absent. Lorsqu'il est utilisé, c'est le plus souvent sous l'angle du problème de la réductibilité de la biologie à la physique ; le mot est alors une sorte de concession à un mode d'expression archaïque, le vrai problème étant celui de la réduction. La conclusion de notre sondage quantitatif est donc claire. Le concept de vitalisme ont disparu de la littérature contemporaine de philosophie de la biologie. Le mot n'apparaît que rarement, de manière non problématisée. C'est un terme muséal, réservé aux historiens.

3. QUE S'EST-IL PASSÉ ?

Pour comprendre ce qu'il s'est passé, il vaut la peine de revenir sur quelques textes majeurs de la philosophie des sciences néo-positiviste à son apogée, c'est-à-dire autour des années 1960. Nous nous contenterons ici d'examiner les deux ouvrages les plus célèbres à cet égard, deux ouvrages qui depuis un demi-siècle ont plus que tout autre contribué à former des générations de philosophes des sciences : *The Structure of Science* d'Ernest Nagel (1961), et les *Éléments d'épistémologie* de Carl Hempel (originellement publié en anglais sous le titre *Philosophy of Natural Science* en 1965). Ces ouvrages demeurent aujourd'hui des supports majeurs de l'enseignement universitaire de la philosophie des sciences. Les philosophes de la biologie des années 1970-1980 ont construit leur domaine à la fois contre ce genre de philosophie des sciences (au

nom d'une philosophie des sciences plus attentive à des controverses scientifiques particulières, et en quelque sorte moins surplombante], mais aussi en partie dans leur langage. Outre leur représentativité intrinsèque, j'ai choisi ces deux ouvrages parce qu'ils contiennent des développements explicites et importants au vitalisme. Ils permettent de comprendre pourquoi, après eux, le mot et le concept de vitalisme ont été délaissés. Avant de réaliser la présente étude, nous n'avions guère prêté attention à ces développements, tant il nous semblait évident que des philosophes néo-positivistes ne pouvaient pas avoir de sympathie pour le vitalisme, et que le ton ne pouvait donc être que critique, voire condescendant. À relire aujourd'hui ces textes, nous n'avons pas moins de doutes sur cette hostilité des deux philosophes, mais il nous paraît clair que c'était pour un l'objet d'un combat intellectuel important, qu'il convenait de gagner avec les meilleurs arguments possibles. Ces arguments sont tous saufs convenus ; ni Hempel ni Nagel, nous semble-t-il, ne considéraient *a priori* la partie gagnée ; tous deux estimaient essentiels d'affronter le problème (à la différence des philosophes de la biologie d'aujourd'hui, qui s'en désintéressent totalement, laissant la question aux historiens, ou à des traditions philosophiques, plutôt littéraires ou continentales, avec lesquelles ils estiment n'avoir pas grand chose à partager).

Les réflexions de Hempel et de Nagel sur le vitalisme présentent beaucoup de points communs, mais on observe aussi des nuances significatives. Hempel a exposé avec élégance les raisons qu'il y a de ne pas laisser place au vitalisme dans les recherches scientifiques. Son argumentaire exclusivement positiviste, et donc prudent, a son exact pendant chez Nagel, qui a cependant été plus loin en faisant place à des arguments théoriques en faveur du mécanisme. Comme on l'a parfois fait remarquer, Nagel ne peut être classé sans nuance comme un 'philosophe des sciences néo-positiviste' ; son œuvre est marquée par une inflexion naturaliste (Nagel 1954, Rosenberg 1997). Aussi examinerons-nous successivement les argumentaires anti-vitalistes de ces deux biologistes anti-vitalistes, la version purement positiviste de Hempel d'abord, puis la version augmentée de Nagel.

3.1. CRITIQUE DU VITALISME : VERSION STRICTEMENT POSITIVISTE

Prenons d'abord acte de la caractérisation du vitalisme et du mécanisme que partagent Hempel et Nagel.

– Le vitalisme, ou plus exactement le 'néo-vitalisme' pose qu'on ne peut expliquer certaines caractéristiques des systèmes vivants en recourant seulement à des principes physiques et chimiques. Pour rendre compte de ces caractéristiques, « il faut les rapporter à de nouveaux facteurs, d'une espèce inconnue dans les sciences physiques, notamment les entéléchies ou forces vitales' (Hempel [1966] 1972, p. 157).

– Le mécanisme pose quant à lui que « les organismes

vivants ne sont rien d'autre que des systèmes physico-chimiques très complexes' (Hempel [1966] 1972, p. 157), et que, par conséquent « tous les processus biologiques "peuvent être expliqués en termes physico-chimiques d'une manière non équivoque" » (Nagel 1961, p. 412. Nagel cite J. Loeb, *The Mechanistic Conception of Life*, 1912).

L'argumentation anti-vitaliste de Hempel tient en deux volets, développés en deux chapitres distincts de son ouvrage.

En premier lieu, Hempel dénonce l'irresponsabilité des hypothèses vitalistes du point de vue de l'explication. Les hypothèses vitalistes portent sur des entités et des processus inaccessibles à l'observation directe, comme c'est souvent le cas dans les sciences physiques. En mécanique, par exemple, la masse ou la gravitation ne sont pas non plus des observables. Mais, alors que la mécanique newtonienne permet de dériver des conséquences testables, les théories vitalistes ne disent pas dans quelles conditions les entéléchies ou le principal vital qu'elles invoquent entrèrent en action. Ces théories ne disent jamais ce qui devrait se produire ; elles se contentent de dire après coup que le facteur invisible s'est manifesté. Les comparaisons traditionnelles entre les hypothèses vitalistes et les hypothèses newtoniennes ou les hypothèses physiques en général sont donc non fondées (Hempel [1966] 1972, p. 111-112).

Le second volet de l'argumentation hempélienne vise à clarifier l'opposition entre mécanisme et vitalisme. Si le mécanisme a raison, il faut admettre les deux assertions suivantes (Hempel [1966] 1972, p. 158-164). (1) Toutes les propriétés des êtres vivants peuvent être complètement *décrites* grâce aux concepts de la physique et de la chimie. De manière plus technique, ceci implique que les termes des théories biologiques soient traductibles en termes physiques et chimiques : c'est la thèse de la réductibilité des termes théoriques. (2) Tous les aspects du comportement des êtres vivants peuvent être *expliqués* au moyen de lois et de théories physico-chimiques ? C'est la thèse de la réductibilité des lois de la biologie à celles de la physique et de la chimie. On reconnaît dans ces deux concepts de réductibilité la distinction faite par Nagel entre 'connectabilité' des termes et 'dérivabilité' des lois (Nagel 1961, p. 433-434).

Prises ensemble, ces deux thèses composent la thèse globale de la possibilité d'une réduction de la biologie à la physique et à la chimie (réductibilité, qui ne signifie aucunement que la réduction existe en fait). À partir de là, Hempel propose de régler la question de l'antagonisme entre vitalisme et mécanisme en développant trois idées.

En premier lieu, la véritable opposition n'est pas entre mécanisme et vitalisme, mais entre la thèse de la réductibilité de la biologie à la physique et à la chimie, et celle de son irréductibilité, thèse également connue sous le nom de la « thèse de l'autonomie de la biologie » (autonomie des concepts et des principes). « La thèse mécaniste affirme qu'il est possible de réduire la biologie à la physique

et à la chimie. La négation de cette thèse est parfois appelée thèse de *l'autonomie de la biologie* ou — ce qui est mieux — de l'autonomie des concepts et des principes de la biologie. Le néo-vitalisme affirme donc l'autonomie de la biologie et ajoute à cette assertion sa doctrine des forces vitales » [Hempel [1966] 1972, p. 159]. Relevons l'argument subtil et redoutable ici mis en place par Hempel : le vitalisme, c'est la thèse de l'autonomie de la biologie (question empiriquement ouverte, comme celle de la réductibilité), *plus* la doctrine selon laquelle il existe des facteurs relevant d'un autre ordre que les facteurs physico-chimiques (thèse irresponsable du point de vue de la recherche empirique, cf *supra*).

En second lieu, à la suite de Nagel, Dans un cas comme dans l'autre, Hempel soutient que la question de la réduction des termes et celle de la réduction des lois ne peuvent être tranchées par des raisonnements *a priori*. « La question de savoir si un terme biologique est 'définissable' au moyen de termes physiques et chimiques seuls ne peut être résolue en contemplant la signification de ce terme, ni par quelque autre procédure non empirique que ce soit » (Hempel [1966] 1972, p. 163). « D'une manière générale les lois biologiques sont explicables au moyen de lois physico-chimiques dans l'exacte mesure où l'on peut établir des lois de connexion appropriées. Et ce point, à son tour, ne peut se décider en recourant à des raisonnements *a priori* ; la réponse ne peut être trouvée que par les recherches biologiques et biophysiques » (Hempel [1966] 1972, p. 164).

Enfin Hempel soutient que le mécanisme bien compris est et ne peut être qu'une « maxime heuristique », qui « enjoint à l'homme de science de persister dans sa recherche de théories physico-chimiques fondamentales des phénomènes biologiques » Dans un cas comme dans l'autre, Hempel et Nagel disent que ces deux questions de définition des termes et de dérivabilité logique ne peuvent être tranchées par des raisonnements *a priori*. « La question de savoir si un terme biologique est 'définissable' au moyen de termes physiques et chimiques seuls ne peut être résolue en contemplant la signification de ce terme, ni par quelque autre procédure non empirique que ce soit » (Hempel [1966] 1972, p. 163). « D'une manière générale les lois biologiques sont explicables au moyen de lois physico-chimiques dans l'exacte mesure où l'on peut établir des lois de connexion appropriées. Et ce point, à son tour, ne peut se décider en recourant à des raisonnements *a priori* ; la réponse ne peut être trouvée que par les recherches biologiques et biophysiques » (Hempel [1966] 1972, p. 165).

3.2. CRITIQUE DU VITALISME : VERSION NAGÉLIENNE

Toutes les thèses de Hempel qu'on vient d'évoquer se retrouvent chez Nagel, et lui ont d'ailleurs été en partie empruntées. Il est difficile au demeurant d'établir ce qui revient à l'un et à l'autre auteur sur le sujet, tant ils ont interagi. On doit cependant noter que

Nagel est plus radical. Il ne se contente pas de dire que la question de la réductibilité est empiriquement ouverte. Il critique âprement les arguments traditionnels des partisans de l'autonomie de la biologie, tout particulièrement ceux avancés par les tenants du caractère 'organismique' de la pensée biologique, forme contemporaine familière du vitalisme. Il s'efforce en particulier de réfuter l'argument tiré du caractère hiérarchique de l'organisation des êtres vivants, et celui consistant à dire qu'ils forment des 'tout' ayant des propriétés non déductibles de leur composition (Nagel 1961, pp. 435-444). Selon Nagel, tout autant que les forces vitales, des termes comme celui de 'totalité indivisible', et des formules comme 'Le tout est plus grand que la somme des parties', ne sont que des cache-misère de notre manque de théories et de descriptions adéquates des phénomènes biologiques. Notre ignorance est grande, admet Nagel, mais aucun argument convaincant ne permet d'écartier *a priori* la possibilité de réduire les théories biologiques aux théories physico-chimiques.

4. LA PHILOSOPHIE DE LA BIOLOGIE CONTEMPORAINE (APRÈS 1970) ET LE VITALISME

Au regard des ouvrages néo-positivistes qu'on vient de commenter, l'attitude de la génération suivante se comprend bien. La 'philosophie de la biologie' des années 1970 est née d'une révolte contre l'approche de la biologie favorisée par la philosophie néo-positiviste des sciences. Le principal reproche qui lui a été fait, et en fait le seul reproche partagé par tous les 'philosophes de la biologie' nouvelle manière, a été d'avoir forcé les sciences particulières à rentrer dans le cadre d'une philosophie générale des sciences, et d'avoir ainsi négligé les continus particuliers des sciences empiriques. Corrélativement, on a reproché à des auteurs tels que Hempel et Nagel d'avoir un peu vite préjugé de l'unité méthodologique de la science et de son unité théorique.

Une fois cela dit, il est aisé de voir que les philosophes de la biologie se sont moulés dans le cadre de discussion mis en place par ces auteurs. En premier lieu, l'antinomie entre les vision mécaniste et vitaliste des phénomènes vitaux a été éliminée au profit de discussions passionnées sur la réductibilité et l'autonomie théorique de la biologie. Comme l'a bien noté Alexander Rosenberg, ce débat a été, et demeure constitutif de la 'philosophie de la biologie' (Rosenberg 1985). Une partie considérable de la littérature a porté sur des questions de réduction : les concepts de la génétique (gène, génotype, phénotype, mutation...) sont-ils réductibles à ceux de la biologie moléculaire ? Le concept de *fitness* est-il un concept physique ? Un grand nombre d'auteurs ont plaidé, dans ces cas particuliers, et dans d'autres, en faveur de l'autonomie, mais pas tous. C'est donc bien le cadre héméliano-nagélien qui a déterminé la structure des débats : les questions de réduction sont devenues des

questions relatives aux descriptions et théories biologiques disponibles.

La théorie de l'évolution a simultanément monopolisé une partie prépondérante des débats. Elle a, pour une majorité d'auteurs, apporté un argument majeur en faveur de l'autonomie des sciences de la vie. Du point de vue de l'évolution, en effet, la plupart sinon toutes les généralisations biologiques apparaissent comme relatives à une histoire unique et contingente. La notion de 'loi biologique' s'en est trouvée ébranlée, ce qui a pu donner l'impression que le cadre néo-positiviste était obsolète. En réalité, là encore, c'est le précepte hémphélien qui a opéré : les questions de réduction, plutôt que d'être abordées abstraitement, ont été rapportées aux théories biologiques disponibles, et en particulier à la plus englobante d'entre elles, la théorie de l'évolution. Bien loin d'avoir favorisé le retour du vitalisme, la discussion sur le caractère contingent des généralisations biologiques l'a rendu plus improbable. Comme l'a dit avec humour le philosophe australien Smart, la biologie est avec les sciences physico-chimiques dans un rapport comparable à celui que l'électronique entretient avec elles. L'électronique, c'est la physique et la chimie plus des réseaux de câblage ; la biologie, c'est la physique et la chimie plus l'histoire naturelle (Smart 1963).

Nous comprenons mieux maintenant mieux maintenant les données statistiques données au début de la présente communication. Le vitalisme a disparu comme terme et surtout comme concept dans la littérature de la philosophie de la biologie car on a préféré se concentrer sur la question de la 'réductibilité ou autonomie de la biologie' plutôt que sur l'alternative 'mécanisme vs. Vitalisme'.

5. CONCLUSION : REGARD SUR L'AUTRE RIVE DE L'ATLANTIQUE

Qu'il nous soit permis pour conclure de jeter un regard sur ce qui se passait de l'autre côté de l'Atlantique lorsque Hempel et Nagel déclaraient la stérilité du vitalisme et la supériorité heuristique des approches d'esprit réductionniste. Dans le même temps, Georges Canguilhem, dans sa fameuse communication de 1946-1947 sur le vitalisme, écrivait qu'« on doit attendre peu d'une biologie fascinée par le prestige des sciences physico-chimiques, réduite ou se réduisant au rôle de satellite de ces sciences » (Canguilhem [1952] 1975, p. 83). Ce que voulait dire le philosophe français, c'est que toutes les grandes découvertes biologiques ont procédé d'une attention à ce que les phénomènes de la vie ont de spécifique, sans chercher à les réduire *a priori* à un corps de doctrines physico-chimiques. De là le 'vitalisme' singulier de Canguilhem, vitalisme élargi ou, comme l'a qualifié François Delaporte, un 'vitalisme rationaliste' (Delaporte 1994). Tout comme le parti-pris mécaniste de

Hempel et de Nagel, le vitalisme de Canguilhem est un parti-pris heuristique plutôt que métaphysique. Cette pensée de Canguilhem, qui est à certains égards en résonance avec l'esprit de la philosophie de la biologie contemporaine, prompte à affirmer la fécondité d'une philosophie des sciences 'régionaliste', est séduisante, mais il nous semble qu'elle a égaré des générations de lecteurs, pour deux raisons.

Premièrement, la question de savoir si le vitalisme a été en pratique scientifiquement plus fécond que le mécanisme ne va pas de soi. Contrairement à ce qu'a dit Canguilhem, et contrairement aussi à l'assertion exactement contraire de Hempel et de Nagel, nous inclinons à penser que le vitalisme comme le mécanisme ont tantôt été heuristiquement féconds, tantôt des obstacles au développement de la science empirique.

En second lieu, même si Canguilhem avait raison du point de vue de la dynamique de la découverte scientifique, cela ne trancherait aucunement la question de savoir si les systèmes vivants ne sont rien d'autre que des systèmes physiques complexes, et si leurs propriétés sont descriptibles et explicables de manière exhaustivement physique. Les préférences cognitives des savants sont une chose ; la cohérence de leur ontologie en est une autre.

Quoi qu'il en soit, le sens que Georges Canguilhem a donné au mot 'vitalisme' place la philosophie de la biologie sous un jour familier. Pour Canguilhem, en effet, le vitalisme bien compris qualifie l'attitude d'une biologie attentive à la spécificité de son objet. Or la philosophie de la biologie telle qu'elle s'est développée depuis les années 1970 est précisément une école de pensée fondée sur le précepte selon lequel la biologie requiert une analyse philosophique adaptée à la spécificité des concepts, théories, et méthodes de la biologie. En outre, la majorité des philosophes de la biologie (quoique certainement pas tous) s'est réclamée de la thèse de l'autonomie des sciences de la vie, et s'est massivement appuyée sur des biologistes (tout particulièrement des biologistes de l'évolution, tels que Mayr, Dobzhansky, Ayala) qui ont plaidé dans le même sens. En ce sens, on pourrait dire que ce sont des 'vitalistes' *sensu* Canguilhem. Toutefois c'est là une acrobatie de langage qui n'a d'autre intérêt que d'ouvrir éventuellement un dialogue.

REMERCIEMENTS

Nous remercions Michael Ruse pour les informations précieuses qu'il nous a communiquées.

RÉFÉRENCES

- BYRON, Jason M. (2007). Whence philosophy of biology, *British Journal for the Philosophy of Science*, 58 : 409-422.
CANGUILHEM, Georges [1952] 1975. Aspects du vitalisme, in

La connaissance de la vie, Paris, Vrin.

LEVY, Arnon (2009). Review of Carl Craver Craver, *Explaining what? Review of explaining the brain: mechanisms and the mosaic unity of neuroscience*, *Biology and Philosophy*, 2: 137-145.

DELAPORTE, François ed. (1994). *A Vital Rationalist. Selected Writings from Georges Canguilhem*, New York, Zone Books.

HULL, David (1969). What philosophy of biology is not, *Journal of the History of Biology*, 2: 241-268.

HULL, David, & RUSE, Michael, eds. (1998). *The Philosophy of Biology*, Oxford, Oxford University Press, coll; "Oxford Readings in Philosophy".

HULL, David (2008). The history of philosophy of biology, in M. Ruse (ed.), *The Oxford Handbook of Philosophy of Biology*, Oxford and New York, Oxford University Press, pp. 11-33.

GAYON, Jean (2009). Philosophy of Biology : An Historico-Critical Characterization, in *French Studies in Philosophy of Science. Contemporary Research in France*, vol. 276 de Boston Studies in the Philosophy of Science, A. Brenner & J. Gayon (eds.), Springer, pp. 201-212.

HEMPEL, Carl (1972). *Éléments d'épistémologie*, Paris, Armand Colin. (Originellement publié sous le titre *Philosophy of Natural Science*, Englewood Cliffs, NJ, Prentice Hall, 1966).

NAGEL, Ernest (1954). *Sovereign Reason and Other Studies in the Philosophy of Science*, Glencoe (Ill.), The Free Press.

NAGEL, Ernest (1961). *The Structure of Science. Problems in the Logic of Scientific Explanation*, London, Routledge & Kegan Paul.

ROSENBERG, Alexander (1985). *The Structure of Biological Science*, Cambridge University Press.

ROSENBERG, Alexander (1997). « A field guide to recent species of naturalism », *The British Journal for the Philosophy of Science*, 47(1) :1-29

RUSE, Michael, ed. (1989). *What the Philosophy of Biology is: Essays Dedicated to David Hull*, Dordrecht, Kluwer Academic Publishers.

RUSE, Michael, ed. (2008). *The Oxford Handbook of Philosophy of Biology*, Oxford, Oxford university Press.

SMART, James Jamieson Carswell (1963). *Philosophy and Scientific Realism*, Londres, Routledge and Kegan Paul.

WHEWELL, William (1840). *Philosophy of the Inductive Sciences*, London, John W. Parker, 2 vols., 1840.

SPICKER, Stuart F. (1987). Biochemical reductionism or obscurantist vitalism. A new passage between Scylla and Charybda, *Biology and Philosophy*, 2 : 509-515.

LA THÈSE DE TURING PHYSIQUE ET L'INFORMATIQUE QUANTIQUE

Florent FRANCHETTE
(M2 Paris I)

RÉSUMÉ

Selon la thèse de Turing, tout ce qui est calculable par une procédure effective (notion intuitive) est calculable par une Machine de Turing (notion formelle). La thèse de Turing physique affirme quant à elle que tout ce qui peut être calculé par un système matériel est calculable par une Machine de Turing. Etant donnée sa nature d'énoncé universel, il ne suffit pas pour prouver cette thèse, de montrer que tout ce qui est calculable par nos systèmes matériels actuels est aussi calculable par une Machine de Turing. En revanche, la thèse de Turing physique serait falsifiée si l'on proposait un mécanisme de calcul qui résoudrait un problème qui ne serait pas décidable par une Machine de Turing. Nous essaierons de répondre dans cet article à la question suivante : pourquoi les récents résultats en informatique quantique apportent de nouvelles perspectives quant à la validité de la thèse de Turing physique ?

1. LA THÈSE DE TURING PHYSIQUE

Une première tentative afin de caractériser ce qu'est une procédure effective est la suivante. Une procédure M , exécutée dans le but d'obtenir un résultat, est dite « effective » lorsque :

1. M est définie à partir d'un nombre fini d'instructions;
2. M produit le résultat désiré en un nombre fini d'étapes;
3. M peut (en pratique ou en principe) être exécutée par un être humain sans l'aide d'aucune machine;
4. M ne demande aucune perspicacité ni ingéniosité de la part de l'être humain qui l'exécute.

Toutefois, notre définition n'est pas suffisante pour caractériser le fait qu'une fonction est calculable par une procédure effective. Un essai afin de formaliser un prédicat dont la signification informelle est celle de « peut être calculé au moyen d'une procédure effective » a été présenté par Alan Turing, qui proposa en 1936, le prédicat formel de « calculé par Machine de Turing ». Dans le but de comprendre en quoi une Machine de Turing est un modèle mathématique représentant une procédure effective exécutée par un être humain, donnons en une définition en citant Turing :

« Un homme en train de calculer la valeur d'un nombre peut être comparé à *une machine* susceptible de se trouver dans un nombre fini d'états. La machine est alimentée avec une *bande* [analogue au papier qu'utilise l'homme] et divisée en cases dans chacune desquelles peut être inscrit un *symbole*. À chaque instant, la liste des comportements possibles de la machine est entièrement déterminée par sa configuration, caractérisée par l'ensemble de ses états et le symbole inspecté. C'est cette configuration qui détermine l'évolution possible de la machine (...) Ce que j'affirme, c'est que ces opérations englobent toutes celles qui peuvent être utilisées pour calculer la valeur d'un nombre (...) Une séquence est dite *calculable* s'il existe une Machine de Turing qui la calcule » (Turing, 1936)

À partir de l'examen de la Machine de Turing, la théorie de la calculabilité, qui est une partie de la logique mathématique, a pu produire une série de résultats définissant les limites de ce qui est calculable par une procédure effective.

En particulier, Turing défendait l'affirmation suivante, nommée thèse de Turing ou thèse **M** : **Une fonction est calculable par une procédure effective si et seulement si elle peut être calculée par une Machine de Turing.** Remarquons le fait que cette thèse n'est pas un théorème car elle lie une notion informelle (celle de procédure effective) à une notion mathématique (celle de calculable par une Machine de Turing). Il est par conséquent impossible de démontrer que toute procédure effective est calculable par une Machine de Turing. En revanche, si la thèse de Turing est correcte, certaines fonctions ne sont pas calculables par une procédure effective car il peut être démontré qu'il n'existe pas de Machine de Turing pouvant les calculer¹.

Cependant, la thèse **M**, à partir de la définition mathématique de la Machine de Turing, détermine uniquement les fonctions pouvant être calculées par des moyens purement mathématiques. Par conséquent, cette dernière n'affirme rien concernant ce qui est calculable à l'aide de moyens extrinsèques aux mathématiques. Ainsi, que se passerait-il si l'on étendait la notion de procédure effective à la physique? Autrement dit, la thèse **M** définit-elle les limites de ce qui est calculable par une procédure effective indépendamment du domaine ou l'on se place? Une première étape afin de répondre à ces questions est d'étudier une version « plus physique » de la thèse de Turing qui concerne la notion de tâche réalisable en temps fini par un système matériel.

La version physique de la thèse de Turing ou thèse **P** s'énonce alors comme ceci : **une fonction est calculable par un système matériel si et seulement si elle peut être calculée par une Machine de Turing.** Cependant, contrairement à la thèse de Turing **M** qui

¹ Par exemple, puisque le problème de déterminer si une formule de la logique du premier ordre est démontrable ou non ne peut être résolu par une Machine de Turing, ce problème ne peut pas être résolu par une procédure effective.

semble être acceptée par la majorité des scientifiques, sa version physique est sujette à des divergences d'opinions quant à son exactitude. En effet, plusieurs scientifiques ont proposé différents mécanismes de calcul qui pourraient remettre en cause la thèse **P** s'ils venaient un jour à exister. Toutefois, cette version physique admet deux interprétations suivant que la phrase « pouvant être calculée par un système matériel » est prise au sens étroit ou au sens large. Au sens étroit, cette phrase doit être interprétée comme « pouvant être calculée par une machine conforme aux lois physiques du monde actuel » tandis qu'au sens large, l'interprétation fait abstraction de la contrainte physique. En interprétant cette thèse dans son sens étroit, la remettre en cause consisterait à proposer un mécanisme de calcul conforme aux lois physiques et pouvant calculer au moins une fonction non calculable par une Machine de Turing. Par conséquent, le calcul ne doit plus être considéré comme un sujet purement mathématique car ce doit être en dernier lieu à la physique de dire si la thèse **P** est juste ou fausse.

Dans ce but, une des propositions les plus prometteuses concernant la création d'une « hypermachine », c'est-à-dire d'un mécanisme pouvant falsifier la thèse **P** et dont la possible création n'est pas contradictoire avec les lois de la physique, provient des modèles de calcul conçus en mécanique quantique à partir des années 1980.

2. LES NOUVELLES PERSPECTIVES DE L'INFORMATIQUE QUANTIQUE

La possibilité d'appliquer la mécanique quantique à l'informatique découle tout d'abord de la confiance que nous accordons à cette théorie. En effet, si nous croyons que la mécanique quantique est la théorie fondamentale d'où dérive toutes les propriétés des objets physiques, alors cette dernière devrait être la base de la description de n'importe quel ordinateur. De plus, la miniaturisation des composants micro-électriques devrait bientôt atteindre l'ordre de 10 nanomètres (10^{-9} m) faisant de la mécanique quantique la théorie la plus pertinente pour créer certains composants informatiques.

Au cours des années 1980, la notion de procédure effective telle qu'elle est définie à travers la thèse de Turing physique, fut soumise à la question de savoir si elle pouvait être étendue aux principes quantiques. D'une part, les efforts initiaux menés par David Deutsch semblent considérer que les notions de calculabilité quantique et de calculabilité mathématique sont identiques. Cependant, des indications récentes établis par Tien Kieu portent à croire le contraire. Dans un premier temps, le modèle « standard » de la computation défendu par Deutsch sera présenté. Puis dans un second temps, la position controversée de Kieu soutenant que les

ordinateurs quantiques peuvent falsifier la thèse de Turing physique sera exposée.

D'après le modèle standard de la computation quantique, qui est une généralisation de la computation classique, l'unité fondamentale d'un ordinateur quantique est le *bit* quantique ou *qubit*, qui est la généralisation d'un *bit* classique. Mais alors qu'un *bit* classique peut prendre l'une ou l'autre des valeurs possibles servant à coder l'information (0 ou 1 par exemple), un *qubit* a la possibilité d'être dans un état superposé de ces deux valeurs. Toutefois, lorsqu'une mesure est effectuée, cette superposition est détruite, ce qui a pour conséquence de révéler une des deux valeurs classiques prises par le *qubit*.

Nous pouvons distinguer trois étapes principales dans la computation d'un ordinateur quantique : la préparation de l'*input*, le calcul et la mesure de l'*output*. Premièrement, la préparation de l'*input* et la mesure de l'*output* doivent être exécutées de telle façon à ne pas perturber les autres qubits qui ne seront pas directement mesurés. La seconde étape, correspondant au calcul proprement dit, est l'étape la plus difficile à réaliser physiquement. En principe, le calcul est causée par l'évolution d'opérations réversibles sur les *qubits*, mais ces derniers doivent être correctement isolés de toute perturbation afin d'éviter le plus possible les effets de décohérence dus à l'environnement.

D'autre part, la supériorité des ordinateurs quantiques sur leurs homologues classiques est due à deux avantages cruciaux. Le premier avantage réside dans un parallélisme massif résultant directement de la superposition des états quantiques. En effet, si chaque *qubit* peut être dans une superposition de deux états, un système constitué de N -*qubits* pourrait avoir accès à 2^N états simultanément. Le second avantage des ordinateurs quantiques prend racine dans « l'intrication quantique » qui n'a pas de contrepartie classique. Cette propriété quantique veut dire que si deux systèmes A et B sont intriqués, leurs propriétés peuvent être reliées même si les deux systèmes sont séparés spatialement.

Enfin, ces caractéristiques ont été exploitées dans le but de réduire la complexité en temps de certains problèmes. L'algorithme le plus célèbre, celui de Peter Shor permet de résoudre dans un temps raisonnable le problème de la décomposition d'un nombre en un produit de facteurs premiers (Shor, 1994). En revanche, malgré les avantages du modèle standard (parallélisme et intrication), David Deutsch a démontré dans les années 1980 que ce modèle était Turing-équivalent, c'est-à-dire qu'il calculait exactement les mêmes fonctions que la Machine de Turing (Deutsch 1985). Toutefois, le modèle standard n'est pas le seul modèle disponible.

En effet, un modèle alternatif de computation quantique employant des processus adiabatiques a été conçu récemment. L'idée de cette méthode est de coder la solution du problème voulant

être résolu dans l'état fondamental $|g\rangle$ d'un opérateur hamiltonien H_P , qui est l'énergie totale du système. Cependant, comme il est plus facile d'implémenter l'hamiltonien que de trouver son état fondamental, nous devons commencer la computation avec un autre état fondamental $|g\rangle$ d'un autre hamiltonien H_I , pouvant lui, être obtenu. Nous devons ensuite transformer l'hamiltonien initial H_I , au cours d'un temps T vers l'hamiltonien H_P qui possède l'état fondamental désiré, ceci à travers un processus dépendant du temps. Enfin, le théorème adiabatique de la mécanique quantique affirme que si le temps de la transformation est suffisamment lent, l'état initial évoluera vers l'état fondamental désiré avec une forte probabilité (Messiah, 1964).

À partir de ce modèle et d'après les idées de Kieu (Kieu, 2003) un ordinateur quantique utilisant les processus adiabatiques pourrait résoudre le 10^{ème} problème de Hilbert qui est indécidable par les Machines de Turing². Autrement dit, si l'ordinateur quantique de Kieu nous permettait de résoudre ce problème de manière effective, la thèse de Turing physique se trouverait falsifiée. L'algorithme de Kieu est de manière schématique le suivant :

1. Tout d'abord, étant donnée une équation diophantienne, nous codons l'équation en un nombre fini d'étapes dans un système ayant une infinité de niveaux d'énergie (tous les atomes en ont une infinité).
2. Puis, nous simulons le processus adiabatique.
3. Enfin, si l'état fondamental peut être obtenu avec une forte probabilité, une mesure praticable en un nombre fini d'étapes est menée sur ce système. Cette dernière extrait une information sur l'équation codée, indiquant qu'elle possède des solutions ou qu'elle n'en possède pas. Pour résumer, l'algorithme peut être vu comme une recherche infinie menée en un temps fini à travers les nombres entiers afin de trouver s'il existe ou non des solutions à l'équation.

3. CONCLUSION

Le résultat de Peter Shor est le résultat de la computation quantique le plus spectaculaire de tout ceux obtenus jusqu'à présent, ce qui tend à confirmer que les ordinateurs quantiques sont plus puissants en terme de vitesse que les ordinateurs classiques. Peuvent-ils pour autant résoudre des problèmes indécidables par les Machines de Turing? Deutsch dans son modèle quantique soutient

² Ce problème est le suivant : est-il possible de concevoir une procédure effective permettant de déterminer si une équation polynomiale à coefficients entiers (une équation diophantienne) a des solutions ou non?

que non : la classe des fonctions qu'un ordinateur quantique calcule est la même que celle d'un ordinateur classique. En revanche, si l'algorithme de Kieu est réalisable et nous n'avons pour l'instant aucune preuve du contraire, la thèse de Turing physique devra être modifiée d'après les ordinateurs quantiques. Toutefois, l'algorithme de Kieu est controversé car il se fonde sur de forts présupposés indispensables au bon déroulement de la procédure (Hagar and Korolev, 2007) :

1. L'exactitude de la mécanique quantique dans la description et la prédiction des processus physiques quelle que soit l'échelle de mesure (du microscopique au macroscopique).
2. Notre capacité à implémenter physiquement certains opérateurs hamiltoniens ayant un nombre infini de niveaux d'énergie.
3. Notre capacité à obtenir de manière physique les états fondamentaux.

Dans tous les cas, ne doutons pas que le débat qui va se poursuivre au sein de la recherche sera riche et passionnant.

RÉFÉRENCES

- DEUTSCH, David (1985). Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer, *Proceedings of the Royal Society of London, A* 400, pp. 97-117.
- HAGAR, Amit and KOROLEV, Alex (2007). Quantum Hypercomputation : Hype or Computation?, *Archiv*.
- KIEU, Tien (2003). Computing the Non-computable, *Contemporary Physics*, 44.
- MESSIAH, Albert (1964). *Mécanique Quantique*, Dunod.
- SHOR, Peter (1994). Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithm on a Quantum Computer, *35th Annual Symp. on Foundations of Computer Science*, pp. 124-139.
- TURING, Alan (1936). On Computable Numbers, with an Application to the Entscheidungs-problem, *Proceedings of the Mathematical Society*, 42.

FAUT-IL RENONCER À LA NOTION D'« INNÉ » ?

Valentine REYNAUD

(doctorante, Lyon III)

Avec le développement des sciences cognitives et de la psychologie du développement, la notion d'« inné » a fait l'objet d'un regain d'intérêt. Pourtant, il n'existe aucun consensus sur ce que pourrait être une *définition satisfaisante de l'innéité*. Parce que la variété des choses à laquelle le concept d'« inné » s'applique est trop importante¹, sa signification reste difficile à saisir. Le sens commun en véhiculant de fausses dichotomies est à l'origine de cette confusion.

LES FAUSSES DICHOTOMIES

Dans le sens commun, la notion d'inné est souvent opposée à celle d'*acquis*. Or, si l'acquisition se définit comme l'apparition d'un trait à un certain moment, il est évident que la plupart des traits, qui n'apparaissent qu'au cours du développement, sont acquis – y compris des traits biologiques dont tout le monde s'accorde pour dire qu'ils sont innés, comme les cheveux ou la puberté, par exemple. En outre, des facteurs environnementaux, non génétiques sont susceptibles de déterminer des caractéristiques biologiques *avant la naissance*. Ainsi, la *présence à la naissance* d'un trait est une propriété ni nécessaire ni suffisante pour prouver son innéité. La reconnaissance de cette trivialité est à la fois ce qui donne du sens à l'inné et en même temps ce qui le rend difficilement identifiable.

On trouve une autre dichotomie dans le sens commun et chez certains penseurs² qui, à la réflexion, se révèle fausse: *l'inné* est souvent opposé à *l'appris*. Or, il est évident qu'aucun organisme ne peut apprendre sans un mécanisme inné d'apprentissage³. C'est une des raisons pour laquelle l'innéité ne peut se réduire à la *maturation* biologique, qui interagit en permanence avec l'apprentissage. Par exemple, un enfant peut apprendre à parler seulement quand son cerveau est prêt pour cela, c'est-à-dire quand son cerveau a réalisé certaines maturations cérébrales. Mais l'acquisition contient en elle-même les résultats cognitifs de l'apprentissage antérieur. La distinction maturation/apprentissage véhicule une fausse idée d'unité et de direction de croissance des *patterns* comportementaux. Le terme de maturation est un moyen d'ignorer le processus réel du

¹ L'« inné » dénote des propriétés aussi hétéroclites que la présence à la naissance, l'universalité, la non acquisition par apprentissage, le caractère adaptatif, ou encore l'insensibilité à la variation des facteurs extérieurs dans le développement. Voir par exemple Samuels, 2007.

² C'est ce que fait notamment Samuels, 2007.

³ Lehrman (1953) ; Block (1981).

V. Reynaud. Faut-il renoncer à la notion d'« inné » ?

développement, l'ontogénie⁴. Au contraire, le problème propre au développement est le développement de nouvelles structures à partir de l'interaction entre structures existantes, l'organisme et l'environnement (interne et externe). La relation de l'organisme à son environnement est ainsi « bidirectionnelle » selon l'expression de Gottlieb (2001), c'est-à-dire qu'elle est constituée d'interactions permanentes. Le phénotype à chaque fois est une conséquence du phénotype précédent⁵, qui est lui-même le produit d'une interaction complexe entre facteurs génétiques et facteurs environnementaux. Ainsi, la confusion liée à la notion commune d'« inné » semble refléter la complexité du développement d'un trait. C'est pourquoi, certains chercheurs⁶, en particulier les défenseurs de la *théorie des systèmes développementaux*, ont récemment prôné le rejet de la notion d'inné.

FAUT-IL RENONCER À LA NOTION D'« INNÉ » ?

Le but de cet article n'est pas de nier l'ambiguïté sémantique du terme « inné », ni la confusion des débats portant sur l'innéité. Mais il n'est pas non plus de défendre une thèse « éliminativiste » comparable à celle des partisans de la théorie des systèmes développementaux. Mon but est plutôt de défendre la thèse suivante : les dichotomies du sens commun sont effectivement fausses, mais l'inexactitude de la conception populaire n'implique pas qu'il faille abandonner la délimitation d'un usage *scientifique* de la notion d'inné. Ce n'est pas parce que l'« inné » est trop connoté dans le sens commun que toute tentative de construire une nouvelle signification de ce terme est vouée à l'échec. Si la confusion de la notion commune d'« inné » reflète la complexité du développement, la question reste néanmoins de savoir si l'explication de celui-ci peut se passer de l'intervention de l'« inné ». Renoncer à *une certaine définition simpliste de l'innéité* et envisager *la possibilité d'une nouvelle conception de l'innéité*, dépassant les vieilles dichotomies, telle est la tâche normative que se donne la philosophie des sciences.

LA THÉORIE DES SYSTÈMES DÉVELOPPEMENTAUX : UNE POSITION ÉLIMINATIVISTE

Récemment, une nouvelle approche a émergé, *la théorie des systèmes développementaux*, qui propose de *renoncer à la notion d'inné*, jugeant que celle-ci reflète une « vision dichotomique » du

⁴ Gottlieb (2001).

⁵ Oyama (1985) ; Lewontin (1974) : "The influence of each factor depends upon the influence of the other factors so that no assignment of fixed weights to genetic, environmental, and noise components of variation is possible". Voir aussi Wimsatt, 1986.

⁶ Par exemple, Karmiloff-Smith, 1992 ; Elman et al., 1996 ; Johnson, 1997 ; Griffiths 1997, Sterelny, 2003 ; Mameli & Bateson, 2007.

développement. Les défenseurs de cette approche ont ainsi émis de sérieux doutes quant à l'utilité scientifique de cette notion qui, relevant de la « biologie populaire », reste, selon eux, enfermée dans un mode de pensée spontanément essentialiste, caractéristique de notre compréhension naïve du monde.

Griffiths (2002) dénonce « le consensus interactionniste » dominant les disciplines qui utilisent la notion d'inné (la biologie moléculaire développementale, l'écologie comportementale, la psychologie cognitive) qui prétend dépasser une dichotomie – gène/environnement – qu'il ne fait pourtant qu'entériner. Ce consensus interactionniste implique la plupart du temps les deux affirmations suivantes :

1. L'inné est défini *génétiquement*⁷.

2. Les causes du développement⁸ se divisent en deux causes symétriques – les *gènes* et *l'environnement*.

Griffiths critique vivement ce consensus, incarnant selon lui la position du sens commun. Il réfute alors les deux assertions évoquées :

1. Griffiths souligne la confusion liée à l'idée de *détermination génétique*⁹ (qu'elle soit pensée en termes de causalité génétique ou de représentation génétique) : le rôle des gènes dans le développement reste très mal compris. Nous savons seulement que les gènes sont requis dans le développement d'un trait. Pour pouvoir aller plus loin, il faudrait connaître l'influence distinctive des gènes, que nous ignorons toujours, à cause de la multiplicité des facteurs qui causent l'existence d'un trait¹⁰. Dans chaque cas, chaque trait est partiellement causé ou représenté *à la fois* dans le génome et l'environnement.

2. Griffiths insiste alors sur la complexité des causes du développement d'un trait. L'existence de plusieurs sources complexes de contraintes sur le développement qui interagissent, entraîne un scepticisme inévitable à propos de la possibilité d'une distinction entre ces différents éléments et d'une évaluation de leur

⁷ Cette assertion est directement héritée de la conception de l'éthologue Konrad Lorenz.

⁸ Le développement signifie l'ensemble des processus qui mènent de la cellule œuf à un organisme adulte.

⁹ Griffiths, 2007 ; Voir aussi Oyama 1983

¹⁰ La notion de vulnérabilité génétique par exemple ne détermine pas l'expression des gènes dans le développement : excepté les rares maladies monogéniques, les facteurs de vulnérabilité dépendent des facteurs environnementaux. Une personne peut rester vulnérable toute sa vie (posséder une certaine configuration de son génome susceptible de provoquer une certaine maladie dans certaines conditions particulières) sans jamais contracter la maladie en question.

V. Reynaud. Faut-il renoncer à la notion d'« inné » ?

influence respective.

Cette nouvelle approche qu'est *la théorie des systèmes développementaux* défend « la thèse de la parité »¹¹, refusant de traiter le gène comme un facteur de classe spéciale qui aurait un rôle causal prédominant par rapport aux autres facteurs développementaux. Elle propose de prendre comme objet d'étude, les « systèmes développementaux », qui regroupent les interactions de l'organisme avec l'ensemble des ressources développementales. Il s'agit d'étudier le développement pour lui-même, en partant des interactions et des processus qui le caractérisent. En d'autres termes, la théorie des systèmes développementaux prétend être la seule approche capable d'élucider « la cascade développementale » à travers laquelle les traits se développent¹².

Griffiths défend alors ce qu'on pourrait appeler « une thèse éliminativiste » à propos de la notion d'inné. Celle-ci en véhiculant une vision restrictive de la causalité, échoue à penser de manière adéquate le développement cognitif qui met en jeu différents éléments, opérations et interactions. Pour Griffiths, supprimer le concept d'inné revient à dénoncer comme vaine la recherche d'une partition des causes du développement d'un être vivant en facteurs génétiques et facteurs environnementaux.

LA CRITIQUE DE LA POSITION ÉLIMINATIVISTE

L'approche de Griffiths a le mérite de souligner l'importance des processus développementaux, longtemps laissés pour compte. Cependant on peut se demander dans quelle mesure cette approche, en refusant l'enquête étiologique, n'en reste pas seulement à un point de vue *descriptif*. Griffiths semble nous dire que la complexité des processus environnementaux ne peut que *se décrire* et non *s'expliquer*¹³. Si l'investigation s'interdit toute tentative de décomposition du système développemental en ses éléments, on peut se demander dans quelle mesure la description des régularités développementales épuise l'approche explicative du développement. On peut imaginer en effet qu'une pleine compréhension de la cascade développementale requiert l'intervention de l'« inné ». Que les régularités développementales soient considérées comme étant le résultat de l'interaction réciproque entre les éléments variés d'un système développemental ne justifie pas le renoncement au terme d'« inné ». Le refus d'une asymétrie explicative en faveur du gène signifie

¹¹ Griffiths & Stotz, 2000 ; Griffiths, 2002 ; Griffiths & Gray, 2005

¹² Griffiths & Stotz, 2000 ; Wolpert et al. 1998, 15 : "Gene expression is only the first step in a cascade of cellular processes that change cell behavior and so direct the course of embryonic development".

¹³ Selon Griffiths lui-même, les régularités observées constituent des « attracteurs » pour les systèmes développementaux. Or, ces attracteurs ne sont « rien de plus qu'un nouveau constat de la régularité observée » (*little more than a restatement of the observed regularity*) (Griffiths et Stoltz, 2000).

l'irréductibilité de l'« inné » au génétique ; il n'interdit pas la formulation d'une *nouvelle conception de l'innéité*.

Il est vrai que la distinction entre les différents facteurs qui entrent en jeu dans la production d'un trait est très difficile à établir, pour ne pas dire impossible. Pour le dire rapidement, il n'existe pas de relation simple entre le génotype et le phénotype qu'il est censé codé¹⁴. D'une part, un génotype peut coder plusieurs phénotypes – les gènes ont un effet « pléiotropique ». D'autres part, il est très difficile d'identifier des séquences génétiques particulières responsables de la production d'un trait sans ignorer les phénomènes de régulation entre les gènes – on parle alors d'un trait « épistatique ». Enfin, les gènes ont une incidence sur les fonctions complexes seulement en agissant sur les structures physiques qui réalisent cette fonction¹⁵. Le développement lui-même est un répertoire dynamique de processus et non pas seulement la réalisation linéaire d'un programme génétique – ce sur quoi insiste à raison la théorie des systèmes développementaux. La question reste néanmoins de savoir *si la nature du développement implique le renoncement à la notion d'inné*.

En fait, le raisonnement de Griffiths semble partir de considérations ontologiques – *la nature du développement* – (qui se reflète dans l'hétérogénéité des propriétés que dénote la notion commune d'« inné ») pour en conclure à l'inutilité épistémique de l'innéité – *le rejet du concept d'inné*. En renonçant à l'approche ontologique, il élimine tout aussi bien la possibilité d'un usage épistémique de la notion d'inné. Une question se pose alors : est-il légitime de renoncer à un usage épistémique de la notion d'inné à partir de considérations ontologiques, alors même que la définition génétique de l'innéité pose problème? La confusion de l'ontologie liée à la détermination génétique ne signifie pas que le terme d'inné n'a aucune valeur dans l'explication.

LA « CONCEPTION DE L'INVARIANCE »¹⁶ : L'INNÉ COMME NOTION DÉVELOPPEMENTALE

L'innéité n'est pas incompatible avec le développement. On peut même aller jusqu'à affirmer que *l'innéité est une notion développementale*. Encore faut-il préciser en quel sens cela s'entend.

La notion biologique de *canalisation*¹⁷ exprime le degré de convergence dans le résultat du processus développemental, malgré

¹⁴ Lewontin, 2001, p. 61 : le phénotype d'un organisme est une conséquence d'une interaction non linéaire entre le génotype et l'environnement pendant le développement. Tout ce que le gène spécifie est une « norme de réaction » sur les environnements.

¹⁵ Par exemple, dans le cas des fonctions cognitives, les gènes altèrent exclusivement la neurophysiologie du cerveau (Buller, 2005, p24).

¹⁶ Thèse dénommée ainsi par Samuels, 1998 et défendue par Ariew, 1999 et Sober, 1998 par exemple.

¹⁷ Waddington, 1952 ; Ariew, 1999.

V. Reynaud. Faut-il renoncer à la notion d'« inné » ?

la variation des stimuli, c'est-à-dire la rigidité développementale. Il existe souvent un haut degré de constance (*robustness*) des phénotypes dans un ensemble bien défini de conditions environnementales. Selon Ariew, un trait est canalisé dans la mesure où son développement est causalement insensible aux variations environnementales et génétiques. Par exemple, les orteils sont plus canalisés que la couleur de la peau, sensible au soleil. La canalisation est un processus où l'état final – le produit du développement – est manifesté en dépit des perturbations environnementales¹⁸.

Il est intéressant de souligner que cette conception appelée parfois « conception de l'invariance » fait de l'innéité une *question de degré*. Ainsi, l'apprentissage peut affecter le développement de traits même hautement canalisés. Par exemple, le langage est un phénomène hautement canalisé mais appris. Selon cette perspective, la question n'est plus de savoir quels concepts ou quelles capacités existent grâce aux gènes et lesquels par l'apprentissage mais la question devient : *comment le développement d'un concept ou d'une capacité cognitive donnés est-il canalisé ?* Dans quelle mesure la variabilité de l'environnement d'apprentissage mènera au même état final du développement ?

Cette conception permet de donner toute son importance au développement, de dépasser les dichotomies simplistes tout en conservant l'innéité grâce à la notion de « canalisation ». Elle engendre une compréhension satisfaisante de la relation complexe existant entre génotype et phénotype : un trait phénotypique est inné pour un génotype donné si et seulement si ce phénotype émergera dans certains environnements développementaux¹⁹. Les traits innés sont ceux dont le développement est invariant dans un ensemble approprié des environnements développementaux.

LA CRITIQUE DE LA « CONCEPTION DE L'INVARIANCE »

Pourtant, la conception de l'innéité comme robustesse développementale ou « canalisation » ne permet toujours pas de *distinguer entre un trait inné et un trait non inné* : en effet, comment savoir si l'invariance du développement n'est pas due à celle de l'environnement ? Comment délimiter un « ensemble approprié d'environnements développementaux » ? Sober suggère que cette délimitation doit se faire pragmatiquement, grâce notamment à des expériences où l'on élève des oisillons isolés de leurs conditions environnementales normales. Cependant, le biologiste Lehrman a bien montré que les expériences de ce type ne prouvent pas grand-chose, dans la mesure où un animal élevé en isolement n'est pas forcément isolé de l'effet de processus et d'événements qui

¹⁸ Cette conception se prolonge dans le concept biologique de « norme de réaction » qui décrit comment le phénotype qu'un génotype développe dépend de l'environnement dont il fait l'expérience. Voir Sober, 1994.

¹⁹ Sober, 1998.

contribuent au développement particulier d'un trait²⁰. S'il n'existe pas de contrôle total des effets possibles de l'expérience, l'invariance ne signifie rien puisque l'on ne sait toujours pas si le trait invariant est insensible aux facteurs environnementaux ou si les facteurs causaux pertinents sont invariants dans l'environnement dans lequel le trait se développe.

En fait, le problème est que, là encore, *la conception ontologique semble guider la conception épistémique*, ce qui revient à dissoudre la notion d'« inné » dans la complexité du développement.

EN QUEL SENS LA NOTION D'« INNÉ » EST-ELLE UNE NOTION DÉVELOPPEMENTALE ?

Il serait peut-être plus fructueux de procéder d'une autre manière : la complexité des processus développementaux, qui se reflète dans notre ignorance des mécanismes réels sous-tendant le développement, justifie *l'inversion de l'ordre du raisonnement*. Étant donnée notre ignorance sur les causes réelles du développement, ce n'est pas la recherche ontologique qui doit guider le point de vue épistémique ; en l'occurrence, le constat de la complexité du développement et des différents facteurs qui entrent en jeu dans sa production, ne légitime pas un abandon de la notion d'inné. C'est, à l'inverse, la démarche épistémique qui peut orienter les assertions ontologiques.

Si l'on peut trouver des raisons diverses de conserver le concept d'inné dans l'explication du développement d'un trait²¹, la raison la plus forte, à mon sens, est que celui-là conditionne celle-ci²². Le souci de produire *une explication convaincante du développement* implique la référence à l'« inné » compris comme la condition de cette explication. L'explication du développement, loin de nous convaincre de délaisser le concept d'inné, en a besoin pour être une véritable explication. En ce sens, elle nous permet de le délimiter plus rigoureusement : l'« inné » apparaît comme la condition de possibilité

²⁰ Lehrman, 1953.

²¹ Le rejet de cette notion reviendrait, en effet, à disqualifier l'ensemble des débats qu'elle a suscités et qu'elle suscite encore, dont l'importance est considérable (aussi bien dans la biologie moléculaire, l'écologie comportementale ou encore la psychologie cognitive). À l'inverse une clarification de cette notion permettrait de maintenir le lien entre ces différentes disciplines.

²² Griffiths et Stotz (2000) eux-mêmes acceptent l'idée de « contraintes faibles » (*soft constraints*) sur le développement. Celles-ci sont contingentes et probabilistes car elles dépendent de différents paramètres qui mettent en place la cascade développementale. Elles émergent et disparaissent comme des parties du processus développemental. Néanmoins, ces contraintes présupposent sans aucun doute que l'organisme possède certaines capacités innées qui leur permet d'émerger lors de l'interaction avec les autres ressources développementales.

des explications développementales²³. *La notion d'inné est nécessaire à l'explication du développement d'un trait.*

Ainsi, l'étude du développement – la tentative de décrire des scénarii possibles de construction des capacités²⁴ – peut permettre d'énoncer une *explication développementale satisfaisante*. Une fois en possession de celle-ci, l'on peut alors se demander quels sont les termes qu'il faut attribuer à l'organisme qui rendent possible cette explication. Ces termes que je qualifie de *conditionnels* et de *primitifs* délimitent ce qu'on peut raisonnablement appeler l'« inné ».

RÉFÉRENCES

- ARIEW, A. (1999). Innateness is canalization: In defense of a developmental account of innateness. In V. Hardcastle (ed.) *Where Biology Meets Psychology*. Cambridge, MA: MIT Press.
- BLOCK, N. (1981). *Readings in Philosophy of Psychology*, volume 2, edited by Ned Block.
- BULLER, D. (2005). *Adapting Mind*, MIT Press, Cambridge.
- GOTTLIEB, G. (2001). A developmental psychobiological systems view: Early formulation and current status- *Cycles of contingency: Developmental systems and evolution*. Edited by Susan Oyama, Paul E. Griffiths & Russell D. Gray. Cambridge. MA: The MIT Press.
- GRIFFITHS P.E. & Gray R.D. (1994). Developmental Systems and Evolutionary Explanation. *The Journal of Philosophy*, Vol 91, No. 6 (Jun., 1994), 277-304.
- GRIFFITHS P.E. & STOTZ, K. (2000). How the mind grows: a developmental perspective on the biology of cognition. *Synthese* 122(1-2): 29-51.
- GRIFFITHS P.E., (2002). What Is Innateness? - *The Monist* - philsci-archive.pitt.edu.
- GRIFFITHS, Paul E, (2002). *The Philosophy of Molecular and Developmental Biology*. The Blackwell Guide to the Philosophy of

²³ Ma proposition de faire du terme « inné » une condition de l'explication développementale ne revient pas à définir l'inné comme un terme primitif au sens de Samuels (2002). Selon lui en effet, l'inné est primitif au sens où il n'est pas appris, c'est-à-dire au sens où il ne peut pas être expliqué par des mécanismes psychologiques. Or, l'explication des traits non innés ne fait pas toujours intervenir des mécanismes psychologiques et, à l'inverse, certains traits innés peuvent être expliqués par des mécanismes psychologiques. Selon moi, l'inné est un terme primitif au sens où l'attribution de certaines capacités innées à l'organisme est une condition de l'explication développementale.

²⁴ C'est-à-dire essayer de comprendre quelle capacité est rendue possible par une autre. Dans le vocabulaire de Wimsatt, 1986, il s'agit d'identifier quelle capacité est « enchâssée » (entrenched) par d'autres.

Science.

LEHRMAN, D. (1953). A critique of Konrad Lorenz's theory of instinctive behavior. In *Cycles of contingency: Developmental systems and Evolution*. Edited by Susan Oyama, Paul E. Griffiths and Russell D. Gray. Cambridge. MA: The MIT Press.

LEWONTIN, R.C. 2001 (1983). Gene, organism and environment: A new introduction- *Cycles of contingency: Developmental systems and Evolution*. Edited by Susan Oyama, Paul E. Griffiths and Russell D. Gray. Cambridge. MA: The MIT Press.

MAMELI, M. & BATESON, P. (2007). The innate and the acquired: useful clusters or a residual distinction form folk biology ? *Developmental Psychology*. 49 (8): 818-31.

OYAMA, S. (1985). *The Ontogeny of Information*. Cambridge: Cambridge University Press.

SAMUELS, R. (1998). "What brains won't tell us about the mind". *Mind & Language*, 13, 4, 548-570.

SAMUELS, R. (2002). Nativism in Cognitive Science. *Mind & Language*, 17, 3, 233-265.

SAMUELS, R. (2004). Innateness in Cognitive science, *TRENDS in Cognitive Sciences*, Vol.8, No3 March 2004, 136-141.

SAMUELS, R. (2007). Is Innateness a Confused Concept ? In P. Carruthers, S. Laurence & S. Stich (eds) *The Innate Mind* vol.3 *Foundations and the future*. Oxford University Press 2007.

SOBER, E. (1994). *From a Biological Point of View: Essays in Evolutionary Philosophy*, Cambridge University Press.

SOBER, E. (1998). "Innate Knowledge" (pp. 794 - 797, vol. 4) *Encyclopedia of Philosophy*, Routledge.

WADDINGTON, C. H. (1952). *The evolution of developmental systems*. Paper presented at the Twenty-eighth Meeting of the Australian and New Zealand Association for the Advancement of Science, Brisbane, Australia.

WIMSATT, W. (1986). « Developmental constraints, generative entrenchment, and the innate-acquired distinction » In William Bechtel, éd. 1986. *Integrating Scientific Disciplines: Case Studies From the Life Sciences*. Dordrecht: Nijhoff, p. 185-208.

WOLPERT L., BEDDINGTON R., BROCKES J., JESSELL T., LAWRENCE P.A. and MEYEROWITZ E.M. (1998). *Principles of Development*. Oxford University Press, New York.

L'ÉTRANGETÉ DÉLIRANTE : LA PHILOSOPHIE À L'ÉPREUVE DE LA PSYCHOPATHOLOGIE CLINIQUE

Sarah Troubé
(Elève à l'ENS)

Traditionnellement, les croyances délirantes sont souvent apparues comme l'envers et le négatif du bon sens, le philosophe les mentionnant alors pour renforcer par exclusion sa conception de la raison. Or, au sein de la philosophie des sciences cognitives, le délire s'est au contraire constitué depuis quelques années comme objet spécifique de réflexion et de débat, dans une interaction étroite avec les neurosciences et la psychiatrie clinique. Les termes de ce débat portent à la fois sur la nature du délire (croyance, imagination, actes de langage sans contenu¹...), et sur son statut par rapport à des troubles primitifs ou de plus bas niveau. Le problème du rapport entre délire et raison ou sens commun est sous-jacent à ces questions. Deux orientations majeures pour les sciences cognitives, l'exploration de troubles fonctionnels de bas niveaux et la relativisation, par les expériences menées en psychologie sociale et cognitive, de la figure d'un agent rationnel idéal, ont en effet rendu possible une reformulation de ce problème. Une hypothèse suppose ainsi que le délire ne serait pas en tant que tel le fruit d'un dysfonctionnement d'ordre cognitif, mais consisterait en une croyance traduisant ou expliquant une expérience pathologique primitive. Elle se fonde sur des études empiriques mettant en évidence, dans des pathologies délirantes, des troubles de certaines composantes de la perception, qui induiraient notamment un sentiment d'étrangeté accompagnant cette perception. On exposera ici les formulations et les difficultés de cette conception en soulignant la fécondité de l'analyse philosophique pour éclairer le délire, et en défendant la nécessité et l'intérêt, pour la philosophie cognitive, d'une attention portée aux manifestations et aux descriptions cliniques des différents types de délire.

LE DÉFI PHILOSOPHIQUE DE LA NATURE DU DÉLIRE

Si le délire semble d'emblée se situer hors de tout sens commun, c'est parce qu'il dégage par son contenu et son mode d'affirmation une bizarrerie manifeste. Il consiste en effet en des affirmations qui s'écartent de la réalité objective et des croyances communément et socialement partagées, et qui sont maintenues sur le mode d'une certitude à toute épreuve. Cette bizarrerie est particulièrement saillante dans les cas où l'affirmation délirante ne semble pas seulement tenir de l'invraisemblable au vu des faits (telle l'idée d'un complot dirigé contre le sujet, ou l'idée qu'un proche a été remplacé par un sosie), mais paraît contredire les schémas habituels

¹ (Currie, 2000) ; (Bayne & Pacherie, 2005) ; (Berrios, 1991)

S.Troubé. L'étrangeté délirante

du raisonnement logique ou causal : affirmation d'être en état de mort, ou dirigé à distance par la volonté de Dieu ou de Satan, par exemple².

Si le fait que le sujet tienne son délire comme reflétant véridiquement un état du monde, et le fait que ces affirmations semblent bien posséder un contenu représentationnel, ont traditionnellement amené à assimiler le délire à un type de croyance, ces derniers exemples mettent en lumière l'ampleur des difficultés qu'une telle conception adresse à la philosophie analytique de la connaissance. Elle pose tout particulièrement question pour le processus de formation des croyances, les contraintes holistiques pesant sur leur consistance, et l'attitude épistémique adoptée envers elles³. Comment une croyance peut-elle être ainsi formée en ne semblant s'appuyer sur aucun élément de réalité, et générer une certitude, plutôt que d'être révisée comme une croyance fautive, ou maintenue avec un engagement moindre, comme peuvent l'être, par exemple, certaines croyances superstitieuses ?

Si la définition du délire en termes de croyance semble intuitivement séduisante en ce qu'elle lui confère une visée de sens et semble ainsi correspondre au point de vue des patients sur leurs affirmations, ces difficultés ont conduit certains auteurs à nuancer ou critiquer cette conception doxastique. Cette dernière n'aboutit-elle pas en effet à effectuer une telle révision du concept traditionnel de croyance, qu'il semble plus adéquat de penser que le délire n'a que l'apparence de la croyance, qu'il relève en réalité d'une autre nature et se présente faussement, tant au sujet délirant qu'au clinicien, sous le déguisement de la croyance ? Il a été défendu dans cette perspective que l'idée de contenu devait elle-même être remise en cause, les affirmations délirantes consistant dans des actes de langage vides de sens et ne visant aucun référent dans le monde⁴. Une autre conception, conciliant à la fois l'idée de contenu et le refus de la nature doxastique du délire, propose de considérer ce dernier comme une attitude relevant de l'imagination : le délire serait alors le résultat d'une erreur se situant au niveau métacognitif, le sujet méconnaissant cette nature imaginative et formant alors la croyance

² L'idée d'un complot est très fréquente dans les délires de la paranoïa et de la schizophrénie dite paranoïde. L'idée du sosie et celle de la mort du sujet sont respectivement considérées comme caractéristiques de syndromes spécifiques, le syndrome de Capgras, sur lequel nous reviendrons, et le syndrome de Cotard. L'idée d'une action causale à distance sur le sujet se retrouve quant à elle fréquemment dans la schizophrénie, et est partie prenante de ce qui a été nommé le délire d'influence.

³ (Engel, 2001)

⁴ (Berrios, 1991)

de second niveau que ce contenu relève d'une croyance⁵.

Affirmer que la nature du délire ne relève pas de la croyance permet de rendre plus compréhensible ce qui apparaît du point de vue de la conception doxastique comme une incohérence dans l'attitude du sujet délirant : contradictions avec les faits et avec ses croyances générales, et, dans de nombreux cas, comportements et affects en inadéquation avec le contenu délirant⁶. Si, à l'inverse, la conception doxastique peut rendre compte des cas où le contenu délirant guide effectivement l'action du sujet, et s'appuyer sur la relativisation des exigences pesant sur le fonctionnement des croyances, l'un des arguments invitant à considérer le poids de cette conception tient également dans la reformulation, au vu de certaines données empiriques, du problème du rôle de l'évidence dans la formation du délire. Cette reformulation suggère en effet que loin d'être le fruit d'un dysfonctionnement métacognitif ou d'un processus irrationnel de formation de la croyance contraire à l'évidence des faits, le délire serait une croyance dont la certitude serait fondée rationnellement sur une évidence perceptive altérée de la réalité.

Une telle hypothèse met alors le philosophe au défi de décrire le statut de la croyance qui découle de cette évidence, en la faisant entrer dans la sphère du compréhensible, tout en rendant compte de sa spécificité par rapport à d'autres types de croyances irrationnelles. En effet, si la croyance délirante se forme à partir d'une expérience, elle semble acquérir une signification compréhensible. Mais elle ne peut pour autant être rapportée à une simple croyance fautive ou fantaisiste, sous peine de voir son caractère pathologique échapper à toute explication. Les résultats expérimentaux de la neuropsychopathologie appellent ainsi des descriptions, des évaluations et des comparaisons, qui sont autant de questions adressées à l'analyse conceptuelle philosophique : par exemple, quelles sont les propriétés d'une expérience perceptive d'étrangeté⁷ ? Quels sont les modes de passage d'une expérience à une croyance, et quel rôle ou quelle fonction remplit alors la croyance par rapport à cette expérience ? Dégager les implications des résultats

⁵ (Currie, 2000). Cette conception fait notamment appel à deux éléments que nous ne pouvons approfondir dans le cadre de cet article : la compréhension des états d'imagination comme actes de simulation, renforçant la possibilité de prendre l'imagination pour l'état mental qu'elle simule, et les travaux liant la schizophrénie à un dysfonctionnement de l'agentivité, altérant la capacité à reconnaître la simulation comme un acte dont le sujet est l'agent. La conception de Currie semble par ailleurs en miroir de celle qui est proposée par Campbell, selon laquelle les affirmations délirantes consistent en des croyances, mais n'ont pas le contenu qu'elles semblent posséder : (Campbell, 2001).

⁶ Voir (Bayne & Pacherie, 2005) pour une discussion de cette question, et une défense de la conception doxastique.

⁷ (Ratcliffe, 2008)

expérimentaux pour le rapport du délire au sens commun place la philosophie dans une position de jonction entre la neuropsychopathologie et les expériences de psychologie cognitive portant sur les propriétés et le fonctionnement de ce sens commun.

DE L'ÉTRANGETÉ PERCEPTIVE À LA CROYANCE DÉLIRANTE

Face aux difficultés que pose l'hypothèse d'une expérience altérée à la source du délire, différentes formulations en ont été proposées. En effet, d'un côté, cette hypothèse conforte le caractère rationnel de la croyance délirante : le contenu étrange du délire serait hérité du contenu étrange de la perception, et la certitude serait naturelle, le contenu d'une expérience perceptive pouvant être considéré comme une preuve particulièrement digne de confiance pour former et affirmer une croyance⁸. Cependant, que la croyance délirante devienne ainsi moins hermétique fait d'autant plus ressortir son écart avec le sens commun : le contenu invraisemblable de la plupart des délires ne vient-il pas renforcer l'étrangeté de la perception, plutôt que l'exprimer ou lui donner sens ? L'attitude rationnelle face à une expérience étrange ne serait-elle pas de l'expliquer par une croyance cohérente avec les autres éléments de réalité, les autres croyances du sujet et les croyances communes, et de prêter attention au fait que les autres ne semblent pas partager cette expérience étrange ?

Face à cette difficulté, il est à présent majoritairement admis que l'explication du délire dite à « un seul facteur » doit être reformulée dans une version dite « à deux facteurs », c'est-à-dire que la présence d'une expérience pathologique de bas niveau doit être associée, pour expliquer les caractéristiques du délire, à des biais spécifiques affectant la formation de la croyance : par exemple, le fait de passer trop rapidement d'une donnée d'expérience à une croyance sans en examiner suffisamment le caractère logique ou vraisemblable, ou encore, d'attribuer la cause de cette expérience à autrui ou au monde plutôt qu'à un dysfonctionnement affectant l'expérience propre⁹. Sur la question du statut de la croyance par rapport à l'expérience, deux versions s'opposent : l'une conçoit cette croyance comme une explication de l'expérience¹⁰, l'autre comme une approbation ou une traduction de son contenu. Il semble que cette dernière version rende davantage compréhensible l'adoption et le maintien de la croyance délirante sans examen critique¹¹. Considérer la croyance délirante comme une explication rend plus difficilement compte de son contenu, qui est alors sous-déterminé par le contenu de la perception, mais aussi de sa fonction, au sens où si une explication vise à réduire le caractère étrange ou inhabituel d'un

⁸ [Maher, 1974]

⁹ [Davies et al., 2001]

¹⁰ [Maher, 1974]

¹¹ [Pacherie, 2008]

phénomène, que cette visée soit atteinte dans la croyance délirante semble peu évident.

DÉLIRES MONOTHÉMATIQUES ET POLYTHÉMATIQUES

Ces différentes versions doivent être discutées non pas seulement à l'aune d'une analyse de la nature et du fonctionnement des croyances, mais aussi des manifestations cliniques du délire. Les descriptions fournies par la psychiatrie constituent ainsi une mise à l'épreuve des travaux philosophiques amenés par les résultats expérimentaux, et une donnée essentielle invitant le philosophe à en proposer des redescriptions, à formuler les problèmes spécifiques aux différents types de délire, et à questionner les conditions d'application de ses modèles.

L'hypothèse d'une expérience altérée amenant la formation du délire illustre particulièrement cette nécessité d'une attention portée à la réalité clinique, car elle a été formulée à partir de deux types de délire. Le premier est le délire schizophrénique, caractérisé par une élaboration souvent complexe et un enchevêtrement de thèmes et de croyances : c'est un délire « polythématique ». La multiplicité des hypothèses sur les déficits ou troubles présents dans la schizophrénie ne permet pas actuellement une description unique de ce qui pourrait constituer l'altération perceptive primitive. Il a été supposé qu'il pourrait s'agir d'un sentiment d'étrangeté, ou de signification, reposant notamment sur un trouble de la capacité à discriminer, hiérarchiser et contextualiser les stimuli¹². Le second type de délire, qui a suscité des propositions plus précises et amené à leurs différentes versions, est le syndrome de Capgras : il est dit « monothématique » car il tourne autour d'une seule croyance circonscrite, celle qu'un proche du sujet a été remplacé par un sosie. Cette croyance surgirait suite à l'absence d'une composante de la perception des visages familiers, celle d'un sentiment implicite de familiarité : la reconnaissance formelle demeurerait quant à elle intacte, amenant ainsi rationnellement le sujet à l'idée du sosie¹³.

Les difficultés soulevées par ces deux types de délire sont distinctes sur de nombreux points. Ainsi, la version dite « à deux facteurs » semble poser davantage problème pour les délires monothématiques, à propos desquels elle a pourtant été forgée, car elle pose la question de savoir pourquoi les biais de raisonnement restent cantonnés à cette croyance isolée¹⁴. Les délires polythématiques, quant à eux, rendent plus difficile d'espérer faire correspondre chaque croyance délirante à une altération particulière de la perception. Si une altération générale est supposée, se pose alors le problème de la sous-détermination du contenu de la

¹² (Maher, 1974 ; 1999)

¹³ (Ellis & Young, 1990)

¹⁴ (Pacherie, 2008)

croissance par rapport à cette altération globale, qui rend difficile d'appliquer telle quelle la version d'une approbation ou d'une expression directe du contenu perceptif dans la croyance. Concevoir le rapport entre l'expérience et la croyance comme un rapport explicatif a ainsi pour avantage de prendre en compte le caractère d'élaboration des croyances et leur construction en système dans ces délires.

DU DÉLIRE COMME CROYANCE AU DÉLIRE COMME CONSTRUCTION

La généralisation des propositions formulées à partir du syndrome de Capgras aux délires polythématiques dépend pour une grande part des recherches expérimentales, sur la schizophrénie notamment. Cependant, à côté d'une approche posant la question légitime de la généralisation de délires considérés comme « plus simples » vers des délires paraissant « plus complexes », et d'une approche cherchant à décrire l'expérience perceptive de la schizophrénie, nous aimerions souligner la fécondité des questions qu'adressent au philosophe les délires élaborés polythématiques. Dans la mesure où une explication « à deux facteurs » paraît particulièrement appropriée pour ces délires, l'étude des biais invite à porter attention non seulement au passage entre l'expérience et la croyance, mais aussi à l'articulation des croyances entre elles et à leur agencement en un système. En effet, ces délires, parfois élaborés en systèmes très développés et rigoureux, posent d'emblée non pas la question du délire comme contenu d'une croyance, mais comme construction ou schéma global d'interprétation.

C'est en cela que l'idée d'un délire possédant le statut d'une explication paraît séduisante : le corrélat positif du problème de la sous-détermination est que cette hypothèse invite à poser la question de l'articulation des croyances en théorie, de la mise en œuvre des liens logiques ou causaux, ou encore du processus d'extension progressive des croyances délirantes. Évaluer cette hypothèse appelle ainsi une analyse du contenu de l'expérience perceptive altérée, mais également de la structure du délire : par exemple, le passage progressif d'un vécu d'étrangeté à la construction du délire, l'implication du renforcement des croyances entre elles pour le problème de la certitude, ou encore la possibilité de croyances possédant des statuts différents, l'élaboration secondaire du délire pouvant répondre à un besoin de justification de la croyance centrale. Interroger l'« épistémologie » propre à ces délires invite alors à une confrontation avec les explications construites par le sens commun, et avec la fonction qu'elles peuvent revêtir. Même si l'hypothèse d'un statut explicatif du délire s'avérait rencontrer des difficultés insurmontables, l'attention portée à la clinique des délires polythématiques apparaît ainsi essentielle : elle appelle, en soulevant des questionnements que ne posent pas les croyances isolées des délires monothématiques, des contributions particulièrement

fécondes au sein de la philosophie cognitive.

RÉFÉRENCES

- BAYNE, T. & PACHERIE, E. (2005). "In defence of the doxastic conception of delusions". *Mind and Language* 20, 163-188.
- BERRIOS, G. (1991). "Delusions as 'wrong beliefs': A conceptual history". *British Journal of Psychiatry* 159, 6-13.
- CAMPBELL, J. (2001). "Rationality, meaning, and the analysis of delusions". *Philosophy, Psychiatry & Psychology* 8, 2/3, 89-100.
- CURRIE, G. (2000). "Imagination, delusion and hallucinations". In M. Coltheart and M. Davies (eds), *Pathologies of Belief*. Blackwell : Oxford, 167-182.
- DAVIES et al. (2001). " Monothematic delusions : Toward a two-factor account". *Philosophy, Psychiatry & Psychology* 8, 2/3, 133-158.
- ELLIS, H.D. & YOUNG, A.W. (1990). "Accounting for delusional misidentifications". *British Journal of Psychiatry* 157, 239-248.
- ENGEL, P. (2001). "Peut-on parler de croyances délirantes ?". In J. Chemouni (dir), *Clinique de l'intentionnalité*. In-press : Paris, 157-173.
- MAHER, B. (1974). "Delusional Thinking and Perceptual Disorder". *Journal of Individual Psychology* 30, 1, 98-113.
- MAHER, B. (1999). "Anomalous experience in everyday life : Its significance for psychopathology". *The Monist* 82, 4, 554-570.
- PACHERIE, E. (2008). "Perception, emotions and delusions : Revisiting the Capgras Delusion". In T. Bayne & J. Fernandez (eds), *Delusions and Self-Deception*. Psychology Press : Hove, UK, 105-123.
- RATCLIFFE, M. (2008). "What is a feeling of unfamiliarity ?". *Philosophy, Psychiatry & Psychology* 14, 1, 43-49.

LES ÉPOUVANTAILS TRANSPARENTS N'EXISTENT PAS. NOTE POUR UNE PHILOSOPHIE EXPÉRIMENTALE À LA FRANÇAISE Xavier Kieft (Paris Sorbonne)

La philosophie expérimentale a le rare mérite d'avoir réuni autour d'elle (c'est-à-dire sous sa bannière ou contre elle) un très grand nombre d'acteurs du monde philosophique en moins d'une dizaine d'années. Ce mérite ne tient pas seulement à la lourde artillerie communicationnelle que ses principaux représentants et ses *aficionados* ont mis à son service, particulièrement via Internet, à son titre, ses blogs, son acronyme branché et aux frontières du réel (*x-phi*), son hymne, son logo au fauteuil enflammé, ses immanquables produits vestimentaires dérivés. Il tient aussi – et au fond, surtout – à l'intérêt critique de cette nouvelle tendance¹. Elle a, dès ses débuts², fait preuve d'une vertu aussi simple que salutaire : la capacité à remettre en cause des idées reçues. En l'occurrence, il s'agissait de montrer que certaines de nos intuitions théoriques étaient mal fondées et non universellement valides, l'intuition semblant dépendre du contexte dans lequel on la sollicite, que ce dernier soit culturel, moral, ou discursif³.

En réalité, la chose ainsi assénée n'a rien d'une révélation fracassante, et quelques décennies d'études sur le holisme de la croyance ont très largement préparé le terrain de cette découverte⁴. Mais ce qui change ici et participe pour une part importante au succès de l'*x-phi* est la manière dont elle combat la philosophie de salon⁵. Le philosophe expérimental (le philosophe *x*) se rend sur le terrain et interroge les gens, compose des questionnaires et collecte des réponses qui constituent *ses data*. Puis il classe et interprète ces réponses de façon à en faire saillir les spécificités, ce qui le conduit à discuter la pertinence d'intuitions jusque-là ordinairement admises⁶. La spécificité de cette philosophie nouvelle consiste donc

¹ L'article de J. Knobe traduit ici, « qu'est-ce que la philosophie expérimentale ? » est une très brève présentation de ce « mouvement ».

² Voir par exemple, J. Weinberg, S. Nichols & S. Stich, « Normativity and epistemic intuitions », *Philosophical Topics*, 29, 2001, p. 429-460.

³ Le lecteur intéressé que la brièveté de « Qu'est-ce que la philosophie expérimentale ? » aura laissé sur sa faim se reportera avec profit au texte plus complet et engagé de J. Knobe et S. Nichols : « An experimental philosophy manifesto », J. Knobe et S. Nichols (ed.), *Experimental philosophy*, Oxford, OUP, 2008, p. 3-16, ainsi qu'à la très complète bibliographie accessible sur <http://experimentalphilosophy.org>.

⁴ Voir, entre autres, W.V.O. Quine & J.S. Ullian, *The Web of Belief*, New York, McGraw Hill, Inc., 1978.

⁵ « La philosophie de salon » est ici une adaptation non littérale de l'*armchair philosophy*, la philosophie en fauteuil. Pour un appel pré-*x-phi* à l'abandon du fauteuil, on peut par exemple voir D.C. Dennett, « Out of the armchair and into the fields » [1988, repris dans *Brainchildren*, Cambridge, Mass. MIT Press, 1998, p. 289-306].

⁶ Pour le dire de manière plus subtile : « La philosophie expérimentale utilise la méthode philosophique traditionnelle qui consiste à rapporter des intuitions et y adjoint l'analyse

X. Kieft. Les épouvantails transparents n'existent pas

essentiellement dans la méthode qu'elle déploie pour l'analyse conceptuelle.

Aussi suggestif qu'ils soient, les résultats de l'*x-phi* demeurent encore essentiellement négatifs et propédeutiques, voire préparatoires. On n'apprend en effet peu de choses sur une intuition en constatant ce qu'elle n'est pas et en remarquant que telles ou telles intuitions, quoi que formellement semblables, sont en fait divergentes⁷. Mais, indéniablement, une nouvelle voie se dégage. C'est pourquoi il est légitime de s'interroger plus avant sur le traitement des *data* du philosophe *x* en rapport avec son objet de recherche. Mais se pourrait-il qu'il n'y ait rien de tel que des intuitions⁸?

Les « intuitions » à propos desquelles on mène l'enquête, si elles soutiennent fonctionnellement⁹ les réponses verbales, doivent être estimées avoir une certaine consistance et être propres à la personne interrogée, de sorte que dans des conditions similaires d'interrogation, le même sujet répondra de manière similaire. Si ce n'est pas le cas, les réponses ne seront dues qu'au hasard (pour autant qu'il existe) ou à des facteurs extérieurs causant eux-mêmes les dites « intuitions », qui ne formeront alors plus qu'un concept fourre-tout et vide de sens. Dans ce cadre, le philosophe *x* ne disposera d'aucun objet d'étude réel et s'appuiera seulement sur un concept *ad hoc* d'intuition. Or, le réquisit de consistance est assez vraisemblablement intenable, pour plusieurs raisons, dont je ne présenterai ici rapidement que quelques échantillons – le paradoxe de l'entreprise *x* consistant précisément dans l'effort qu'on y déploie pour faire la preuve de l'impossibilité de satisfaire à ce réquisit fondamental.

1. D'abord, une personne à qui l'on pose plusieurs fois à peu près la même question (que ce qui varie et cause l'« à peu près » soit le contexte, la répétition de la question ou une légère modification de son énoncé) ne donne pas forcément la même réponse. Les indécis l'expérimentent chaque jour, comme les amoureux inquiets¹⁰, les

statistique de manière à critiquer les affirmations des philosophes concernant ce que les intuitions faisant autorité devraient être » (J.J. Prinz, « Empirical philosophy and experimental philosophy », J. Knobe et S. Nichols, ed., *Experimental philosophy*, op.cit., p. 199 ; je traduis).

⁷ On distinguera ici soigneusement deux perspectives : celle qui exhibe la diversité des « intuitions » effectives des gens sur le terrain et celle qui fait valoir cette diversité contre la pertinence des « intuitions » philosophiques de salon. Voir E. Sosa, « Experimental philosophy and philosophical intuition », *Philosophical Studies*, 132, 2006, p. 99-107 (repris dans J. Knobe & S. Nichols, ed., *Experimental philosophy*, op. cit., p. 231-240). Pour d'autres critiques subtiles, voir A. Kauppinen : « The rise and fall of experimental philosophy », *Philosophical Explorations*, 10, 2007, p. 95-118.

⁸ Je paraphrase ici la question fatale soulevée par S. Stich, l'un des parrains et fondateurs de l'*x-phi*, dans son maître-livre : *From folk Psychology to Cognitive Science: The Case against Belief*, Cambridge, Mass., 1983, p. 228 : « *Could it turn out that there are no such things as beliefs?* ».

⁹ J'utilise ici « fonctionnellement » dans le sens présenté dans un article *x* suggestif à propos de ces questions : J. Knobe & J.J. Prinz, « Intuitions about consciousness: Experimental studies », *Phenomenology and Cognitive Science*, 7, 2008, p. 67-83.

¹⁰ « Tu m'aimes ? Dis, tu m'aimes ? » - combien de fois cette question doit-elle être répétée pour

petits chefs¹¹ et les hommes politiques¹².

2. Ensuite, les réponses possibles du sujet sont toujours anticipées (et donc inventées) par le philosophe x ¹³, de sorte qu'elles sont avant tout les siennes – et, dans un sens encore très indéterminé « intuitives ». Il s'agit d'ailleurs là de l'un des impératifs du processus expérimental auquel satisfait la méthode d'enquête x de prédilection : le QCM. Pourtant, il arrive que les réponses proposées ne correspondent pas à tout ce que peuvent penser les gens (pour autant qu'ils pensent bien quelque chose). Par exemple, les questions posées lors du *Moral Sense Test* de M.D. Hauser et alii¹⁴, qui se rapportaient à des formulations ou des variantes du « problème du wagon »¹⁵, ne permettaient pas toujours des réponses de principe du type « ce comportement est inacceptable » : le sujet testé devait parfois se contenter de quantifier le nombre de victimes ou de rescapés pour évaluer l'acceptabilité ou l'utilité d'une action.

3. De surcroît, l'« effet de sondage » est sans doute trop négligé par les philosophes x . Or, s'il est compréhensible (mais non problématique, comme je l'ai indiqué) que l'enquêteur x anticipe les réponses possibles du sujet interrogé, force est de constater qu'un sondé anticipe les attentes (voire les pensées et l'interprétation supposées) du sondeur, ce qui modifie parfois considérablement les réponses données, qui peuvent ne pas restituer de manière transparente l'avis dudit sondé, et ne le font sans doute que

que la réponse proposée soit négative ?

¹¹ « Vous êtes sûrs ? » suffit parfois à obtenir une autre réponse que celle qui a été formulée en première instance. J'ai pu faire l'expérience de ce point sur un bon nombre de sujets placés dans une situation d'infériorité hiérarchique (toutes compétences intellectuelles étant égales par ailleurs). Les étudiants de philosophie passant une épreuve orale, qu'on aurait pu croire plus aguerris et formés à la ténacité intellectuelle, font en effet ici des cobayes exemplaires.

¹² La « pédagogie » politique contemporaine consiste souvent à « réexpliquer » les projets de lois impopulaires, sans en modifier une ligne.

¹³ La chose est aussi vieille que la science expérimentale elle-même. Voir par exemple la distinction des trois sortes d'expériences décrites par J. Rohault dans la préface de son *Traité de physique* [1671].

¹⁴ <http://moral.wjh.harvard.edu>

¹⁵ Un wagon hors de contrôle roule sur une voie ferrée sur laquelle sont attachées cinq personnes. Un commutateur permet de faire bifurquer le wagon sur une seconde voie sur laquelle une autre personne est attachée. Faut-il actionner le levier du commutateur ? – Le « problème du wagon » (*Trolley problem*) est apparu dans les discussions de l'utilitarisme moral, notamment chez P. Foot (« The problem of abortion and the doctrine of the double effect », *Oxford Review*, 5, 1967, p. 5-15), qui introduit la notion de *side-effect* – dont l'équivalent militaire est le dommage collatéral. L'article de P. Foot est une laïcisation de réflexions classiques en théologie morale (voir par exemple J. Mangan, « An historical analysis of the principle of double effect », *Theological Studies*, 10, 1949, p. 41-61 et, sur le « double effet », P.A. Woodward, ed., *The Doctrine of Double Effect*, Notre Dame, University of Notre Dame Press, 2002). Le dilemme est ensuite repris sous son nom actuel par J.J. Thomson (« Killing, letting die and the trolley problem », *The Monist*, 59, 1976, p. 204-217). Il constitue le fonds de commerce des questionnaires moraux de l'*x-phi*. Le fameux « *Knobe effect* » [ou « *side-effect effect* »], qui n'est pas pour rien dans la popularisation de l'*x-phi*, dérive évidemment des interrogations sur le « double effet ». Voir J. Knobe, « Intentional action and side effects in ordinary language », *Analysis*, 63, 2003, p. 190-193. Ce dernier article, dont le premier cas est repris pour la discussion de la théorie de l'action dans « Qu'est-ce que la philosophie expérimentale ? », peut faire figure d'acte de naissance de l'*x-phi*.

X. Kieft. Les épouvantails transparents n'existent pas

rarement¹⁶. Seule une minorité des électeurs du Front National se présente comme telle aux sondages, nous répète-t-on chaque soir d'élection, tout particulièrement depuis le 21 avril 2002. De même, l'envie de proposer un discours cohérent ou de représenter une position engagée incite parfois à modifier les réponses qui se présentent les premières¹⁷, sans que l'enquêteur puisse légitimement choisir de remettre en cause telle ou telle réponse particulière effectivement donnée.

D'ailleurs, tout comme l'anticipation des attentes de l'interrogateur donne à l'interrogé l'impression d'une marge de manœuvre possible, l'arrachage de l'aveu relève d'une stratégie d'emprise sur le sujet sondé. Une partie des enjeux de la relation des intéressés est alors très indépendante du contenu propositionnel de l'échange, et peut influencer considérablement le type des réponses par rapport à l'attente anticipée (comme les questions « intuitivement » anticipées en dépendent dans la constitution du questionnaire). N'a-t-on pas dit que les Français avaient voté « non » au référendum sur la Constitution européenne du 29 mai 2005 pour exprimer, non pas leur avis sur la constitution, mais leur mécontentement face à une classe politique quasi-unanimement favorable au « oui »¹⁸? Ne peut-on expliquer le refus de la majorité des mêmes Français de se vacciner contre l'effroyable grippe A H1N1 durant l'hiver 2009-2010 par la volonté de ne pas se plier aux attentes d'un gouvernement trop enclin à s'immiscer dans la vie privée des gens¹⁹? Il se pourrait bien que le plus dupé ne soit pas celui à qui il revient de faire le constat de l'instabilité de ses « intuitions » par le biais de l'analyse des résultats d'un questionnaire, mais celui qui pense qu'un avis exprimé restitue bien quelque chose qu'on peut nommer « une intuition »²⁰.

La philosophie x est pleine d'un entrain salubre. Puisse-t-elle en profiter pour enrichir davantage ses investigations en assaisonnant ses recettes *cum grano salis* « à la française ». Descartes, notre gloire nationale, le disait admirablement dans le *Discours de la méthode* : « Pour savoir quelles étaient véritablement leurs opinions, je devais plutôt prendre garde à ce qu'ils pratiquaient qu'à ce qu'ils disaient [...] à cause qu'en la corruption de nos mœurs il y a peu de gens qui veuillent dire tout ce qu'ils croient »²¹. Fort heureusement, le

¹⁶ Tel est l'un des premiers points dont on averti un enquêteur lors de sa formation dans un institut de sondage. Du moins, c'est la première chose que j'ai apprise lors de ma propre formation de sondeur (CREDOC, Paris, 1996).

¹⁷ L'enjeu de cette stratégie étant d'éviter l'insupportable « dissonance cognitive » théorisée par L. Festinger (*A Theory of Cognitive Dissonance*, Stanford, SUP, 1985).

¹⁸ Je ne soutiens pas ici que cette raison soit déterminante, ni ne nie qu'il ne s'agit là que d'une « intuition de salon ».

¹⁹ *Idem*.

²⁰ Je suggère seulement ici aux philosophes x de consacrer plus d'un vœu pieux aux sciences sociales. Voir, par exemple, J. Elster, *Explaining Social Behavior. More nuts and Bolts for the Social Sciences*, Cambridge, CUP, 2007.

²¹ Troisième partie : AT VI, p. 23.

« faire semblant » (*pretence*) est aussi une préoccupation importante des philosophes *x*²². Pourquoi ne pas alors agrémenter l'enquête sur les causes et les moyens supposés de cette « prétention » par la question de son sens ? Certes, les sonnettes ne sont totalement étrangères à personne²³, mais le sujet interrogé ne peut-il avoir une autre raison de répondre ce qu'il me répond que le sens (obvie ou réfléchi) de ma question²⁴ ? Peut-être n'y a-t-il pas d'interlocuteur naïf à la fin d'un interrogatoire harassant (comme l'est manifestement un questionnaire *x* pour le profane).

Faute de prêter attention aux stratégies d'énonciation et de réponse²⁵, le philosophe *x* ne pourra que proposer des remarques critiques relatives à l'appréciation du comportement verbal²⁶, et rien sur les « intuitions » mêmes, sauf à assimiler celui-ci et celles-là – mais alors, il semblera impossible d'arriver à une phase constructive de la recherche, ce qui serait dommage.

Entre temps, à vaincre sans péril, on triomphe sans gloire. Les protocoles de la philosophie *x* clochent car, pour être valides, les expériences doivent postuler une certaine transparence et une certaine naïveté comportementaliste, quand les résultats produits visent précisément à souligner l'absence d'une telle transparence : les « intuitions » sur lesquelles l'*x-phi* fait porter sa charge mettent en effet au jour des concepts d'états mentaux complexes ou épais, c'est-à-dire des états mentaux on ne peut plus opaques, qui impliquent d'autres états mentaux plus simples ou fins²⁷. En l'état actuel de l'enquête, donc, rien ne prouve que mon discours soit le reflet de mes intuitions. Et pour le dire, je n'ai pas besoin de me lever de mon fauteuil.

Au reste, selon l'ancien paradigme « intuitif » dont la nouvelle philosophie ne s'est pas encore débarrassée, comme le prouve son ardeur à le combattre, trouver les raisons ultimes des réponses obtenues relève de la gageure. La citation de Descartes proposée à l'instant était scandaleusement tronquée. La voici maintenant dans son en entier. Descartes y réfute de la manière la plus claire qui soit l'idée de transparence à soi-même de l'esprit qu'on a souvent

²² Voir S. Nichols & S. Stich, *Mindreading. An Integrated Account of Pretence, Self-awareness and Understanding other Minds*, Oxford, OUP, 2003, chap. 2.

²³ Voir le succès d'*On Bullshit* d'H.G. Frankfurt (Princeton, PUP, 2005 ; trad. fr. *De l'art de dire des conneries*, Paris, 10/18, 2006).

²⁴ On pourrait aussi se poser la question suivante : Pourquoi suis-je donc si enclin à qualifier l'interrogé de sujet ? – Autrement dit : de quoi ou de qui, de quel ordre, est-il le sujet ?

²⁵ Je pense non seulement à ce que la lecture de Foucault et Derrida pourrait apporter à nos philosophes *x*, mais même aux recherches qui ont fait le succès de D.C. Dennett, notamment *The Intentional Stance*, Cambridge, Mass., MIT Press/Bradford, 1987 – que P. Engel avait traduit sous le titre : *La stratégie de l'interprète* (Paris, Gallimard, 1990).

²⁶ Sur cette notion, voir B.F. Skinner, *Verbal Behavior*, New York, Appleton-Century-Crofts, 1957.

²⁷ Voir les objets de l'étude de S. Nichols & S. Stich : *Mindreading, op. cit.*, chap. 3 à 5. Voir également, dans un genre beaucoup plus classique, Descartes, *Regulae ad directionem ingenii* (*Règles pour la direction de l'esprit*), III à VI.

X. Kieft. Les épouvantails transparents n'existent pas

attribuée au *cartesian mind* et que la philosophie *x* critique encore. Ce faisant, il nous rappelle que les épouvantails transparents n'existent pas. « Pour savoir quelles étaient véritablement leurs opinions, je devais plutôt prendre garde à ce qu'ils pratiquaient qu'à ce qu'ils disaient ; non seulement à cause qu'en la corruption de nos mœurs il y a peu de gens qui veulent dire tout ce qu'ils croient, mais aussi à cause que plusieurs l'ignorent eux-mêmes ; car l'action de la pensée par laquelle on croit une chose, étant différente de celle par laquelle on connaît qu'on la croit, elles sont souvent l'une sans l'autre ».

QU'EST-CE QUE LA PHILOSOPHIE EXPÉRIMENTALE ?¹

Joshua Knobe
(Yale)

Joshua Knobe est professeur assistant à l'Université Yale au département de philosophie et pour le cursus de sciences cognitives. Pionnier de la philosophie expérimentale, ses travaux s'attachent à aborder les problèmes philosophiques à l'aide de méthodes expérimentales empruntées aux sciences cognitives. Récemment, comme il en est discuté dans cet article, il montra que les intuitions que forment les gens au sujet de l'intentionnalité d'une action peuvent être biaisées par des considérations morales – effet de « l'effet secondaire », parfois aussi appelé « effet Knobe ».

Depuis les premiers jours de la philosophie analytique, une pratique commune a été d'en appeler aux intuitions pour des cas particuliers. De manière caractéristique, le philosophe présente une situation hypothétique, puis fait une supposition de la forme : « Dans ce cas, nous dirions sûrement que... ». Cette supposition concernant les intuitions des gens constitue une partie d'une démonstration relative à une théorie un peu plus générale sur la nature de nos concepts ou de notre usage du langage.

Un aspect déconcertant de cette pratique est qu'elle fait rarement l'usage de méthodes empiriques standard. Bien que des philosophes fassent assez fréquemment des suppositions concernant « ce que les gens diraient ordinairement », ils les fondent rarement en *interrogeant* réellement les gens et en cherchant à cerner des types dans leurs réponses. Cependant, ces dernières années, un certain nombre de philosophes ont essayé de confronter à des tests ces suppositions concernant les intuitions en utilisant des méthodes expérimentales afin de déterminer ce que pensent vraiment les gens au sujet de certaines hypothèses. Parfois, les résultats furent extrêmement surprenants.

Nous traiterons ici des applications de cette nouvelle méthodologie dans trois domaines de la philosophie : la philosophie du langage, la théorie de l'action et la querelle du libre arbitre.

LA PHILOSOPHIE DU LANGAGE

L'un des recours à l'intuition les plus marquants de la philosophie analytique récente a été l'histoire de Gödel et de Schmidt, imaginée par Saul Kripke :

¹ Traduction par Raphaël Verchère et Luc-Etienne de Boyer des Roches de l'article « What is Experimental Philosophy? » de Joshua Knobe disponible en version originale sur Internet à l'adresse suivante : <http://www.unc.edu/~knobe/ExperimentalPhilosophy.pdf>

Supposez qu'en fait Gödel n'ait pas été l'auteur du théorème [dit « de Gödel »]. Un homme appelé « Schmidt » ... aurait en réalité été l'auteur du travail en question. Son ami Gödel aurait obtenu d'une manière ou d'une autre le manuscrit qui lui aurait été par la suite attribué. Selon le point de vue adopté [la théorie « descriptiviste », qui était alors populaire], lorsque l'homme lambda utilise le nom « Gödel », il veut en réalité faire référence à Schmidt, car Schmidt est l'unique personne satisfaisant la description « l'homme qui a découvert l'incomplétude de l'arithmétique ».²

Pourtant, les lecteurs de cette histoire s'accordaient presque tous à dire que le terme « Gödel » n'avait pas comme référent Schmidt. Toute théorie qui déclarait que Schmidt était le référent de « Gödel » fut par conséquent considérée comme incorrecte.

Cependant, une question se posa de savoir si *tout le monde* partageait cette intuition, ou si elle ne l'était que par les catégories de personnes habituées à lire la philosophie anglo-américaine. Les philosophes Edouard Machery, Ron Mallon, Shaun Nichols et Stephen Stich ont récemment conduit une étude empirique pour résoudre cette question. On raconta à tous les sujets l'histoire de Gödel et Schmidt. Mais le protocole de l'étude incluait un ressort surprenant. Certains sujets étaient Américains, d'autres étaient des habitants de Hong Kong. Comme on pouvait s'y attendre, les sujets américains partageaient les intuitions de la plupart des philosophes analytiques. Mais les sujets de Hong Kong présentaient un type de réponses assez différent. Parmi eux, la majorité disait que le mot « Gödel » avait en fait comme référent Schmidt.

Ce récent résultat (ainsi que des résultats similaires provenant d'études en éthique et en épistémologie) suggère que les Asiatiques pourraient ne pas partager un grand nombre des intuitions sur lesquelles des théories philosophiques largement partagées sont basées.

THÉORIE DE L'ACTION

Les gens font généralement un distinguo entre les actions accomplies *intentionnellement* (par exemple, prendre un verre de vin) et celles accomplies *inintentionnellement* (renverser ce verre de vin sur la chemise de quelqu'un). Toutefois, il s'avère assez difficile de dire précisément en quoi ce distinguo consiste. Les gens ont souvent des intuitions claires quand il s'agit de savoir si une action particulière peut être tenue ou non pour « intentionnelle », mais l'aspect particulier de l'action qui décide de ces intuitions est souvent obscur.

Dans une série d'expériences récente, mon équipe et moi-

² Saul Kripke, *Naming and necessity*, Cambridge, MA, Harvard University Press, 1972/1980, pp. 83-84.

même avons montré que les intuitions des gens peuvent en fait être affectées par les qualités *morales* de l'action elle-même. Nous avons fabriqué des couples de situations qui étaient très similaires à bien des égards mais qui différaient quant à leur signification morale. Dans certains cas, les sujets étaient beaucoup plus disposés à dire qu'une action était accomplie intentionnellement lorsque celle-ci était moralement mauvaise que lorsqu'elle était moralement bonne.

Ainsi, considérons par exemple la situation suivante :

Le vice-président d'une compagnie va trouver le président du conseil d'administration et dit : « nous songeons à débiter un nouveau programme. Il nous aidera à augmenter nos profits, mais, en même temps, il détériorera l'environnement. »

Le président répond : « Détériorer l'environnement m'est totalement égal. Je veux juste faire autant de profit que possible. Lancez le nouveau programme. »

Ils lancent le nouveau programme. Et, comme on pouvait s'y attendre, l'environnement est détérioré.

Face à cette situation, la plupart des gens disent que le président a *intentionnellement* endommagé l'environnement.

Mais supposons que l'on remplace le mot « détériorer » par « préserver ». La situation devient alors :

Le vice-président d'une compagnie va trouver le président du conseil d'administration et dit : « nous songeons à débiter un nouveau programme. Il nous aidera à augmenter nos profits, et il préservera également l'environnement. »

Le président répond : « Préserver l'environnement m'est totalement égal. Je veux juste faire autant de profit que possible. Lancez le nouveau programme. »

Ils lancent le nouveau programme. Et, comme on pouvait s'y attendre, l'environnement est préservé.

Face à cette seconde situation, peu de sujets disent que le président a intentionnellement préservé l'environnement.

Ce résultat indique-t-il que des considérations morales jouent en fait un rôle dans notre concept d'action intentionnelle ? Ou cela montre-t-il seulement que nos jugements peuvent parfois être déformés par des sentiments de responsabilité ? Plusieurs théories rivales ont été proposées mais aucun consensus réel n'a encore émergé. En clair, des recherches supplémentaires sont nécessaires.

LA LIBERTÉ DE LA VOLONTÉ

Si tout ce que nous faisons est dans un certain sens déterminé par les lois physiques, de sorte qu'un énorme ordinateur serait en principe capable de prédire la moindre de nos actions, pourrions-nous toujours être moralement responsables des décisions que nous prenons ? Beaucoup de philosophes ont soutenu que la réponse intuitive à cette question est *non*. Robert Kane écrit ainsi :

D'après mon expérience, la plupart des gens ordinaires sont au départ des incompatibilistes spontanés. Ils croient qu'il y a une sorte de conflit entre la liberté et le déterminisme ; l'idée que la liberté et que la responsabilité puissent être compatibles avec le déterminisme leur semble être à première vue un « bourbier³ » (William James) ou « un misérable⁴ subterfuge » (Emmanuel Kant). Les habiles démonstrations des philosophes doivent tirer les gens ordinaires de cet incompatibilisme spontané.⁵

Il est clair que cette supposition au sujet des intuitions des gens est empirique, et il doit être possible de la tester en utilisant des méthodes empiriques standard.

Dans une série d'expériences récentes, c'est justement ce que firent les philosophes Eddy Nahmias, Thomas Nadelhoffer, Jason Turner et Steve Morris. On raconta à des sujets l'histoire suivante :

Imaginez qu'au siècle prochain, on découvre toutes les lois de la nature, et que l'on construise un super-ordinateur capable de déduire, à partir de ces lois et de l'état actuel de toutes les choses du monde, exactement tout ce qui arrivera à l'avenir dans le monde, à n'importe quel moment. Il pourrait tout examiner de la structure du monde, et pourrait entièrement prédire comment il sera avec une exactitude de 100%. Supposez qu'un tel super-ordinateur existe et examine l'état de l'univers à une certaine heure le 25 mars 2150 après J.-C., vingt ans avant que ne naisse Jeremy Hall. L'ordinateur déduit ensuite avec certitude, à partir de ces informations et des lois de la nature, que Jeremy va dévaliser la Fidelity Bank

³ A « *quagmire of evasion* ». Loÿs Moulin traduit par « échappatoire ». Voir William James, « Le dilemme du déterminisme » in *La volonté de croire*, trad. Loÿs Moulin, Paris, Flammarion, p. 167.

⁴ Emmanuel Kant, *Critique de la raison pratique*, trad. Jean-Pierre Fessler, Paris, GF Flammarion, V, 96, p. 211.

⁵ Robert Kane, « Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism », *Journal of Philosophy*, 96, p. 217.

à 18h00 le 26 janvier 2195. Comme toujours, la prédiction du super-ordinateur est correcte. Jeremy dévalise la Banque de la Fidélité à 18h00 le 26 janvier 2195.

On demanda alors aux sujets si Jeremy était moralement blâmable d'avoir dévalisé la banque : 83% des sujets répondirent que oui. Ce résultat remet en question l'opinion répandue selon laquelle les gens considèrent d'ordinaire le déterminisme comme étant incompatible avec la responsabilité morale.

Des résultats similaires furent obtenus dans une étude conduite par Rob Woolfolk, John Doris et John Darley. Ces chercheurs montèrent des scénarios à propos de personnes désirant accomplir une action immorale, et qui sont ensuite placées dans des circonstances dans lesquelles elles ne peuvent que réaliser cette action. (Par exemple, une histoire concerne une personne qui veut tuer quelqu'un et qui est ensuite forcée de tuer cette même personne par de vils terroristes qui lui ont administré une « drogue de la soumission » qui la rend incapable de résister à leurs ordres.) Bien que les sujets de ces expériences savaient que l'agent ne pouvait rien faire d'autre que ce qu'il a fait, la plupart avaient l'impression qu'il était responsable de son action. Une fois de plus, ce résultat semble ébranler le postulat selon lequel les gens sont des « incompatibilistes spontanés ».

CONCLUSION

La philosophie expérimentale est un domaine de recherche relativement nouveau et le plus gros du travail reste encore à faire. Cependant, il apparaît que son champ d'application s'étend extrêmement rapidement. Nous verrons sûrement un grand nombre de résultats surprenants dans les années à venir.

TRAVAUX CITÉS

- KNOBE, J. (2003). Intentional Action and Side Effects in Ordinary Language. *Analysis*, 63, 190-193.
- MACHERY, E., MALLON, R., NICHOLS, S., & STICH, S. (2004). Semantics, Cross-Cultural Style. *Cognition*, 92, B1-B12.
- NAHMIAS, E., NADELHOFFER, T., MORRIS, S., & TURNER, J. (2004). *Surveying Free Will: Folk Intuitions about Free Will and Moral Responsibility*. Unpublished manuscript. Florida State University.
- WOOLFOLK, R., DORIS, J., & DARLEY, J. (2004). *Attribution and Alternate Possibilities: Identification and Situational Constraint as Factors in Moral Cognition*. Unpublished manuscript. Princeton University.

RÉPHA
Revue Étudiante
de Philosophie Analytique

Numéro 2
Printemps 2010