

L'humanité de l'homme bouleversée par les neurosciences ?

Actes du colloque du Réseau Blaise Pascal, avril 2011

CONNAÎTRE

REVUE SEMESTRIELLE ASSOCIÉE AU RÉSEAU BLAISE PASCAL

Cahiers de l'Association Foi et Culture Scientifique

N° 36-37 (numéro double) Décembre 2011

Rédacteur en chef: Dominique GRÉSILLON

Comité de rédaction

Jacques ARSAC, Marie-Claire GROESSENS-VAN DYCK, Marc le MAIRE, Thierry MAGNIN, Jean-Michel MALDAMÉ, Bernard MICHOLLET, Bernard SAUGIER, Remi SENTIS, Christoph THEOBALD

ISSN: 1251-070X

Membre honoraire: Jean LEROY

Comité de gestion Marcelle L'HUILLIER, Christian MALET

Ce numéro : 18 Euros

« Connaître », 13 Rue Amodru, 91190 GIF sur Yvette evry.catholique.fr/IMG/pdf/AFCS_connaitre.pdf 91afcs@orange.fr

ABONNEMENTS (voir encadré en dernière page)

CONNAÎTRE

Cahiers de l'Association Foi et Culture Scientifique Réseau Blaise Pascal

SOMMAIRE

N°36-37, décembre 2011

L'humanité de l'homme bouleversée par les neurosciences ?

Colloque du Réseau Blaise Pascal, 2 et 3 avril 2011, Sainte-Foy-lès-Lyon

Éditorial	Philippe Deterre	4	
Conférences pl	énières		
La naissance des neurosciences sociales	Marc Jeannerod	6	
Neurosciences, déterminisme et libre arbita	re Bernard Feltz	32	
Neurosciences et théologie de la foi-confia	nce Alexandre Ganoczy	48	
Neurosciences + éthique = neuroéthique	Bernard Baertschi	62	
Table ronde P	Philippe Deterre (Modérateur)		
Deux carrefours th	ématiques		
Neurosciences et âme humaine	Jean-Michel Maldamé	110	
Compte-rendu du carrefour	Brice de Malherbe	123	
Neurosciences et psychanalyse : le cas Thé	érèse d'Avila Michel Simon	126	
Compte-rendu du carrefour	Michel Dion	141	
Abonnements, anciens numéros		143	

Éditorial

Philippe Deterre

Des cerveaux et une humanité

L'homme n'est qu'un roseau, le plus faible de la nature, mais c'est un roseau pensant. Il ne faut pas que l'univers s'arme pour l'écraser : une vapeur, une goutte d'eau suffit pour le tuer. Mais, quand l'univers l'écraserait, l'homme serait encore plus noble que ce qui le tue, puisqu'il sait qu'il meurt, et l'avantage que l'univers a sur lui, l'univers n'en sait rien. I

Cette pensée de Blaise Pascal dit bien toute l'ambition du colloque dont vous trouverez ici les Actes. D'abord, parce que le Réseau qui l'a organisé a pris le nom du grand physicien croyant - Réseau Blaise Pascal (Sciences, Cultures et Foi)² – et surtout parce que ce colloque abordait la thématique des sciences du cerveau, cet organe « roseau » s'il en est, puisque protégé uniquement par la fine couche osseuse de la boîte crânienne. Pourtant, il constitue le siège de notre conscience individuelle, de nos théories scientifiques, de toutes nos sagesses y compris religieuses, et de nos convictions les plus hautes, comme celles des « droits de l'homme ».

Ce colloque s'est ouvert sur une conférence de Marc Jeannerod, professeur émérite de physiologie de Lyon, sur la psychologie cognitive de la relation sociale. De nombreuses personnes de notre réseau doivent beaucoup à ce grand scientifique lyonnais, disparu prématurément deux mois après le colloque. En outre, Alexandre Ganoczy, théologien, a présenté dans le même colloque une conférence issue de son dialogue fructueux avec le neurobiologiste. Au milieu d'autres³, la publication de ces Actes voudrait être un hommage de gratitude envers l'homme « noble » profondément chrétien et passionnément scientifique qu'a été Marc Jeannerod.

³ Voir http://l2c2.isc.cnrs.fr/en/jeannerod/?language=en#S7

Connaître | Nº 36-37 | décembre 2011

.

¹ Blaise Pascal *Pensées* n°347, selon la classification de Léon Brunschvicg

² Voir le site de ce réseau : http://sciences-foi-rbp.org

Tous les articles "invités" ainsi que quelques "présentations de carrefours" appartenant au programme du colloque sont publiés dans ce numéro double 36-37. Cependant quelques communications ordinaires et des présentations de carrefours qui n'ont pu trouver place dans le volume limité de ce numéro ont été publiées dans le numéro 35.

Le programme du colloque a été construit par l'équipe de coordination du Réseau Blaise Pascal, et en particulier par Christiane Trimoulet, Mireille Leduc, Claude Béguin, Pierre Bourdon, Philippe Deterre, Michel Dion, Jean-Michel Maldamé, Bernard Michollet, Fabien Revol, Bernard Saugier, Rémi Sentis. La publication du colloque a bénéficié d'une subvention de la Fondation Templeton.

La naissance des neurosciences sociales

Marc Jeannerod †,1

La psychologie cognitive comporte depuis ses débuts un volet « social », ce terme étant entendu dans le sens d'une psychologie prenant en compte les relations entre les individus. Le problème des relations interindividuelles a longtemps été abordé sous l'angle des théories psychanalytiques. Ce modèle, qui a été détaillé dans mon ouvrage "*Le cerveau volontaire*"², donne un rôle important, dans le développement des processus qui aboutissent au sentiment de confiance, aux relations parentales. Les relations parentales (où se distinguent les rôles respectifs de la mère et du père) doivent être envisagées dans le cadre d'une société fondée sur la notion de famille, notion dominante sinon unique dans les sociétés développées comme les sociétés occidentales, et dont on retrouve la trace dans la plupart des autres sociétés connues.

Comme le décrit bien le modèle psychanalytique, la fonction principale des parents est de déterminer puis de délimiter l'environnement affectif du petit enfant. L'enfant et ses parents forment une cellule où se tissent des relations affectives spécifiques, sur le mode de la triangulation. Les parents ont une individualité propre qui conditionne leur situation dans le triangle : c'est cette situation qui change selon les formes familiales qui se sont imposées dans les différentes sociétés, tandis que l'organisation triangulaire de base reste invariante d'un type social à l'autre. Si l'on s'en tient à ce modèle, les parents sont donc des individus socialement identifiés par leur comportement vis-à-vis de l'enfant, comportement de type « maternel » ou de type « paternel », même si leur rôle étroitement biologique n'intervient pas dans cette identification. Dans de nombreuses sociétés, la répartition de ces rôles entre les membres du groupe ne tient en effet pas compte de la fonction

_

¹ Institut des sciences cognitives, Lyon. Depuis la présentation de cet exposé, le Professeur Marc Jeannerod est décédé à Lyon le 1^{er}, juillet 2011.

² M. Jeannerod, *Le cerveau volontaire*, Paris, Odile Jacob, 2009.

de géniteur. Le statut de « parent » donné à tel ou tel membre du groupe au sein du noyau familial est un statut de type social plutôt que biologique.

Dans cet article, nous décrivons une autre approche des relations interindividuelles, celle de la psychologie cognitive. Cette approche envisage le problème des relations interindividuelles à un niveau purement objectif où les individus ne sont pas identifiés par leur rôle social mais par leur position dans un champ d'interaction. Autrement dit, le problème est abordé au niveau de la distinction entre « soi » et un autre indifférencié. Cette approche se distingue donc de l'approche psychanalytique par son caractère plus fondamental, au sens où elle étudie les relations en quelque sorte « avant » l'affectivité, telles qu'elles sont déterminées par le rôle que jouent les individus en tant qu'agents dans une situation de communication. A l'inverse de la psychanalyse qui laisse le champ libre au discours sur les sentiments, la psychologie cognitive représente une forme minimaliste de l'analyse des relations interindividuelles. Nous verrons cependant dans la suite comment elle permet aussi d'aborder la question de l'affectivité dans ces relations.

Il est important de comprendre la confiance à partir d'une reconstruction de ses mécanismes de base, y compris neurologiques. La psychologie cognitive permet d'envisager un point de vue expérimental, c'està-dire à partir de situations où les différents facteurs de ces relations peuvent être systématiquement manipulés. Cette nouveauté a été à l'origine de grands progrès dans la compréhension des relations entre individus, en ouvrant la voie aux études portant sur les mécanismes cérébraux sous-jacents à ces relations et sur certaines de leurs perturbations pathologiques. C'est récemment, au cours des vingt dernières années, que le problème des relations interindividuelles est devenu un thème de recherche inscrit dans le programme des neurosciences. Cette évolution parait logique, puisque, comme dans d'autres domaines, les concepts et les méthodes de la psychologie cognitive ont constitué la base des paradigmes expérimentaux des neurosciences. On parle alors, dans le cas qui nous intéresse ici, de « neurosciences sociales », pour lesquelles il existe depuis peu une revue (Social Neuroscience), et une société savante (Society for Social Neuroscience) basée aux Etats-Unis. Un

autre avantage de l'approche de la psychologie cognitive, qui ne sera cependant pas évoqué ici, est qu'elle n'est pas limitée aux relations sociales chez l'homme, mais peut être étendue au monde animal dans son ensemble.

1. Communication directe et communication indirecte

La notion de confiance et son corollaire, la notion de méfiance, qui font l'objet de cette étude, sont des notions éminemment sociales, puisqu'elles impliquent une interaction entre deux pôles, le pôle où se situe celui qui fait (ou ne fait pas) confiance, et le pôle où se situe celui qui inspire (ou n'inspire pas) confiance. Les situations expérimentales qui se proposent d'étudier les bases psychologiques et biologiques de la confiance doivent donc tenir compte de cette dualité. Le problème se pose cependant de manière différente selon que la relation entre les deux pôles est directe ou indirecte. La relation directe suppose la présence physique des deux individus en interaction : les signaux émis par un des protagonistes sont perceptibles par l'autre et l'analyse de ces signaux sociaux donne accès aux mécanismes qui permettent de comprendre les émotions, les actions et les états d'esprit de l'autre. La relation directe comporte donc la notion de réciprocité (la relation de confiance entre deux collaborateurs, entre un malade et son médecin). La relation indirecte, au contraire, est à sens unique, au moins dans l'immédiat. Elle peut s'exercer au travers de médias, vis-à-vis d'un tiers absent ou d'une entité abstraite, par exemple dans le cadre d'une enquête de confiance comportant des questions du type : « Le Président de la République vous inspire-t-il confiance ? ». Dans ce cas, le caractère impersonnel du tiers implique le caractère unidirectionnel de la relation

Cette distinction entre relation directe et relation indirecte recouvre en fait une complémentarité entre deux modalités de la communication entre les individus. D'un côté, une conception centrée sur le corps et ses manifestations physiologiques, celle d'une communication « par le corps » (embodied), fortement automatisée, et fondée sur des signaux physiologiques issus du corps des individus impliqués dans l'acte de communication ; de l'autre, une conception centrée sur des mécanismes cognitifs, qui relève d'une approche

proprement « mentaliste », opérant par l'intermédiaire de processus mentaux sans passer par des signaux corporels. Dans ce chapitre, nous envisageons successivement ces deux modalités, tout en faisant apparaître leur caractère complémentaire : la communication directe par le corps est en effet souvent le point de départ d'une communication plus mentalisée et indirecte, comme nous le verrons à propos de certaines formes de psychothérapie.

Les neurosciences cognitives (dans leur version sociale) tendent à privilégier la thèse d'une communication s'exprimant par l'intermédiaire du corps. Le corps émet et reçoit des signaux qui ont une valeur sociale, et le système nerveux semble posséder une organisation qui permet de décoder efficacement ces signaux : la façon dont ces signaux sont traités par le système nerveux permet de décrire des systèmes spécialisés, des réseaux nerveux responsables du comportement de communication entre individus. Même si cette capacité est universelle dans le monde animal, elle atteint chez les primates et particulièrement chez l'homme un haut degré de développement qui permet la communication de la signification d'actions, d'émotions et d'états d'esprit.

La communication par le corps et le rôle des signaux sociaux

Depuis Charles Darwin, de nombreux travaux ont permis de décrire les signaux qui transmettent aux congénères des informations sur l'état interne (mental) de celui qui les émet. Expression faciale, posture, démarche, intonation de la voix, constituent un répertoire commun à tous les individus d'une même espèce. Darwin a insisté sur le caractère spécifique de ces signaux, qui permet à chaque individu d'exprimer de façon univoque ses émotions et ses sentiments. Il existe sur ce sujet des descriptions exhaustives des composantes de chaque type d'émotion, aussi bien en termes des muscles de la face qui y sont impliqués (Paul Ekman), que des corrélats végétatifs qui les accompagnent (Scherrer). Cette forme de communication ne s'applique évidemment qu'aux informations effectivement exprimées, ce qui exclut les états internes ne comportant pas ce type d'information, les idées ou les savoirs, par exemple.

Le corollaire de cette communauté d'expression d'une émotion ou d'un sentiment est la communauté de leur reconnaissance par les congénères. C'est là en effet que se situe le problème de la relation : non pas seulement dans l'élaboration des signaux correspondant à un état mental donné, mais dans la reconnaissance de ces signaux par les autres individus. Les expressions faciales, pour prendre cet exemple, sont reconnaissables : nous sommes capables de détecter sur un visage l'expression de la peur, de la surprise, de la tristesse, de la joie, etc. Dans la ligne de ce qui vient d'être dit sur le répertoire de signaux sociaux dont dispose une espèce, il existe également une importante littérature sur la perception de ces signaux et leur contribution à la reconnaissance des émotions. Chez le singe, l'étude de neurones d'une région du cortex visuel (le cortex inféro-temporal) démontre l'existence d'un mécanisme de codage des expressions faciales (mouvements et postures des lèvres, des veux, etc.) ainsi que des postures de la tête (Perrett). Chez l'homme, la région homologue du sillon temporal supérieur est le siège d'activations importantes, repérées par les techniques de neuroimagerie, lors de la présentation de visages porteurs d'expressions émotionnelles. Une autre région du cortex visuel dans l'aire occipitale latérale est activée de manière spécifique lors de la présentation de parties du corps, immobiles ou en mouvement (extra-striate body area, EBA).

Il existe donc une correspondance étroite entre expression et reconnaissance : je reconnais l'expression de la joie sur le visage d'un autre, alors même que je ne connais pas ce qu'exprime mon propre visage ; mais ma façon d'exprimer la joie est immédiatement reconnue par l'autre. (Darwin expliquait cette correspondance par des dispositions innées : la synergie musculaire de la face correspondant à chaque émotion serait innée, comme le serait la reconnaissance de cette synergie). Une des raisons de cette intercommunicabilité des signaux sociaux est leur origine « biologique », qui les distingue des autres stimuli visuels : les signaux sociaux sont « intentionnels » au sens où ils se réfèrent à un auteur et où ils sont orientés vers un but déterminé. Leur forme reflète les états mentaux – intentions, désirs, croyances – que leur auteur construit dans le cadre d'une action ou d'un acte de

communication. C'est ainsi que les mouvements qui composent les actions intentionnelles ont une cinématique particulière qui les distingue des déplacements occasionnés par des causes mécaniques. Outre le fait que la perception de ces signaux relève de systèmes neuronaux spécifiques, leur statut particulier est attesté par de nombreuses observations. Ainsi, au cours du développement, les visages et leurs expressions (sourires) sont reconnus de manière très précoce. A l'inverse, certaines lésions cérébrales affectent la reconnaissance des visages tout en épargnant la reconnaissance des autres stimuli.

Le développement de ce mode de communication par l'intermédiaire des signaux sociaux mérite un intérêt particulier. Quelle est l'origine de ces mécanismes de la perception des signaux biologiques intervenant dans la communication entre individus? L'observation du comportement de nouveaunés et d'enfants en bas âge apporte des éléments de réponse à cette question. En effet, la précocité des réponses à ce genre de signaux est telle qu'elle semble exclure l'acquisition de ces mécanismes à la suite d'un apprentissage. Comment peut-on poser ce genre de questions à un nouveau-né de quelques jours? Plusieurs méthodes ont été mises au point, dont la méthode dite de « succion non nutritive ». Chacun a pu remarquer qu'un bébé en train de boire son biberon tête plus ou moins rapidement d'un moment à un autre. D'une manière générale, il tête plus rapidement lorsqu'il est intéressé par les événements qui se déroulent autour de lui. Ce comportement peut donc servir d'indice pour mesurer l'intérêt ou le désintérêt qu'un bébé prête à une stimulation, une voix ou un visage, par exemple. Dans ce genre d'expérience, on donne à sucer au bébé une tétine vide (non nutritive) reliée à un dispositif d'enregistrement. Cette méthode a permis de déterminer qu'un enfant de quelques jours est capable de reconnaître la voix de sa propre mère par opposition à la voix d'une mère étrangère : la succion augmente lorsque la voix de la mère est présentée. L'indice que l'enfant utilise pour reconnaître cette voix est l'intonation : en effet, le fait de présenter la voix maternelle en faisant dérouler la bande magnétique à l'envers, procédé qui modifie profondément l'intonation, empêche toute reconnaissance. Ce résultat

s'explique par la longue exposition de l'enfant à la voix maternelle au cours de son séjour dans l'utérus : le seul aspect de la voix qui soit reconnaissable après la traversée du liquide amniotique et des tissus maternels est l'intonation. D'autres résultats, encore plus intéressants peuvent être obtenus à l'aide de cette méthode de succion non nutritive. On peut ainsi montrer que l'enfant de guelques jours préfère les sons de langage à toute autre sorte de bruit. Surtout, il est capable de déterminer si l'intonation d'un énoncé est correcte : il marquera son étonnement (en augmentant la fréquence des mouvements de succion) si une pause est marquée dans une phrase au mauvais endroit. Cela ne signifie évidemment pas que le bébé a une quelconque connaissance linguistique; cela signifie qu'il dispose en naissant d'un « organe » prêt à fonctionner pour reconnaître les sons de langage et pour leur appliquer un certain nombre de règles élémentaires. C'est à partir de ces règles qu'il pourra progressivement organiser les sons de langage en syllabes, en mots et en phrases, avant d'être capable de les produire lui-même. Ainsi le langage, signal et social par excellence, est-il enraciné dans le cerveau humain dès la naissance. Cet exemple du traitement précoce du langage ouvre sur la question de la cognition néonatale dans son ensemble. A ce stade, la performance est sans doute limitée par le degré de maturation des systèmes cognitifs : dans le cadre de la communication, l'élément essentiel que constitue la Théorie de l'esprit ne se met en place qu'à partir de la seconde ou de la troisième année. Dans l'intervalle, le bébé, pour émettre comme pour recevoir, utilise la gamme des signaux biologiques.

Représentations partagées et empathie

Ces propriétés particulières des signaux sociaux leur confèrent un caractère biface : le signal perçu se calque directement sur le signal exprimé. De même que l'expression correspond à un état affectif, la perception de cette expression ne s'arrête pas au stade de son identification : l'émotion exprimée par un congénère est non seulement reconnue, elle est également immédiatement éprouvée par celui qui l'observe. Un visage joyeux déclenche chez celui qui le voit un sentiment de joie, un visage triste déclenche un

sentiment de tristesse (qui peut, chez l'enfant aller jusqu'à produire des pleurs ou d'autres manifestations émotionnelles de la tristesse). Ces observations démontrent le degré de « pénétration » du signal émotionnel (et, comme nous le verrons, des signaux sociaux en général) dans le système nerveux de l'observateur. On peut en donner plusieurs exemples à propos de deux émotions caractéristiques, le dégoût et la peur.

Dans le cas du dégoût, des modifications de l'activité cérébrale sont enregistrées dans la partie antérieure du cortex de l'insula dans trois conditions : lorsque le sujet est exposé à une odeur dégoûtante, lorsqu'on lui montre l'image d'un objet répugnant, et enfin (et surtout) lorsqu'on lui montre l'image d'un visage exprimant le dégoût (Wicker et al, 2003). Cette empathie pour le dégoût de l'autre, s'observe aussi pour la peur. Dans ce cas, en effet, les novaux amygdaliens sont activés aussi bien chez celui qui a peur que chez celui qui observe un visage exprimant la peur. La lésion bilatérale de ces mêmes novaux provoque d'ailleurs à la fois l'impossibilité d'exprimer la peur et l'impossibilité de reconnaître l'expression de peur sur un visage. On pourrait ajouter à ces exemples, celui de la douleur, dont il sera question plus loin: l'activation enregistrée chez un sujet lorsqu'on lui applique une stimulation tactile douloureuse est semblable à celle qui est provoquée, chez le même sujet, par la vision de quelqu'un d'autre qui reçoit la stimulation douloureuse (Keysers et al, 2004). Dans tous ces cas, le signal social (l'expression du visage de celui qui éprouve l'émotion) joue le rôle d'un stimulus qui déclenche l'activation des zones cérébrales correspondant à cette émotion. Cette activation, toutefois, ne représente que la partie la plus automatique de la « réponse » de l'observateur à la situation, point sur lequel nous reviendrons lorsqu'il sera question de l'empathie.

Ces données peuvent être interprétées en considérant la notion de « représentations partagées ». Le réseau nerveux qui code l'expression du dégoût, par exemple, et celui qui code sa reconnaissance sont en grande partie superposés : c'est le cas chez les deux protagonistes impliqués dans l'expérience ci-dessus, où l'un des deux exprime son dégoût devant un stimulus visuel qui lui inspire le dégoût et où l'autre qui l'observe reconnaît

son expression et éprouve à son tour du dégoût. Si, comme le montrent les données d'IRM, les zones impliquées ont la même répartition topographique chez les deux protagonistes, on peut dire qu'ils partagent la même représentation et peuvent, du même coup, éprouver le même état affectif. Cette notion de superposition et de partage a été enrichie par la découverte, dans le cortex prémoteur du singe, d'une catégorie particulière de neurones, les neurones miroir. Un neurone miroir est typiquement actif dans deux situations : lorsque l'animal exécute un mouvement (de la main, par exemple, lors de la saisie d'un morceau de nourriture), et lorsqu'il observe le même mouvement exécuté devant lui par un congénère ou un expérimentateur. Le même neurone est donc impliqué à la fois dans l'exécution et la perception du même acte moteur. Sa décharge représente le même acte, considéré de deux points de vue différents : le point de vue « en première personne » lors de son exécution (« c'est moi qui le fais »), et le point de vue « en troisième personne » lors de son observation (« c'est lui qui le fait »). Les neurones miroir constituent donc un modèle de représentation partagée, directement applicable dans le cas de la représentation d'une action : « je comprends ce que fait l'autre parce que son action occupe dans mon cerveau le même espace que si je le faisais moi-même (et réciproquement) ». Ce modèle peut être aisément généralisé à la compréhension des émotions qui représentent un cas particulier des actions. Peut-on franchir un pas de plus, et le généraliser aux états mentaux en relation avec les actions ou les émotions? Il a été proposé que les neurones miroir codent, non seulement l'action elle-même, mais le but vers lequel elle est dirigée, c'est-à-dire l'intention de son auteur. Cette interprétation, toutefois, a ses limites : ainsi, le même mouvement (atteindre un objet) peut faire partie de plusieurs actions différentes sur ce même objet (le prendre, le heurter pour le faire tomber, etc.). Comment l'observation du mouvement d'atteinte pourrait-elle déterminer, à elle seule, l'intention finale qu'il contient? Il est nécessaire pour cela de disposer d'informations sur le contexte affectif et cognitif dans lequel se déroule l'action, informations dont les neurones miroir ne disposent en principe pas : le neurone miroir serait donc, par définition, un neurone « empathique »

(Discussion in Jacob et Jeannerod, 2003). Pour reprendre la distinction faite ci-dessus entre relation directe et relation indirecte entre individus, les neurones miroir seraient concernés seulement par la relation directe, c'est-à-dire par la présence de signaux sociaux directement perceptibles et interprétables sur l'action, l'émotion ou l'état mental/affectif du protagoniste de la relation. D'autres mécanismes devront être envisagés pour rendre compte de la reconnaissance et le partage d'états mentaux pour lesquels ces signaux sont absents. Ce sera le but de la section de ce chapitre consacrée à la Théorie de l'esprit (voir plus loin).

Un concept proche de celui du partage de représentations est le concept d'empathie. Pour Théodor Lipps, un de ceux qui l'ont le mieux développé à la fin du XIXème siècle, l'empathie (Einfühlung) constitue un des mécanismes de base de la connaissance de soi-même et des autres. Nous ne pouvons en effet, pensait Lipps, comprendre les autres par la voie de la perception, qui ne peut nous révéler leurs sentiments, leurs intentions ou leurs désirs; ni en procédant par analogie, en projetant sur eux nos propres expressions (que nous ne percevons d'ailleurs pas). Pour lui, le fait de voir une expression sur un visage réveillerait automatiquement les influx nécessaires à la production de cette expression, et ces influx induiraient l'état affectif interne correspondant à cette expression. Autrement dit, les impulsions motrices induites par la vue de l'expression sur le visage de l'autre incluent la tendance à ressentir cet état affectif. La vision de l'expression correspondrait à un « début d'imitation », une « imitation interne ». C'est par ce moyen, pensait Lipps, que nous devenons conscients de l'existence des autres. Freud avait fréquemment fait appel à l'Einfühlung, qu'il utilisait dans le sens de « se mettre à la place de l'autre et tenter de le comprendre ». « Avec la perception d'un geste déterminé, disait-il dans "Le mot d'esprit et sa relation avec l'inconscient", est donnée l'impulsion de le représenter par une certaine dépense. Ainsi donc, en accomplissant l'acte de 'vouloir comprendre' ce geste, d'en avoir l'aperception, je me comporte [...] tout à fait comme si je me mettais à la place de la personne observée ». Dans le cas d'un geste comique, c'est la comparaison entre le geste observé et la représentation du mien propre (que j'aurais moi-même accompli à sa place) qui est source de comique et déclenche le rire : « L'origine du plaisir comique [...] provient de la comparaison entre l'autre personne et notre propre moi – c'est à dire de la différence quantitative entre la dépense d'empathie et la dépense propre... ». Plusieurs exemples très explicites sont donnés sur cette notion de différence quantitative : c'est ainsi que la différence entre un mouvement attendu par un spectateur et ce mouvement effectivement réalisé par l'acteur peut être source de comique, quand, par exemple, « je prends dans une corbeille un fruit que je crois lourd mais qui, pour me tromper, est un objet creux, une imitation en cire. En partant en l'air, ma main trahit que j'avais préparé une innervation trop grande pour la fin visée, et c'est pour cela qu'on rit de moi » (Freud).

Max Scheler, dans son livre sur la sympathie, critique la position de Lipps en soulignant que l'Einfühlung (traduit dans le texte français de ce livre par « intuition projective »), ne peut rendre compte de la participation aux états affectifs d'autrui. Elle peut servir à concevoir ou à comprendre ces états affectifs, mais non à les éprouver. Selon Scheler, il faut, pour franchir cette étape, un mécanisme préalable de connaissance des relations entre le moi et ses sentiments. Cette relation, nous ne pourrons jamais la saisir directement chez autrui, mais toujours à travers l'aspect qu'elle prend chez nous-mêmes : étant donné le lien intime qui existe entre notre moi et nos sentiments, nous pouvons également admettre qu'autrui possède un lien du même type. Cette formulation, très proche de celle que nous examinerons à l'occasion de la Théorie de l'Esprit, incite à distinguer plusieurs niveaux et plusieurs formes au sein du phénomène d'empathie. Le concept d'empathie a en effet été redéfini et précisé dans le cadre des neurosciences sociales. Une distinction semble devoir s'imposer entre une empathie de bas niveau, caractérisée par sa survenue automatique et involontaire, et une empathie contrôlée selon le contexte et le type de relation au cours de laquelle elle se manifeste. L'empathie automatique prend souvent la forme d'un mimétisme où le comportement de l'observateur copie celui de l'observé : on peut ranger dans cette catégorie les phénomènes de résonance motrice observés chez les animaux vivant en groupe, la contagion des émotions (fou-rire, bayement),

l'imitation dans sa forme compulsive. A ce niveau, l'observateur se trouve placé dans la perspective de celui qu'il observe et tend à épouser son état moteur ou émotionnel, l'empathie aboutissant alors, comme semblaient l'envisager certains des auteurs classiques évoqués plus haut, à une fusion indifférenciée des protagonistes au sein d'un état commun. L'empathie contrôlée, au contraire, est sélective : elle cherche à éviter la fusion et à préserver l'individualité des protagonistes : celle de l'observateur qui se maintient à distance de l'observé, et celle de l'observé qui est l'objet, de la part de l'observateur, d'une relation choisie et adaptée aux circonstances, plutôt que d'une réponse automatique et incontrôlée. L'empathie sous contrôle pourrait ainsi correspondre à la sympathie telle que la définit Scheler, en associant le mécanisme de partage de représentations à celui de la reconnaissance de l'autre en tant qu'individu différencié.

Une forme de psychothérapie, celle élaborée par Carl Rogers, est explicitement fondée sur cette possibilité. Selon cette méthode, le thérapeute doit utiliser la « compréhension empathique » pour percevoir la situation émotionnelle du patient, tout en conservant sur « inconditionnellement positif » et en faisant preuve d'une « acceptation chaleureuse » de son expérience. Rogers insistait sur le fait que la relation doit être réciproque : le patient doit savoir qu'il est l'objet d'une compréhension empathique et, en retour, le thérapeute se guide sur la réponse qu'il reçoit du patient (voir Bozarth, 2009). Cette notion de réciprocité, où l'estime de soi se trouve renforcée par l'estime témoignée par les autres, fait partie intégrante de nombreux systèmes théoriques en psychologie et en pédagogie. Le but est la « mise en confiance » du sujet par la participation active, la valorisation de ses émotions positives (joie, fierté, contentement, résilience), le développement de son sentiment de créativité (voir Dolto, Bandura). Nous reviendrons sur ce point ultérieurement.

L'existence de niveaux différents d'empathie incite à se demander s'ils ne correspondraient pas, en fait, à différents stades de l'établissement d'une relation affective entre individus. Dans un premier temps, la réponse immédiate et automatique permettrait, par réplication, de détecter l'existence de l'état affectif de celui qu'on observe ; dans un second temps, interviendrait le mécanisme de contrôle et de sélection qui permettrait la différenciation entre soi et l'autre. Cette hypothèse, défendue par ailleurs par des auteurs comme Goldman (2006) ou Wegner (Hodges et Wegner, 1997), semble être accréditée par les données récentes des neurosciences qui révèlent que l'anatomie fonctionnelle de l'empathie n'est en effet pas la même selon qu'il s'agit du processus automatique ou du processus contrôlé. La description de ces données sera faite ici à partir d'un exemple précis, celui de l'empathie pour la douleur.

Une stimulation douloureuse provoque chez celui qui est stimulé l'activation d'un réseau complexe comprenant, d'une part, les zones sensorielles situées dans la région du cortex pariétal du côté opposé au côté stimulé, et d'autre part, les zones codant l'aspect affectif du stimulus dans le cortex pariéto-insulaire et la région cingulaire antérieure des deux côtés. Ces deux sous-réseaux correspondent à des composantes différentes de l'expérience douloureuse, et peuvent d'ailleurs être dissociés l'un de l'autre : dans une condition pathologique appelée « asymbolie à la douleur », seul le réseau sensoriel est activé. C'est dire qu'on peut « ressentir » une douleur, la localiser à la partie du corps concernée, en percevoir les qualités perceptives sans l'« éprouver », c'est-à-dire sans être affecté par l'émotion négative qui accompagne normalement ce type de stimulus. On se rappellera que cette dissociation a servi de justification, en leur temps, aux tentatives de « psychochirurgie » de la douleur mises au point pour pallier les effets de douleurs chroniques dues à des maladies incurables. La lobotomie frontale était alors supposée déconnecter le réseau sensoriel du réseau affectif que l'on localisait déjà, non sans raison, dans les lobes frontaux : le patient était ainsi censé devenir « indifférent » à sa douleur.

Qu'en est-il pour un observateur qui assiste au spectacle de la douleur chez un autre ? S'il ne peut, en toute logique, ressentir la douleur de l'autre, puisqu'il n'est pas lui-même directement atteint par le stimulus douloureux, il peut toutefois éprouver la composante émotionnelle qui fait partie de cette expérience douloureuse. Puisque cette composante, comme nous l'avons vu,

peut être détachée de la composante sensorielle, elle peut donc en principe être éprouvée en tant que telle, c'est à dire en l'absence d'un stimulus provoquant la douleur. Les données des neurosciences vont effectivement dans ce sens. Tania Singer, enregistrant l'activité cérébrale chez des sujets, soit soumis à un stimulus douloureux, soit assistant à l'administration du même stimulus à un autre sujet, a montré que le réseau sensoriel n'est activé que dans la première condition. En revanche, le réseau affectif est activé, non seulement lorsque le sujet recoit le stimulus douloureux, mais aussi lorsqu'il observe ou anticipe la stimulation infligée à l'autre (Singer et al. 2004). La même activation du réseau affectif est également enregistrée lors de la présentation de photos de visages exprimant la douleur (Botwinick et al, 2005). Le fait que, dans ces expériences, l'activation soit enregistrée en l'absence de relation particulière entre les protagonistes indique bien que la vision ou l'évocation d'une situation douloureuse entraîne une réponse empathique « spontanée » et incontrôlée. Notons que, pour Scheler, la connaissance des états affectifs d'autrui ne passe pas par un mécanisme de simulation, mais par une perception immédiate et directe qui nous donnerait accès à ce qu'il appelait une « sorte de grammaire universelle, valable pour tous les langages d'expression et fournissant le critère suprême à l'aide duquel nous sommes à même de comprendre tous les genres de mimiques et de pantomime » (p 58). Cette position, qui se rattache somme toute à celle de Darwin, ne tient évidemment pas compte, pour ce qui est de Scheler, des progrès plus récents en matière de neuroscience. Elle est néanmoins défendue par des auteurs contemporains (voir une discussion de ce point dans de Vignemont et Jacob, 2010).

De l'empathie à la préférence sociale

Cette réponse automatique peut cependant être modulée par des influences provenant du contexte dans lequel se déroule l'expérience, ou de l'état mental de l'observateur. Tania Singer et ses collaborateurs, dans l'expérience déjà citée, avaient personnalisé la situation en faisant appel à des sujets vivant en couple et donc supposés éprouver de l'empathie l'un pour

l'autre : dans ce cas, l'activation de l'insula et cortex cingulaire témoignant de la réponse affective était significativement augmentée par rapport à celle qui était enregistrée chez des sujets ne se connaissant pas. Cette modulation de la réponse empathique par des facteurs personnels, si elle n'est pas surprenante en elle-même, ouvre cependant des perspectives nouvelles pour les neurosciences sociales. On pourrait en effet tenter de généraliser ce type de paradigme aux autres modalités de relations entre individus, où la réponse cérébrale deviendrait un test « objectif » (indépendant du sujet) des réponses affectives et émotionnelles qui conditionnent la préférence sociale. On voit ici comment un jugement, donc un acte de discernement intervient pour dépasser le niveau purement émotionnel de l'empathie non contrôlée. La préférence sociale contribue alors à justifier la confiance : on fait confiance « en connaissance de cause ».

Prenons comme exemple le traitement des expressions faciales, dont nous avons vu plus haut l'importance parmi les autres stimuli sociaux. Rappelons que ce traitement implique essentiellement le sillon temporal supérieur, où se fait l'identification consciente de l'expression, ainsi que l'amygdale et le cortex orbito-frontal, où se fait la reconnaissance émotionnelle automatique. Dans l'étude de Winston et al (2002), on demandait à des sujets dont l'activité cérébrale était enregistrée de porter un jugement social sur des photos de visages qui leur étaient présentées. La question posée portait sur le degré de confiance (trustworthiness) que le sujet pouvait accorder à chaque visage (« ce visage vous inspire-t-il confiance ou non? »). A côté de cette tâche de jugement explicite, une autre série de photos de visage était présentée, avec une question portant sur l'âge de la personne représentée et ne concernant donc pas la signification sociale de l'expression, ce qui n'empêche pas que cette expression était néanmoins traitée de façon implicite et automatique. Dans la situation de jugement explicite sur la confiance, l'activité du sillon temporal supérieur était globalement plus élevée que dans la situation où le jugement portait sur l'âge. Quant à l'amygdale, elle était activée dans les deux situations, avec un niveau plus élevé pour les visages n'inspirant pas confiance. En résumé, la réponse cérébrale

émotionnelle (celle de l'amygdale) provoquée par la vision d'une expression faciale est la même, qu'il s'agisse d'un jugement implicite sur la confiance qu'inspire ce visage, ou d'une réponse automatique ne comportant pas de jugement.

Cette réponse implicite et automatique est-elle modulable ? Singer et ses collaborateurs (2006) ont transféré leur paradigme d'observation d'un stimulus douloureux à une situation où les sujets ne se connaissaient initialement pas, mais établissaient une relation au cours de l'expérience. Il s'agissait d'une situation de jeu d'argent où les sujets observés étaient des comparses qui pouvaient être, soit « bons joueurs » en coopérant avec les autres pour favoriser les gains des deux joueurs, soit « mauvais joueurs » en refusant de coopérer. Les observateurs classifiaient rapidement les joueurs en personnes sympathiques (et donc dignes de confiance) ou antipathiques (et suscitant la méfiance) en fonction de leur comportement dans le jeu. Lorsque, dans une seconde phase de l'expérience, des stimuli douloureux (fictifs) étaient appliqués aux joueurs, les modifications de l'activité cérébrale enregistrées chez les observateurs tenaient compte de leur statut de bon ou de mauvais joueurs : une réponse empathique pour la douleur était enregistrée lorsque la douleur était infligée à des comparses sympathiques, mais non à ceux qui avaient refusé de coopérer. Pour ces derniers, au contraire, une réponse était enregistrée dans des régions activées lorsque le sujet reçoit un renforcement positif, régions associées aux sentiments de plaisir, de satisfaction ou de « récompense » (striatum ventral et novau accumbens). Ainsi, la réponse cérébrale devient ici un indicateur des attitudes sociales : la préférence accordée aux personnes fiables et coopératives se marque par une réponse empathique, alors que le rejet suscité par les personnes peu fiables se marque par la satisfaction de les voir punis. La conclusion donnée par Singer et al (2006) à l'ensemble de leurs résultats est intéressante. Selon ces auteurs, ces résultats permettent de donner aux théories sur les préférences sociales une base neurologique. Si, comme le disent ces théories, les gens approuvent les gains acquis par d'autres lorsqu'ils les perçoivent comme des personnes honnêtes, alors qu'ils désapprouvent les mêmes gains acquis par des personnes malhonnêtes, ce serait en effet parce qu'ils se fient aux réponses provoquées dans certaines zones de leur cerveau par l'observation du comportement, des attitudes et des expressions des autres. Notons que, dans cette expérience, la différence d'empathie à l'égard des mauvais joueurs était moins marquée chez les observateurs de sexe féminin que chez ceux de sexe masculin ; de même, les hommes présentaient une réponse de « récompense » plus marquée que celle des femmes lorsque la douleur était infligée aux mauvais joueurs. Ce serait peut-être pour la même raison, suggèrent les auteurs, que, dans une société, les individus de sexe masculin apparaissent comme ayant un rôle prédominant pour le maintien du droit et la punition pour violation de la loi.

Même si elle est modulable par des facteurs comme le contexte de l'interaction entre individus ou comme l'expérience acquise au cours de cette interaction, la réponse cérébrale concernant la sympathie ou la confiance reste une réponse automatique. Peut-on en déduire que c'est sur ce type de réponse que se déterminent les comportements « pro-sociaux » d'aide ou de coopération qui rentrent dans la définition de l'altruisme réciproque? Cette question a été examinée chez des sujets dans une situation de jeu où le gain était fonction du degré de coopération entre les deux partenaires (Rilling et al, 2002). Lorsque la coopération était bonne, les zones cérébrales impliquées dans le traitement de la récompense étaient activées, et cette activation augmentait au cours du jeu aussi longtemps que la coopération se prolongeait. Toutefois, ces mêmes zones étaient désactivées dès que le sujet soupçonnait son partenaire de ne pas jouer le jeu de la coopération. La conclusion des auteurs est que l'activation de ces régions pourrait contribuer à l'altruisme réciproque, et donc au maintien de relations sociales de coopération entre les individus, grâce au sentiment positif de récompense qui lui serait associé. Du même coup, les comportements égoïstes se trouveraient inhibés.

L'empathie, comme nous l'avons vu tout au long de la première partie de cet article, est bien le moteur de l'altruisme; mais seule l'empathie contrôlée peut être le moteur de l'altruisme réciproque, facteur de lien social et de coopération, alors que l'altruisme unidirectionnel est au contraire un

facteur d'isolement. Telle était déjà l'intuition de Carl Rogers: la confiance est dans la réciprocité. Dans la psychothérapie d'inspiration Rogerienne, un premier pas est franchi par le psychothérapeute, mais sa démarche empathique, pour être efficace, doit être perçue et comprise par le patient pour qu'il fasse confiance au psychothérapeute. Cette conception, nous l'avons vu, n'est pas très éloignée de celle de Winnicott et de ses successeurs. Le va et vient entre empathie, confiance en l'autre et confiance en soi semble jouer un rôle capital pour l'équilibre cognitif, ainsi que le souligne Philippe Jeammet. Le regard positif porté par l'autre est un déterminant de l'estime de soi : on se sent alors valorisé et renforcé dans sa créativité. Faute de cette attitude positive de l'autre, la perte de la confiance en soi peut aboutir à la violence et même à l'autodestruction. L'absence de valorisation et a fortiori la dévalorisation, qui tuent la confiance en soi ont les répercussions que l'on sait dans le milieu scolaire ou le milieu de travail.

Dans toute interaction sociale, la composante automatique de l'empathie joue le rôle d'amorce. Une fois le processus lancé, un contrôle doit intervenir pour orienter la relation vers une dimension pro-sociale. Ici se situent des processus cognitifs où sont représentées les valeurs et les croyances du sujet. Ces processus font l'objet de la section suivante sur l'approche mentaliste.

2. L'Approche mentaliste

La communication sociale par le corps, déclenchée automatiquement par les signaux sociaux, est donc essentielle pour établir une relation et amorcer une coopération entre les individus. L'empathie, qui est le moteur de ce mécanisme, a été un des principaux objets d'étude des neurosciences sociales : ainsi ont pu être délimités les contours d'un « cerveau empathique » qui inclut essentiellement les régions du cortex préfrontal et de l'amygdale traitant les émotions, ainsi que celles qui répondent au renforcement positif dans le striatum ventral. Ce mode de communication par l'empathie, cependant, est loin d'être exclusif. La compréhension de l'autre, condition nécessaire pour une communication réussie et pour l'établissement d'une

relation de confiance, implique aussi la possession, par le sujet, d'un système de représentations détaché des réactions aux signaux sociaux, d'une « théorie », lui permettant de comprendre le monde social à partir de la connaissance de règles implicites. Ce système de représentations et de règles est de nature mentale : il élabore des jugements, des inférences et des prédictions sur le comportement des autres membres du groupe. En psychologie cognitive, cette approche « mentaliste » des états mentaux des autres, s'oppose à l'approche « simulationniste » qui faisait l'objet du chapitre précédent. « Théorie » comme attitude détachée de l'immédiateté et de l'automatisme des signaux sociaux, « règles », « jugements » : tout cela fait barrage au danger subjectiviste d'une empathie purement émotionnelle.

La Théorie de l'esprit

Les recherches en psychologie cognitive ont popularisé, sous le terme de « Théorie de l'esprit » (Theory of Mind, ToM), la capacité particulière qu'ont les êtres humains (et qu'ils partagent, semble-t-il, avec très peu d'autres espèces) d'attribuer à leurs congénères des états d'esprit (Premack et Woodruff, 1978). Le fait de disposer soi-même de croyances, d'intentions, de désirs permet, par une sorte d'analogie ou d'inférence, d'en attribuer aux autres. Il s'agit donc d'une activité représentationnelle au second degré : je possède une représentation de mes états mentaux, et j'en infère que mes semblables ont aussi le même genre de représentations.

Cette capacité des humains de former des représentations de représentations (la capacité « méta-représentationnelle ») témoigne de l'existence d'un ensemble de mécanismes (d'un module, au sens que la psychologie cognitive donne à ce mot), d'un type particulier. En effet, alors que les autres modules de notre système cognitif (les modules de premier ordre) traitent de concepts en relation avec les choses du monde extérieur, le module méta-représentationnel est un module de second ordre qui traite de croyances, de désirs et d'intentions fondés sur le contenu des modules de premier ordre. Ce module est « multi-tâches » : dans le cadre qui nous intéresse ici, Dan Sperber définit son rôle comme « la capacité à comprendre

et catégoriser le comportement non pas en termes de simples mouvements corporels, mais en termes d'états mentaux sous-jacents. » (La contagion des idées, 1996). Comprendre le comportement en termes de mouvements corporels, nous l'avons vu, relève de la simulation et de l'empathie. Ici, il s'agit d'une compréhension détachée de la réalité observable, fondée sur la représentation des états mentaux des autres. Cette propriété confère au module de second ordre une flexibilité particulière, qui fait que les métareprésentations qui en sont issues ne cadrent pas nécessairement avec les croyances intuitives que nous tirons des modules de premier ordre, ceux qui sont en contact avec la réalité. On peut ainsi voir voisiner dans notre système cognitif des représentations contradictoires au sujet des mêmes faits ou des mêmes événements. Pascal Boyer décrit de nombreux exemples tirés de la pensée magique où les croyances magiques contredisent les observations objectives fondées sur les lois de la physique.

La Théorie de l'esprit, avec sa structure méta-représentationnelle qui lui permet d'inférer les états mentaux des autres sans avoir recours à l'observation directe, a donc toute sa place dans la communication entre les individus. Il ne s'agit plus, comme je viens de la rappeler, d'une communication de premier niveau au travers de la réponse empathique et la simulation automatiques du comportement observable des autres ; il s'agit au contraire d'une communication explicite fondée sur des connaissances sur la « psychologie » des autres. Grâce à cette expertise universellement répartie, je peux deviner les états d'esprit de mes congénères et faire des prédictions à leur sujet, non seulement sur leur comportement à mon égard, mais également sur ce qu'ils pensent de moi, sur les conséquences de mon comportement à leur égard, et ainsi de suite.

Mensonge et méfiance

En contrepartie, on peut se poser la question du degré de fiabilité de la communication que permet la Théorie de l'esprit : le rapport avec autrui peutil être digne de confiance s'il repose sur des croyances détachées de la réalité et qui, de surcroît, seraient manipulables ? Autrement dit, l'existence même de la Théorie de l'esprit, qui, comme nous allons le voir, rend possible le mensonge et la dissimulation, ne justifierait-elle pas une généralisation de la méfiance plutôt que de la confiance ? L'aphorisme « gouverner, c'est faire croire » attribué à Nicolas Machiavel, expert en la matière, suggère que le mensonge et la traîtrise, pourraient représenter des moyens de pouvoir autrement plus efficaces que la sincérité et la transparence. Dans cette section, nous analysons le développement de la Théorie de l'esprit chez l'enfant, ce qui nous permettra de définir les bases psychologiques des comportements mensongers.

Comme d'autres modules cognitifs (en particulier celui qui permet l'acquisition du langage), la Théorie de l'esprit est présente très précocement chez tous les individus sains, mais subit une évolution au cours du développement. Une série de tests réalisés chez des enfants de différents âges a révélé, vers l'âge de 3-4 ans, l'existence d'un stade critique marqué par l'acquisition de la notion de « fausse croyance ». Alors qu'avant ce stade l'enfant considère que ses propres états mentaux sont partagés par tous les autres, il devient capable d'envisager que d'autres puissent avoir des états d'esprit différents des siens propres. L'acquisition de la notion de fausse croyance chez l'enfant a pu être étudiée grâce à des tests simples et ingénieux que l'on doit à l'école anglaise de psychologie de l'enfant, dont le plus connu est celui dit de « Sally et Ann » (Baron-Cohen et al, 1985). Dans une des variantes de ce test, l'enfant à tester se trouve dans la pièce en compagnie de Sally et Ann. Deux récipients opaques A et B sont placés sur la table. Ann place ostensiblement un bonbon dans le récipient A, après quoi Sally sort de la pièce. En l'absence de Sally, Ann transfère le bonbon dans le récipient B. On pose alors à l'enfant la question suivante : « Lorsque Sally va revenir, où va-t-elle chercher le bonbon? ». Pour le tout jeune enfant, chez qui la Théorie de l'esprit ne s'est pas encore développée, Sally ne peut rechercher le bonbon que là où il se trouve réellement, c'est-à-dire dans le récipient B : n'est-ce pas là que l'enfant lui-même sait qu'il se trouve? Comment Sally pourrait-elle penser autrement que lui ? Ce n'est que lorsque la capacité d'attribuer à Sally une fausse croyance se sera développée, que l'enfant pourra donner la bonne

réponse, à savoir que Sally va d'abord regarder dans le récipient A, là où elle croit que le bonbon se trouve encore.

L'apparition de la notion de fausse croyance, d'après la réponse à ce genre de test, se situe autour de l'âge de 4 ans. Elle pourrait en réalité être plus précoce : d'autres expériences, où l'on tient compte du comportement de l'enfant (par exemple, des déplacements de son regard) et non plus seulement de ses réponses explicites, montrent qu'il commence à attribuer des croyances fausses à un interlocuteur dès l'âge de 18 mois. Cette acquisition ouvre de nouvelles possibilités pour l'enfant : elle lui permet de prédire le comportement de l'autre, non pas en fonction de ses propres crovances, mais en fonction de croyances qu'il lui attribue. Il devient capable de « manipuler » le contenu mental de l'autre en lui communicant des informations fausses : tel est le sens de comportements comme faire semblant, dissimuler, tromper ou mentir, que tous les enfants savent utiliser à partir de quatre ans. Chez l'adulte, le pas est vite franchi, commente Sperber, pour que nous ayons des désirs sur les états mentaux des autres, ou que nous formions l'intention de les modifier. Tout en sachant que la ToM peut produire des résultats erronés, surtout si elle consiste uniquement à attribuer ses propres états mentaux à autrui. La réalité de l'autre peut en effet fortement différer de ce que je lui attribue. Une dose de méfiance serait donc une qualité d'une ToM prête à l'autocritique.

Le développement de la Théorie de l'esprit chez l'enfant va de pair avec l'acquisition d'une sorte de « vigilance épistémique » à l'encontre de la tromperie. Cette capacité se manifeste normalement chez tout un chacun par une double tendance : d'une part, la tendance à faire confiance aux autres et à considérer que les informations qu'ils fournissent sont fiables ; et d'autre part la tendance, encore plus forte, à être vigilant à l'encontre du risque de communication d'informations fausses. Cette tendance conduit à préférer l'interaction avec des individus identifiés comme bons et bienveillants. L'étude du développement de cette préférence chez l'enfant montre qu'elle s'établit de manière précoce. Lorsqu'on présente à des enfants d'un an un dessin animé qui met en scène (sous la forme de figurines géométriques) un

« méchant » et un « gentil », on constate qu'ils s'intéressent davantage (regardent plus longtemps) le « gentil » que le « méchant » (Premack et Premack, 1997). L'enfant est donc capable très tôt de faire la différence entre malveillance et bienveillance et peut-être, par là même, de déterminer à qui faire confiance. Olivier Mascaro, dans sa thèse consacrée à ce problème de la vigilance épistémique, rapporte de nombreuses recherches chez de jeunes enfants tendant à montrer que ceux-ci savent évaluer les informateurs et former des préférences durables en faveur de ceux qui leur paraissent fiables (Mascaro & Sperber, 2009). Dans le but d'étudier ce phénomène, Mascaro et ses collaborateurs ont inventé une tâche voisine de la tâche de fausse croyance décrite plus haut. Il s'agit de la tâche de « fausse communication » : l'enfant doit retrouver dans quelle boîte un objet a été caché. Un informateur, figuré par une poupée, donne à l'enfant une information fausse sur la boite où se trouve l'objet. L'expérimentateur, de son côté, donne à l'enfant des indices lui permettant de croire que cette information est fausse. Pour réussir la tâche, l'enfant doit donc ignorer la boite désignée par la poupée, ce qu'il ne parvient à faire qu'à partir de l'âge de 4 ans. On remarquera que ce résultat est cohérent avec celui de la maîtrise de la tâche de fausse croyance, qui se produit également à cet âge.

Théorie de l'esprit et activité cérébrale

Les travaux sur l'anatomie fonctionnelle de la Théorie de l'esprit se confondent en partie avec ceux qui concernent la cognition sociale au sens large et qui incluent les interactions entre soi et l'environnement social, la connaissance des règles sociales, la compréhension des intentions des autres. Les tâches qui sont utilisées pour ces travaux consistent à observer des situations présentées sous la forme de dessins, de bandes dessinées ou de vidéos et qui requièrent de porter un jugement sur l'état d'esprit d'une autre personne : les tâches de fausse croyance évoquées plus haut en font évidemment partie. On peut ainsi comparer l'activation produite par des tâches d'attribution de différentes formes d'intentions à soi-même ou à d'autres personnes, intentions qui peuvent elles-mêmes être d'ordre « privé »

(il va/je vais éclairer la lumière pour pouvoir lire) ou d'ordre « communicatif » (il demande / je demande la direction à suivre en montrant une carte). Dans les deux cas, l'activation siège dans le cortex préfrontal médian: la mentalisation au sujet d'intentions privées active la partie postérieure de cette région du cortex, tandis que les intentions communicatives impliquent sa partie plus antérieure (revue dans Amodio et Frith, 2006). D'autres régions sont également concernées par les tâches qui relèvent de la Théorie de l'esprit. La cognition sociale, comme les autres aspects de la cognition, nécessite l'intervention de réseaux connectant le traitement de la conscience de soi, de la motivation, de l'anticipation des conséquences d'actions, du jugement par inférence, réseaux qui incluent, non seulement le cortex préfrontal médian, mais également d'autres régions de ce vaste ensemble. Ces résultats convergent avec les multiples travaux sur les effets de lésions pathologiques siégeant dans les différentes régions du cortex préfrontal. On sait depuis longtemps que les lésions affectant les régions médianes et ventrales altèrent le comportement social. Il en résulte une interprétation des situations sociales. mauvaise des comportements socialement inappropriés qui traduisent l'incapacité à anticiper et comprendre les intentions et les états d'esprit des autres. Rappelons que les lésions altérant la compréhension des émotions et l'empathie ont un siège plus ventral, dans la région du cortex orbitaire. Ces régions sont mises en jeu dans des tâches nécessitant la prise de décision dans des dilemmes moraux à caractère social (Greene et al, 2001).

Enfin, une autre forme de pathologie, celle des enfants présentant des troubles envahissants du développement cognitif, dont le type le plus connu est l'autisme infantile, reste à inventorier sur le plan des altérations de l'activité cérébrale. Contrairement aux autres enfants, les enfants autistes présentent des troubles du développement des aptitudes sociales : privés de cette capacité, ils ne peuvent exploiter les signaux sociaux de leurs congénères (expressions faciales, position du regard), et deviennent incapables d'entrer en communication avec d'autres individus. Cette incapacité est confirmée par l'échec de ces patients dans les tests classiques de la Théorie de l'esprit,

comme le test de fausse croyance auquel ils échouent (Baron-Cohen et al). Ce fait attire l'attention sur l'existence de mécanismes cérébraux dont le bon fonctionnement serait nécessaire à l'apparition de la Théorie de l'esprit, et dont la maturation, pour des raisons génétiques ou autres, se révélerait déficiente chez les enfants autistes.

En conclusion

Les neurosciences cognitives apportent donc une contribution essentielle à la compréhension des relations de confiance/méfiance. Une des principales indications qu'on peut tirer des travaux de psychologie cognitive et de neuro-imagerie décrits ici est que ces relations s'établissent à plusieurs niveaux différents de la communication entre les individus. Nous distinguerons ici deux niveaux, celui de la communication par le corps et celui d'une communication plus abstraite, qui utilise des mécanismes purement mentaux, sans expression visible, et donc inaccessibles directement de l'extérieur.

En ce qui concerne la communication par le corps, son mécanisme central est celui de l'empathie. Il s'agit d'une communication directe, véhiculée par des signaux « sociaux » (expressions du visage, direction du regard, postures et mouvements, intonations de la voix, etc.). Les signaux sociaux émis par un des partenaires de la relation ont la particularité de pouvoir être immédiatement reconnus par l'autre partenaire, et inversement. Il en résulte une communication automatique qui assure l'amorçage de la relation, avec des réactions d'acceptation ou de rejet, de sympathie ou d'antipathie, formes implicites de la confiance ou de la méfiance.

L'autre forme de communication, celle qui ne passe pas par les signaux du corps, est fondée sur la « mentalisation » de la relation. Il s'agit donc d'une forme de communication indirecte qui ne nécessite pas la présence physique des partenaires : ce qui compte, ce n'est pas le corps visible et audible du partenaire, c'est plutôt son contenu mental, ses états d'esprit. Ceux-ci peuvent être compris et connus grâce à des mécanismes cognitifs dont l'ensemble constitue la Théorie de l'esprit. Cette forme de communication remplit une

fonction bien différente de celle qui est assurée par la communication par empathie : à l'abri des réactions immédiates, elle permet l'établissement de relations à plus long terme, fondées sur des critères rationnels d'avantages et de crédit mutuel.

Les deux niveaux de communication décrits ici mettent en jeu des mécanismes cérébraux différents, et peuvent être atteints par des processus pathologiques distincts. La communication par le corps fait intervenir les mécanismes sensoriels de décodage des signaux sociaux émis par les partenaires, ainsi que le système d'expression et de reconnaissance des émotions situés dans la partie la plus ventrale du cortex préfrontal et des structures sous-corticales comme les noyaux amygdaliens. La communication mentalisée, tout en utilisant aussi les marqueurs émotionnels, met en jeu le système d'anticipation et d'évaluation cognitive localisé principalement dans la région médiane du cortex préfrontal.

Neurosciences, déterminisme et libre arbitre¹

Bernard Feltz²

Introduction

Les neurosciences, ou sciences du système nerveux central, connaissent des développements considérables ces dernières années. Les rapprochements entre la neurologie et la psychologie cognitive ont conduit à des avancées majeures et au renouvellement d'une approche rigoureuse du comportement humain en lien avec le fonctionnement du système nerveux. Par ailleurs, plus récemment, ces travaux ont donné lieu à des synthèses audacieuses, parfois de la part des scientifiques eux-mêmes, qui proposent des conceptions plus larges de la conscience et envisagent les conséquences des travaux scientifiques sur une conception de l'humain³. Enfin, en philosophie, ces thèmes sont au coeur d'un courant de recherche important qui s'est développé au départ des travaux en philosophie des sciences et en philosophie du langage⁴.

C'est dans ce contexte général que s'inscrit cette contribution. Les développements des neurosciences sont susceptibles de modifier notre conception de l'humain, notre conception du rapport au corps, notre conception des rapports à l'animal, voire à la nature : c'est toute la question de la spécificité de l'humain qui est en cause.

Je voudrais aborder cette question complexe à partir d'un angle d'approche particulier : la question du déterminisme. Il faut souligner un paradoxe qui marque profondément le monde des chercheurs. Notre système social, et en particulier tout notre système juridique, repose sur la présupposition que le comportement humain se caractérise par une certaine

² Institut supérieur de Philosophie, UCLouvain (bernard.feltz@uclouvain.be)

Connaître | Nº 36-37 | décembre 2011

¹ Texte publié in : *La Revue Nouvelle*, **65**, 3, Mars 2010, 60 -71.

³ Cf. A. Berthoz, J.P. Changeux, F. Crick, A. Damasio, G. Edelman, M. Jeannerod, F. Varela.

⁴ Cf. P. Churchland, D. Dennet, F. Dretske, F. Jackson, Th. Nagel, J. Searle.

dimension de libre arbitre. Lorsqu'un individu est jugé par un tribunal, s'il est considéré comme responsable, on présuppose qu'il aurait pu ne pas commettre l'acte pour lequel il est poursuivi. Même si on accepte le principe de circonstances atténuantes, on ne remet pas en cause cette caractéristique fondamentale du comportement humain qui est marqué par une capacité de choix. C'est d'ailleurs ce qui justifie certaines décisions juridiques de non responsabilité; dans certains dossiers, la personne incriminée est jugée irresponsable par un collège d'experts qui estiment que la personne en question ne dispose pas de cette capacité de choix: le comportement de cette personne est jugé comme strictement déterminé par des facteurs qui échappent à son contrôle. Cela manifeste à l'inverse combien le comportement d'un individu jugé responsable est considéré comme relevant ultimement de sa décision, autrement dit de son libre arbitre.

Une telle présupposition n'est pas celle des neuroscientifiques. Chercher à expliquer, pour le neuroscientifique, revient au contraire à tenter de mettre en évidence un lien de causalité entre le fonctionnement du système nerveux et le comportement de l'individu. Pour le neuroscientifique, présupposer le libre arbitre revient à renoncer au projet explicatif. Les neuroscientifiques s'inscrivent donc d'emblée dans une perspective déterministe où l'objectif de la recherche est précisément d'établir les liens de causalité entre fonctionnement du système nerveux et comportement humain.

Ce paradoxe n'est pas anodin ; il touche au fondement de notre conception de l'humain : en quel sens peut-on parler d'individu libre, si on prend au sérieux les déterminations corporelles du comportement ? Je voudrais tenter d'éclairer cette question en proposant une démarche en trois temps. En un premier temps, je me référerai à la conception développée récemment par un grand connaisseur des neurosciences et tout à la fois un homme d'une grande culture philosophique : H. Atlan prend appui sur le concept d'autoorganisation pour tenter de penser le rapport entre fonctionnement du système nerveux et comportement humain en lien avec les conceptions de Spinoza. En un deuxième temps, je proposerai une réflexion basée sur l'histoire de cette question dans le champ de la philosophie des sciences. Enfin, en une troisième

partie, je voudrais montrer que le déterminisme défendu par H. Atlan n'est pas inéluctable et que, en lien avec les théories neuroscientifiques récentes, une interprétation plus ouverte à une liberté effective est défendable.

1. Auto-organisation et liberté : quand H. Atlan rencontre Spinoza

H. Atlan se réfère à la problématique de l'auto-organisation pour éclairer ce paradoxe d'un comportement déterminé par le système nerveux et pourtant considéré comme indéterminé sur le plan social⁵. Un des modèles classiques d'auto-organisation, auquel se réfère H. Atlan, consiste en des réseaux d'automates booléens, reliés en une structure où chaque automate comporte deux entrées et deux sorties. Sans entrer dans des considérations techniques, précisons simplement que de tels réseaux d'automates booléens sont des systèmes strictement déterministes, dont le comportement est donc prédictible par calcul au moyen d'ordinateurs puissants. Un réseau de 20×20 automates, en fonction de la valeur initiale des automates et de la répartition des opérateurs, va donner à terme une configuration où un grand nombre d'automates prennent une valeur qui reste définitivement stable, tandis que d'autres restent définitivement instables. Les ensembles d'opérateurs qui présentent des comportements analogues constituent des formes qui n'ont de sens qu'au niveau macroscopique de l'ensemble du réseau.

H. Atlan suggère de considérer ces réseaux d'automates booléens comme des archétypes de ce que l'on pourrait appeler l'émergence de formes nouvelles. Dans le réseau 20×20, cette émergence est parfaitement prédictible et calculable. Mais si on considère des réseaux plus complexes tels que 1000×1000, voire, pour être dans l'ordre de grandeur des réseaux de cellules nerveuses dans le cerveau, de l'ordre de 1 million × 1 million, on entre dans un registre où le calcul devient impossible pour deux raisons : il est impossible de définir l'ensemble des conditions initiales – répartition des divers automates et

⁵ H. Atlan, « Le monisme radical du corps et de l'esprit », in *Les étincelles du hasard. Tome 2 : Athéisme de l'Ecriture*, Paris, Seuil, 2003, pp. 199-274.

valeur de départ de chaque opérateur (0 ou 1) – et d'autre part le calcul requerrait des ordinateurs d'une puissance considérable et le temps de calcul prendrait lui-même des valeurs inacceptables.

Pour H. Atlan, les réseaux d'automates booléens présentent donc toutes les caractéristiques permettant la modélisation de l'émergence : une structure au comportement strictement déterministe produit des figures macroscopiques qui sont pourtant imprévisibles, parce qu'incalculables à la fois en raison de l'impossibilité d'une connaissance exhaustive des conditions initiales et en raison du caractère trop complexe des calculs qui seraient nécessaires à la prédiction. Le déterminisme au niveau microscopique est parfaitement compatible avec une imprédictibilité au niveau macroscopique.

C'est ce phénomène paradoxal qui est visé par le concept d'autoorganisation : pour H. Atlan, l'auto-organisation c'est l'émergence de propriétés nouvelles, imprédictibles, au sein d'un système déterministe. L'émergence, c'est cette imprédictibilité de fait du comportement d'un système strictement déterministe.

Les cellules nerveuses de notre cerveau ne se comportent certes pas comme des réseaux d'automates booléens. Elles présentent une connectivité beaucoup plus complexe et, par ailleurs, elles présentent une plasticité, une capacité de réarrangement, que les réseaux d'automates booléens sont incapables de simuler. C'est donc en pleine conscience du caractère simplificateur du modèle de référence que H. Atlan propose de rapprocher le comportement des réseaux d'automates du comportement du cerveau. Il s'agit de montrer qu'il est parfaitement possible d'attribuer, au niveau macroscopique, un comportement imprédictible à un organisme pourtant complètement déterminé au niveau microscopique.

Le concept de « liberté » ou « libre arbitre » renvoie précisément à cette impossibilité de prédiction associée à la complexité du système nerveux, lequel système connaît pourtant un fonctionnement strictement déterministe. Le concept de « liberté » porte sur le niveau macroscopique. Le déterminisme caractérise le niveau microscopique. Le concept de « liberté » renvoie fondamentalement à l'impossibilité de calculer le passage du microniveau au

macroniveau. En ce sens, la « liberté » est une illusion, puisque le système est déterministe, mais une illusion nécessaire puisque, dans les faits, les comportements macroscopiques sont imprévisibles. Dans ce contexte, le macroniveau est illusoirement irréductible au microniveau.

H. Atlan rapproche une telle position de la conception de la liberté chez Spinoza. Ce penseur de la deuxième moitié du dix-septième siècle, précédant de quelques années les travaux de Newton, propose une conception panthéiste de la nature où l'humain participe à la vie de Dieu. Le déterminisme intégral, qui caractérise l'univers, est au fondement de la science nouvelle et renvoie en même temps à l'omniscience de Dieu. C'est précisément parce que le déterminisme est intégral qu'une science est possible. Et la liberté de l'humain est cette illusion nécessaire, produit du déterminisme, et fondement d'une vie sociale démocratique à imaginer. Spinoza est marqué par ce paradoxe de l'affirmation d'un déterminisme intégral et de l'anticipation d'une réflexion sur le fonctionnement démocratique de la société.

Il faut souligner l'élégance de la position de H. Atlan qui propose une version contemporaine parfaitement cohérente et remarquablement argumentée de l'architecture spinoziste. La compatibilité est posée entre l'affirmation d'une liberté à l'œuvre, même illusoire, et d'un déterminisme qui marque la démarche neuroscientifique.

Je ne peux m'empêcher d'exprimer certaines réticences devant une telle conception, réticences liées à la fois aux options philosophiques et aux considérations scientifiques. Cette conception d'une liberté illusoire liée à un déterminisme intégral au niveau microscopique ne permet pas de rendre compte de la distinction socialement admise entre personne responsable et personne irresponsable. Dans l'irresponsabilité, le déterminisme est socialement acté. Mais là où la responsabilité est reconnue, le déterminisme est nié. La conception spinoziste ne me paraît pas susceptible de rendre compte de cette distinction, pourtant fondamentale dans notre système social. L'argument n'est certes pas décisif, un consensus social n'a pas valeur de vérité, mais il mérite d'être mentionné. D'autre part, et c'est un des enjeux de cette contribution, je voudrais montrer que cette position ne s'impose pas au vu de

l'évolution des neurosciences contemporaines. La présupposition d'un déterminisme intégral du microniveau relève d'une logique que les neurosciences contemporaines n'imposent pas, ou tout au moins que certaines théories du fonctionnement du système nerveux permettent d'éviter.

Avant d'approfondir cet argument, je voudrais souligner l'ancrage philosophique des positions déterministes et non déterministes par un rapide parcours historique de la question.

2. Liberté et déterminisme : perspective historique

Le 17^{ème} siècle est le siècle de l'émergence de la science moderne. La science galiléenne s'est construite dans le registre à la fois de la rupture et de la continuité avec le monde médiéval. Le concept de force et le rejet de la cause finale au profit de la causalité efficiente marquent une rupture claire avec le rapport médiéval à la nature. Par contre, la présupposition d'un monde déterministe est en stricte continuité avec l'omniscience du Dieu médiéval où la compatibilité entre omniscience de Dieu et liberté humaine a toujours fait problème. Rappelons également que ce même 17^{ème} siècle est le siècle du jansénisme avec Cornélius Jansen et Blaise Pascal qui tendent à défendre une perspective de prédestination, conséquence logique d'une omniscience à l'origine d'un monde strictement déterministe.

Les philosophes Spinoza et Leibniz s'inscrivent clairement dans cette tradition d'un monde strictement déterministe et d'un Dieu rationnel omniscient à l'oeuvre de manière immanente dans la nature. Les scientifiques Newton, Laplace, Einstein s'inscrivent dans cette même tradition en termes certes plus dégagés des considérations théologiques, mais néanmoins sous la modalité d'une présupposition d'un monde strictement déterministe.

Le « démon de Laplace » a pris dans cette perspective une valeur emblématique. C'est dans ses réflexions sur les statistiques au début du 19^{ème} siècle que le mathématicien Laplace imagine un « démon » capable d'observer, en un instant donné, la position et la vitesse de chaque masse constitutive de l'univers, et capable d'en déduire l'évolution universelle vers le

passé comme vers l'avenir. Ce démon constitue bien l'affirmation d'un univers strictement déterministe, en dehors de toute référence à la divinité. La physique moderne à son origine s'est donc inscrite dans une conception déterministe de l'univers, et cette conception marque profondément l'histoire de la physique jusqu'à la période contemporaine.

Le début du 20^{ème} siècle a vu l'émergence d'une autre conception à partir de deux lieux distincts. La physique quantique, tout d'abord, avec la relation de Heisenberg, se prête à une double interprétation, dont l'une ouvre la voie à des espaces d'indétermination dans le monde des particules élémentaires. L'autre lieu vient de la réflexion sur les théories de l'évolution. Bergson, dans son ouvrage L'évolution créatrice, prend une distance radicale par rapport au monde déterministe de Laplace ou Leibniz. Pour Bergson, le déterminisme intégral est négation du temps. Si le monde est strictement déterminé, si le « démon de Laplace » est capable de prédire l'évolution du monde en totalité, le temps n'est qu'illusion, le temps est la mesure de notre ignorance. Si le « démon de Laplace » peut prévoir l'avenir, c'est que l'avenir est complètement inclus dans le passé : « rien de neuf sous le soleil » disait le Qohélet... Il n'y a dès lors aucune surprise... Bergson, quant à lui, défend au contraire une conception forte du temps : il n'y a de temporalité forte que dans un monde où il y a nouveauté absolue, « imprévisible essentiel ». C'est en cela que le concept de « création » prend sens⁶. L'évolution « créatrice », c'est la production de la nouveauté au sens fort et, pour que la nouveauté soit réelle, il faut qu'elle soit imprévisible. Imprévisibilité, nouveauté et temporalité forte sont les diverses facettes d'une même réalité qui renvoie à l'évolution créatrice.

C'est ce concept fort de temporalité qui marque les conceptions de Prigogine et Stengers qui se sont faits les porte-paroles d'une physique articulée à un monde non strictement déterministe. On le sait peut-être moins,

_

⁶ Pour Bergson, la position finaliste tout comme la position mécaniste s'articule à une conception faible de la temporalité comme mesure de notre ignorance. Le finalisme de Leibniz est tout aussi inacceptable que le mécanisme.

mais c'est également ce concept fort de temporalité qui est au coeur de la position de Jacques Monod dans son célèbre ouvrage *Le hasard et la nécessité*.

Monod se réfère explicitement à Bergson et à l' « imprévisible essentiel » pour souligner la contingence de l'évolution biologique. Les mécanismes de production aléatoire de mutations sont complètement indépendants des mécanismes en jeu dans la sélection naturelle. La résultante ne peut donc être le résultat d'une programmation. Il y a une dimension irréductible d'indéterminisme. Bien plus, Monod distingue sélection organismique et sélection environnementale. Il parle donc de deux niveaux de sélection. Une mutation doit tout d'abord s'intégrer dans un organisme et contribuer au fonctionnement de l'organisme en tant que tel. Et, d'autre part, cet organisme s'inscrit dans la dynamique de la sélection naturelle par confrontation aux contraintes environnementales. Ces articulations complexes s'inscrivent clairement dans un monde marqué par la contingence et la nouveauté imprévisible.

Sur le plan de l'histoire des idées, il est surprenant d'observer une telle proximité entre deux univers aussi distants que le vitalisme de Bergson et la biologie moléculaire de Monod : ces deux auteurs se rejoignent sur la conception d'un temps comme lieu de production d'une nouveauté irréductible, ils se rejoignent sur la conception d'un monde non strictement déterministe.

Ce rapide historique tend à montrer combien la question du déterminisme fait débat aussi bien dans le monde des philosophes que des scientifiques. Le débat a connu des apports décisifs au $20^{\rm ème}$ siècle, dans plusieurs domaines scientifiques. Au sein de la biologie, ce débat s'est inscrit dans l'interprétation des théories de l'évolution biologique. À la fin du $20^{\rm ème}$ siècle, le concept de « sélection » a pris dans ce débat une place décisive, comme nous allons le voir dans le contexte des neurosciences.

3. Conscience et sélection : quand Edelman « rencontre » Kant

3.1. Une biologie de la conscience : Edelman et la théorie de la sélection des groupes neuronaux

Edelman est un des plus grands neuroscientifiques contemporains. Prix Nobel de médecine pour ses travaux en immunologie, il s'est orienté vers les neurosciences depuis les années 1970. La présentation de ses travaux ici proposée est plus que succincte⁷. L'essentiel est d'en dégager les enjeux philosophiques.

La théorie de la sélection des groupes neuronaux vise les mécanismes de mise en place de la structure fine des connexions dans le cerveau. Ces connexions se comptent par milliards et les modèles explicatifs de la mise en place de cette connectivité sont confrontés à des difficultés majeures. L'originalité de l'approche edelmanienne, comme celle de Changeux d'ailleurs, est d'introduire une logique sélectionniste dans cette explication. Prenons l'exemple de l'apprentissage de la marche. À un stade particulier de l'évolution de l'enfant, en fonction des mécanismes de programmation génétique, on observe un haut degré de connectivité entre des cartes des divers centres – moteurs, sensoriels, émotions – qui fait que toute entrée dans un centre comporte des informations transmises dans les autres centres. C'est le processus de cartographie globale qui connaît un haut niveau de redondance.

L'apprentissage de la marche se fait par la méthode des essais et erreurs. La structure redondante du système nerveux permet à l'enfant d'essayer des modes de déplacement qui sont plus ou moins efficaces, par conséquent plus ou moins bien gratifiés. Sur le plan neurologique, les circuits utilisés sont renforcés, et ce d'autant plus qu'ils sont utilisés. L'association des comportements avec le système des émotions conduit à une gratification des attitudes les plus efficaces ce qui mène l'enfant à reprendre et améliorer ces

⁷ G. Edelman, Biologie de la conscience, Paris, Odile Jacob, 1992; (avec G. Tononi) Comment la matière devient conscience, Paris, Odile Jacob, 2000; Plus vaste que le ciel. Une nouvelle théorie générale du cerveau, Paris, Odile Jacob, 2004; La science du cerveau et la connaissance, Paris Odile Jacob, 2007.

attitudes et aboutit au renforcement des circuits neuronaux qui sous-tendent le comportement le plus adapté. La structure finale des circuits stabilisés est donc liée au comportement de l'enfant. Les circuits neuronaux non utilisés s'affaiblissent, voire disparaissent.

Le processus d'apprentissage correspond dès lors à un processus de sélection par stabilisation des circuits les mieux adaptés : l'adaptation à l'environnement est le fruit du processus d'apprentissage, qui est un processus épigénétique et non programmé.

Ce qui vient d'être décrit pour l'apprentissage de la marche pourrait l'être de la même manière pour l'apprentissage du langage. Cela signifie que la structure du cerveau d'un individu adulte est le fruit de ses divers apprentissages, y compris de son appartenance culturelle : l'apprentissage du chinois conduit à des connectivités différentes de l'apprentissage du français.

Un tel schéma de base appelé « sélection des groupes neuronaux », Edelman l'intègre dans une conception évolutive de la conscience. Pour Edelman, les animaux ont une conscience, une conscience qu'il qualifie de « primaire » en ce sens que les mammifères supérieurs, par exemple, ont une capacité de représentation qui leur donne une maîtrise de l'environnement que n'ont pas les animaux qui en sont dépourvus. Les stratégies de chasse des prédateurs, de même que les stratégies d'évitement des proies, dénotent un rapport à l'environnement médiatisé par des représentations qui intègrent les apprentissages. Edelman parle même de « concept » sans langage : la proie est capable de reconnaître un prédateur ; de même, le prédateur est capable d'élaborer des stratégies d'approche. Cependant, chez l'animal, la conscience est liée aux stimuli de l'instant présent. Ce sont les informations de l'instant qui mobilisent les processus d'apprentissage mémorisés dans la conscience animale : Edelman parle de la conscience animale comme d'un « présent remémoré ». L'animal est esclave du présent. Sa conscience primaire est marquée par l'immédiateté dans le temps.

A la différence de l'animal, l'humain dispose d'un langage articulé qui lui permet d'avoir accès à ses représentations indépendamment des stimuli de l'instant présent. L'évocation du simple concept de « lion » fait que chacun est

capable de mobiliser la représentation du lion, de « rendre présente » l'image du lion. Chez l'humain, le recours à la représentation est donc délié de l'instant. Ce qui conduit à la conscience de la distinction entre réel et imaginaire, entre passé, présent et futur. Cette conscience conduit à une distanciation qui mène à la « conscience d'ordre supérieur », la « conscience d'être conscient », dira Edelman.

La conscience s'inscrit donc dans une perspective d'évolution phylogénétique. On perçoit bien que la conscience primaire est un atout dans le processus de sélection naturelle. On peut parler d'une coévolution des prédateurs et des proies sur ce point. On perçoit également que la conscience d'ordre supérieur, cette capacité de représentation dégagée de l'instant, a été un atout majeur pour les humains dans les relations avec les autres espèces.

Sur cette base, Edelman esquisse une conception de l'humain que l'on pourrait résumer autour de trois concepts : idiosyncrasie, conscience intentionnelle, libre arbitre. L'idiosyncrasie renvoie à l'unicité de chaque individu. À partir du moment où la structure fine du cerveau est liée aux comportements, deux individus avec même patrimoine génétique, deux vrais jumeaux, ont des cerveaux différents : la théorie de la sélection des groupes neuronaux établit le caractère absolument unique de chaque individu, y compris au niveau biologique. Chaque individu est le fruit d'une histoire unique spécifique. D'autre part, la conscience humaine est une conscience intentionnelle qui oriente l'humain vers l'action. Par sa capacité de représentation, l'humain est capable d'anticiper et d'élaborer des plans d'action en fonction d'un système de représentation. Enfin, Edelman parle explicitement d'un certain niveau de libre arbitre, liberté sous contraintes, bien entendu, mais espace qui échappe au déterminisme radical.

Comment se pose la question du déterminisme dans ce contexte ? C'est ce que je voudrais évoquer maintenant, en dialogue en un premier temps avec la philosophie des sciences contemporaine, en interaction avec la position kantienne ensuite.

3.2. Sélection et déterminisme

Je voudrais montrer tout d'abord que l'explication sélectionniste en jeu dans la théorie de la sélection des groupes neuronaux rompt avec l'explication classique de la biologie moléculaire. En effet, depuis la découverte de l'ADN en 1953, la biologie a développé un mode d'explication déterministe qui associe le fonctionnement à la structure des molécules qui constituent l'organisme. L'hémoglobine est l'exemple type de ce mode d'explication. La fonction de captation de l'oxygène est une conséquence de la structure de la protéine. La structure ou la forme de la protéine est une conséquence des acides aminés est causée par la séquence des nucléotides dans l'ADN, siège du programme génétique. On est donc dans le registre d'un déterminisme biologique radical : la structure détermine la fonction, le programme génétique détermine la structure.

Dans la théorie de la sélection des groupes neuronaux, l'explication rompt avec ce schéma déterministe. En effet, nous avons vu que la structure finale du cerveau dépend des processus d'apprentissage, donc du comportement. La structure finale du cerveau ne dépend donc pas exclusivement de l'ADN. Le macroniveau a un impact sur le microniveau. Dans ce sens, on peut parler d'une « causalité descendante » : le comportement a un impact sur la structure biologique du cerveau. D'un point de vue synchronique, le comportement est sans doute produit par la structure. Mais d'un point de vue diachronique, la structure est également le produit du comportement. On peut donc parler d'une circularité dans les rapports structure-fonction. Ce n'est pas seulement la structure qui détermine la fonction, mais bien également la fonction qui détermine la structure.

De manière plus précise, si on considère l'apprentissage du langage, le parler chinois implique des connexions neuronales distinctes du parler français : la structure finale du cerveau est donc dépendante du type de langage appris. Le langage est bien un produit culturel, qui implique un vivre-ensemble. Cela signifie qu'il y a sens à dire que la structure du cerveau est déterminée autant par la culture que par la biologie : le cerveau est donc un produit biologico-culturel!

D'un point de vue philosophique plus général, on peut donc dire que, si la théorie de la sélection des groupes neuronaux est pertinente, le déterminisme biologique est scientifiquement indéfendable. L'explication sélectionniste enlève toute pertinence à une tendance majeure de l'explication biologique de la fin du 20^{ème} siècle, le déterminisme génétique. En un certain sens, on peut dès lors dire que Bergson et Monod pourraient bien avoir raison : le biologique est marqué de manière irréductible par un imprévisible essentiel.

Il faut cependant être prudent. La fin du déterminisme génétique ne signifie pas nécessairement la fin du déterminisme. Il est toujours possible de plaider pour un déterminisme environnemental : c'est du reste la thèse classique des courants behavioristes en psychologie.

Avec la fin du déterminisme biologique, une première étape importante est donc franchie sur la voie d'un concept possible de libre arbitre. Peut-on être plus affirmatif dans une prise de distance par rapport au déterminisme environnemental? Un petit détour par la philosophie kantienne va peut-être nous aider à voir plus clair.

3.3. Kant, Edelman et la question de la liberté

Au 18^{ème} siècle, le philosophe E. Kant a abordé avec rigueur la question du déterminisme, précisément en articulation à la physique newtonienne à laquelle il adhérait complètement. Kant définit la liberté comme une capacité d'inaugurer une nouvelle chaine causale. La liberté, c'est une rupture dans l'enchaînement des causes. C'est la possibilité d'une cause nouvelle non liée à une cause précédente.

Dans *La critique de la raison pure*, Kant montre, au niveau de ce qu'il appelle l'antinomie de la causalité, qu'il est impossible de démontrer que l'être humain est libre. En effet, le concept même de liberté est porteur de contradictions et ce, quelle que soit la thèse défendue. C'est le propre d'une antinomie chez Kant que chaque terme de la contradiction est porteur, en luimême, d'une difficulté. En effet, si l'on défend la thèse que la liberté existe, le monde devient incompréhensible. Si, à tout moment, un nouvel enchaînement de causes peut être inauguré, on se trouve devant un monde chaotique,

inexplicable, capricieux, totalement imprévisible... Or tel n'est pas notre monde!

D'autre part, si on nie la possibilité d'inauguration d'une nouvelle chaine causale, on se trouve devant une régression à l'infini des chaines causales, ce qui est tout aussi inacceptable. Chacune des affirmations conduit donc à une impasse. C'est le propre des antinomies. Cela signifie donc que, sur le plan de la raison pure théorique, de la raison en tant qu'elle donne accès à la connaissance, la question de la liberté ne peut être résolue. En strict rigueur de termes, on ne peut « savoir » si l'homme est libre. La raison pure théorique nous laisse devant un indécidable quant à la question de la liberté. On ne peut démontrer que l'humain est libre, comme on ne peut démontrer qu'il n'est pas libre. La possibilité existe donc réellement.

Par contre, toujours selon Kant, l'être humain se caractérise par un comportement éthique et, pour qu'un comportement éthique soit possible, l'être humain doit être libre. Il nous faut donc postuler la liberté pour rendre compte de la dimension éthique du comportement humain. La liberté résulte donc d'un postulat de la raison pratique, de la raison en tant qu'elle justifie l'action bonne. Dans ce contexte, il y a bien liberté effective — on se distancie du concept de liberté illusoire de Spinoza - ; cependant, le rapport à la liberté ne résulte pas des conclusions du savoir mais relève d'un postulat de la raison pratique. En d'autres termes, on n'est pas sûr que l'humain soit libre, ou plutôt cette certitude ne relève pas de la certitude du savoir.

En quoi ces considérations éclairent-elles la question des neurosciences? Les modèles sélectionnistes remettent en cause le déterminisme biologique, mais ne peuvent exclure le déterminisme environnemental. Ce que Kant nous dit, c'est que la position déterministe comme la position non déterministe ne peuvent être fondées de manière apodictique. Le déterminisme, comme le non déterminisme, relèvent d'une position de principe sur le plan du savoir. Dans ce contexte, la position déterministe n'est pas plus rationnelle que la position non déterministe. Puisque la raison pure théorique laisse ouverte la question de la liberté, il y a sens à poser le non déterminisme comme condition de possibilité d'un

comportement marqué par le libre arbitre. Ce positionnement n'est en rien contraire à la raison mais il ne s'impose pas par la raison pure théorique.

Quand il parle des sciences humaines, Jean Ladrière souligne le fait que toute compréhension de l'humain implique une auto-compréhension du sujet scientifique par lui-même : l'impossibilité d'une auto-compréhension absolue, l'impossibilité d'une transparence du sujet à lui-même, induit une circularité irréductible dans le champ des sciences humaines. Le cercle herméneutique, c'est la reconnaissance de l'impossibilité d'une connaissance qui ne repose pas sur des présuppositions.

Quand Edelman parle d'une conscience intentionnelle et d'un libre arbitre, on ne peut dire qu'il démontre que de tels comportements sont à l'oeuvre. Sans doute est-il plus correct de préciser que ces conceptions sont des présuppositions de sa théorie. Comme le dit Jean Ladrière, c'est la fécondité des concepts mis en oeuvre qui, a posteriori, permet de justifier la pertinence de ces concepts.

Il y a donc sens à dire que, sur ce plan, l'oeuvre d'Edelman s'inscrit dans le cercle herméneutique propre aux sciences humaines.

4. Conclusions

Un enjeu majeur de la théorie edelmanienne de la conscience est la remise en cause du déterminisme biologique radical. Cet enjeu est clairement rencontré. Avec Edelman, le déterminisme génétique radical n'est plus scientifiquement défendable. La portée philosophique d'une telle théorie est donc considérable. Un pas important est réalisé pour une collaboration franche et respectueuse entre sciences biologiques et sciences humaines.

Edelman va plus loin dans l'affirmation d'une conscience intentionnelle et dans l'évocation d'un libre arbitre. De telles affirmations vont au-delà des enjeux de la théorie biologique, mais gardent toute légitimité dans une épistémologie des sciences humaines qui intègre le cercle herméneutique.

Cet article s'est ouvert par un exposé de la position spinoziste en lien avec la problématique de l'auto-organisation. La remise en cause du déterminisme biologique souligne les limites du modèle auto-organisationnel

sur base des réseaux d'automates booléens. Le déterminisme strict du macroniveau par le microniveau n'est plus tenable dans la perspective edelmanienne. Le modèle proposé, même à titre analogique, est donc déficient et ne peut plus être retenu. La position spinoziste sur base du déterminisme biologique n'est donc plus défendable.

En lien avec les interprétations kantiennes, on peut montrer que la position spinoziste peut néanmoins encore être tenue sur le plan d'un déterminisme environnemental, mais l'argument perd considérablement de sa force et cette position ne s'impose nullement. Pour Kant, il n'est possible de démontrer ni la liberté, ni l'absence de liberté. En montrant que le déterminisme biologique n'est pas tenable, la science contemporaine affaiblit considérablement la position déterministe et renforce considérablement, sans toutefois la démontrer de manière apodictique, la position d'une liberté effective.

En fait, l'interprétation des positions edelmaniennes en contexte kantien nous conduit à affirmer que la liberté n'est pas nécessairement illusoire. Si des modèles peuvent montrer qu'un déterminisme biologique strict n'est pas défendable, la thèse d'un déterminisme environnemental se voit profondément affaiblie. Un déterminisme environnemental strict reste pensable mais il ne s'impose en aucune manière. La crédibilité de modèles intégrant conscience intentionnelle et libre arbitre se voit singulièrement renforcée. Spinoza n'a pas (nécessairement) raison!8

Les neurosciences contemporaines proposent des théories qui remettent en cause de manière radicale le déterminisme biologique. Avec de telles perspectives, la voie est ouverte pour une réconciliation des sciences biologiques avec les sciences humaines. La voie est ouverte pour une réconciliation des sciences biologiques avec la dignité de l'humain.

⁸ Pour une étude plus détaillée de la question : B. Feltz, « Neurosciences et anthropologie », in M. Delsol, B. Feltz, et M.C. Groessens, *Intelligence animale*, intelligence humaine, Vrin, Paris, 2008, 7-40.

Neurosciences et théologie de la foi-confiance

Alexandre GANOCZY¹

Lorsque nous avons décidé avec Marc Jeannerod, il y a plus d'un an, de mettre en chantier une recherche interdisciplinaire sur l'impact des neurosciences sur la théologie, nous avons choisi comme sujet « la dialectique confiance/défiance ». Ce sujet nous est apparu d'une singulière actualité, particulièrement en France où s'est installé ce que d'autres chercheurs ont pu appeler une « société de défiance ». Nous avons consulté également des travaux de psychologues, de psychanalystes et de philosophes, puisque le sujet fait depuis longtemps, également dans ces disciplines, objet de recherche.

Le théologien avait plusieurs raisons d'y participer. Premièrement, parce que le christianisme bien compris définit la foi en Dieu, sur la base de ses sources bibliques, comme un acte de confiance fait au Fiable par excellence. Deuxièmement, parce que la révolution anthropologique apportée par les neurosciences rendait impossible d'en rester à une conception dualiste de l'homme, celle de l'âme et du corps compris comme des substances ou des systèmes séparables, un dualisme qui se rattache non seulement à Platon et à Descartes, mais ne cesse de s'exprimer jusque dans les documents récents de l'enseignement religieux tant catholique que protestant. Or, à l'ère de la pensée scientifique, c'est une anthropologie holistique qui s'impose, une pensée pour laquelle tout ce qui est mental en nous est entièrement « incarné » et tout ce qui est corporel en nous se « mentalise » à partir de notre cerveau. Nous sommes donc invités à nous comprendre, pour parler avec le philosophe Heimich Rombach, comme des structures vivantes dont les éléments physiques et psychiques ne peuvent exister qu'ensemble et comme un tout, en grec un « holon » d'où l'adjectif holistique.

C'est dans ce contexte que Marc Jeannerod a étudié la confiance en mettant à contribution ce que les neurosciences sociales disent de l'empathie et

¹ Théologien, Paris

de la théorie de l'esprit, ou mieux « theory of mind ». Je vais donc reprendre ces mêmes catégories qui ont d'ailleurs aussi une portée psychologique et philosophique, afin de proposer quelques éléments d'une nouvelle théorie de la foi-confiance

1. Communication directe/indirecte et foi-confiance

Les neurosciences sociales et la psychologie cognitive situent le thème de la confiance et de la méfiance ou de la défiance dans l'aire de la communication interindividuelle dont elles distinguent clairement deux étapes. l'une directe, l'autre indirecte. La première se fait principalement par l'intermédiaire du corps, plus exactement par des signaux sensoriels émis, recus, codés et décodés par les réseaux neuraux dont l'organe central est le cerveau. Pour cette raison, ces phénomènes sont privilégiés par les neurobiologistes et sont l'objet d'une approche expérimentale et clinique. Par contre, la communication indirecte a la préférence de la psychologie cognitive qui met en œuvre une approche que l'on dit « mentaliste ». Le théologien qui cherche à décrire la foi-confiance dans des termes empruntés aux neurosciences tiendra compte de ces deux approches. La première lui servira à établir que la foi est un comportement « incarné » ou comme dit Gerald Edelman « embodied », c'est-à-dire enraciné dans notre système nerveux et dans nos mécanismes sensoriels. L'approche mentaliste l'intéresse en revanche parce qu'elle l'aide à concrétiser comment la foi en Dieu n'est pas une sorte de réflexe conditionné, mais un acte commandé aussi par des facteurs cognitifs, rationnels et des actes volitifs libres qui sont aptes - pour emprunter une expression à Antonio Damasio - « à s'élever au-dessus des diktats des dispositions neuronales »².

Mais quelles sont les caractéristiques de la communication directe et indirecte? La communication directe est spontanée ou même automatique. Elle se produit comme la réaction immédiate à la présence d'un autre individu qu'on voit, entend ou touche ici et maintenant. Le plus souvent les paroles, les

_

² A. Damasio, Le sentiment même de soi : corps, émotions, conscience, Paris, Odile Jacob, 1999, p. 295

mimiques, les gestes du communicant vis-à-vis d'un autre, qui doit par ailleurs être physiquement présent, sont incontrôlés, irréfléchis. C'est à ce niveau que la théorie récente des neurones miroirs³ trouve sa signification: la communication directe et spontanément réactive d'un sujet reflète dans ses propres neurones le comportement, ou les actions de cet autre. Mais cette relation est, pour ainsi dire, tellement viscérale qu'elle ne peut pas être porteuse de sens. Elle se laisse observer par exemple dans le comportement d'un bébé à l'égard de sa mère, ou chez une personne qui se laisse emporter par l'enthousiasme en présence d'un orateur fascinant.

Et pourtant, le théologien serait mal inspiré s'il considérait une telle relation directe comme sans aucune importance pour la foi-confiance authentique. Car elle peut jouer le rôle d'une amorce dans la genèse d'une relation indirecte qui se concrétise dans la foi. On peut supposer que Hans Küng fait intervenir un peu dans le même sens la notion de « confiance fondamentale »⁴, traduisant le « basic trust » d'Erik Erikson. En tous cas, une certaine phénoménologie de la communication directe peut être utile lorsqu'il s'agit de concrétiser l'incarnation de la confiance/méfiance dans nos mécanismes biologiques et neuraux.

Venons-en à la description de la communication indirecte. Sa première caractéristique est qu'elle suppose une médiation. Elle est moins spontanée, mais se montre en grande partie réfléchie. Jeannerod pense que la relation indirecte est « à sens unique », mais il ajoute: « au moins dans l'immédiat ». Cette précision a de quoi intéresser le théologien. Du moment en effet qu'il s'agit d'un acte intentionnel qui s'adresse par définition à tel individu ou à tel groupe, il suscite comme naturellement une réaction ou une réponse de ses destinataires, lesquels peuvent être même absents ou seulement à venir. La notion de « tiers absents » qu'on introduit ici ne renvoie-t-elle pas à cette possibilité? Si tel est le cas, la communication indirecte, médiatisée,

³ Rizzolatti, G. et al. (1996) Premotor cortex and the recognition of motor actions Cognit. Brain Res. 3, 131–141; *Les neurones miroirs*, G Rizzolatti, L Folgassi, V Gallese, *Pour la Science*, Janvier 2007, p 44-49

⁴ Hans Küng *Faire confiance à la vie, 2010*, Editions du Seuil

comporte, du moins en puissance, un élément de réciprocité. Elle peut générer un rapport entre message émis et reçu, entre appel adressé et réponse à cet appel, entre transmission effectuée et perçue, entre mémoire exprimée et mémoire reconnue, mais aussi entre vérité affirmée par l'auteur de la communication et vérité vérifiée par le récepteur.

L'idée que la communication indirecte soit intentionnelle et qu'elle puisse s'adresser à un « tiers absent » apporte de l'eau au moulin du théologien. Car est-il aberrant de penser que ceux qui ont communiqué l'Évangile de manière indirecte, par l'intermédiaire de témoins et de messagers, l'ont fait en le destinant à tout un monde de « tiers absents » ? Absents parce que futurs. Ne doit-on pas considérer de la même façon le processus qui a conduit la communauté chrétienne à se doter de sacrements ? De sacrements d'initiation et de célébration qui ont vocation de re-présenter, c'est-à-dire de rendre présentes toujours à nouveau les communications majeures des fondateurs ? N'est-ce pas grâce à ces médiatisations que l'appel originel à la foi-confiance en Dieu est devenu accessible aux hommes de tous les temps qui dès lors pouvaient y répondre en croyant à leur tour?

De toute évidence ces communications incarnées dans des réactions spontanées des premiers croyants, puis transmises à d'autres croyants, ne consistaient pas seulement dans un processus cognitif portant sur les paroles, les actions et la personne du Christ mais elles sollicitaient en même temps l'affectivité des destinataires. C'est ainsi qu'on peut parler, en utilisant une autre notion clé des neurosciences sociales, de leur empathie envers le Christ qui leur était pourtant temporellement et spatialement absent.

2. Contrôle d'empathie et foi-confiance raisonnée.

Empathie... Ce terme, on le sait, est la traduction du mot allemand « Einfühlung » introdui à la fin du XIX^e siècle par le philosophe Theodor Lipps et repris ensuite par Freud. Il désigne un ensemble de pulsions affectives qu'éprouve, d'une manière spontanée, une personne en réagissant à l'état interne émotionnel d'une autre personne, tel qu'il se laisse lire le plus souvent à l'expression de son visage. Par exemple un état de peur, de joie, de souffrance,

de curiosité. Cette lecture de la personne empathique s'accompagne d'une interprétation improvisée et peu réfléchie à partir de ses propres expériences. La personne A projette ainsi son état psychique sur celui de la personne B et cette projection va jusqu'à un rapport de simulation et d'identification entre elles. Comme c'est A qui domine dans ce processus d'identification affective, le risque d'un certain subjectivisme ou, en tout cas, d'erreur est inévitable. A s'expose au danger de manquer la réalité objective de B, de le prendre pour autre qu'il n'est réellement, de l'imaginer entièrement à sa propre image. On sait que Feuerbach s'est appuyé sur cette faillibilité de l'empathie spontanée pour déclarer que la foi chrétienne n'avait aucun statut de réalité puisque son objet n'était que pur reflet de nos désirs insatisfaits.

Il est intéressant de noter que Max Scheler critiquait Lipps déjà avec des arguments semblables. Il qualifia sa théorie de l'empathie de « projection affective »⁵. Tout en reconnaissant que cette attitude spontanée pouvait fort bien amorcer une relation intersubjective et être fondée sur du réel, il mettait l'accent sur son caractère foncièrement égocentrique. Car lorsque A attribue à B les mêmes sentiments qu'il éprouve lui-même, il ne peut jamais arriver à sa connaissance « adéquate », surtout pas à celle de sa « sphère intime »⁶. Pour y arriver, il est indispensable que l'empathie se prête à un contrôle à l'aune de critères plus objectifs.

C'est ici que peut être introduite la question suivante : Étant donné que les neurosciences ont intégré une certaine théorie de l'empathie et qu'elles parlent même de « cerveau empathique », selon quels critères et avec quelles réserves le font-elles? Une solution du problème apparaît possible si l'on rejoint la critique de Scheler, et si l'on distingue « empathie de bas niveau » et « empathie contrôlée ». En quoi consiste ce contrôle ? Dans l'intégration au processus subjectif d'un élément d'objectivité. Ce qui veut dire que A en se tournant vers B renonce à prendre pour seul point de départ son propre vécu et son ressenti subjectif et devient d'emblée attentif à l'altérité de B. Il cherche à

-

⁵ Max Scheler, *Nature et formes de la sympathie. Contribution à l'étude des lois de la vie affective*, Payot et Rivages, 2003, p. 60.

⁶ *Idem* p. 56.

connaître B tel qu'il est en réalité. Il ne se contente pas de l'imaginer. Il veut savoir en quoi il est différent. Au lieu de lui attribuer du sien propre et de s'identifier mentalement avec lui, au lieu de l'imaginer à l'image de sa propre identité, il respecte sa différence et aussi la distance qui le sépare de lui. Dès lors son affectivité n'envahit pas l'autre mais se contente de se positionner par rapport à l'autre.

Sur ce point encore il me semble éclairant de comparer Lipps à Scheler. Le premier emploie comme terme clef « Einfühlung » tandis que le second opte pour « Mitgefühl ». « Einfühlung » indique que A entre pour ainsi dire au-dedans de B, se laissant pousser par ce qu'il éprouve lui-même. C'est un mouvement de pénétration, voire même d'envahissement allant jusqu'à la fusion. En revanche « Mitgefühl » signifie que A s'associe affectivement avec B, se rend proche de lui, sans l'intention de l'envahir. C'est exactement la signification de la notion de « sym-pathie », notion que Scheler préfère à « empathie ». A ses yeux, la sympathie - je cite - « consiste à s'abstraire de soimême, à se dépasser, pour se mettre résolument en présence de l'autre et de son état psychique individuel »⁷.

Jeannerod me semble vouloir dire la même chose quand il parle d'« empathie contrôlée ». Contrôlée par quoi ? Par des états mentaux indépendants de considérations et de sentiments subjectifs. Autrement dit : par une démarche rationnelle qui comporte des questionnements sur les raisons de l'empathie de A pour B, exprimée par telle parole ou tel geste. Cela peut aboutir à un débat ou du moins à un échange d'explications dans le cadre d'un dialogue. Une telle jonction entre le sentiment et le raisonnement est pour la psychologie cognitive indispensable pour qu'un comportement altruiste, amorcé seulement par l'empathie spontanée et unilatérale puisse atteindre le niveau de la réciprocité. En ce qui concerne la confiance mutuelle entre A et B, elle doit s'affirmer dans un processus similaire de contrôle et de vérification, un processus qui ne redoute pas la contradiction ou les objections suggérées par une méfiance rationnelle. Depuis Bachelard nous parlons dans ce contexte

⁷ *Ibidem* p. 117.

de « confiance raisonnée », parce que provoquée et aidée par une « défiance rationnelle ».

J'en viens à mon hypothèse de travail théologique. La foi-confiance en Dieu peut être comprise comme un acte d'empathie contrôlée. Si elle est chrétienne, elle s'adresse fondamentalement au Christ en tant que révélation réelle, parce qu'historiquement accessible, d'un Dieu qui aime l'homme, dont la « philanthropie » selon l'épître à Tite (3,4) s'est vérifiée dans le prophète de Nazareth. Cette foi dépasse le schéma religieux général qui comporte des éléments irrationnels de crédulité et de croyance. La foi-confiance nous invite à croire en quelqu'un de suffisamment identifié, pour ainsi dire « en connaissance de cause ». Elle s'actualise comme acte à la fois cognitif, volitif et affectif. C'est donc une démarche psychologiquement holistique, engageant notre être personnel et social tout entier. La foi-confiance se montre aussi « empathie contrôlée », par le fait qu'elle accepte d'être interrogée, critiquée, mise en doute, contredite, discutée, exposée à la méfiance. Elle ne cherche pas à être sécurisée derrière les murailles d'un dogmatisme rigide. Elle a pour ainsi dire le goût de la compétition sportive dans notre monde de pluralisme philosophique et religieux.

La théologie moderne qui se comprend comme la science ayant pour objet le contenu de cette foi-confiance peut être comprise elle-même comme un outil de contrôle. C'est la raison pour laquelle elle assigne un rôle capital à la méthode historico-critique des textes bibliques.

S'il m'est permis d'ajouter encore une seconde hypothèse de travail, je dirai ceci.

L'élaboration d'une théorie de la foi-confiance qui s'appuie sur des thèses des neurosciences sociales semble rendre possible une reformulation de notre conception de la divinité même de Dieu. On peut se demander pourquoi, si la tradition biblique a pu se représenter Dieu comme Dieu d'alliance sur la base d'une analogie avec le lien conjugal et si les écrits johanniques du Nouveau Testament ont fait de même pour affirmer que Dieu était amour à la ressemblance de personnes qui se donnent sans réserve l'une à l'autre, il nous serait interdit d'appliquer à Dieu lui-même l'analogie de l'empathie contrôlée,

ou plutôt vérifiée. Une empathie ou sympathie vis-à-vis des humains qui s'est vérifiée par le fait même de son incarnation dans l'homme de Nazareth, autant dans son comportement altruiste que dans ses paraboles proches de la vie. Puis ne pourrait-on aller jusqu'à éclairer la structure trinitaire de Dieu en usant de l'analogie de l'empathie/sympathie réciproque? Une telle démarche a été adoptée par des théologiens importants des cinq premiers siècles, par Basile de Césarée en Orient et par Augustin en Occident puis au Moyen Âge par Richard de Saint Victor et Bonaventure. Ils considéraient que l'unité de Dieu n'était ni numérique ni monolithique ni monarchique, mais une unité plurielle résultant de la communication des trois sujets nommés par analogie « personnes ». La raison de cette idée communautaire de Dieu résidait dans la conviction bibliquement établie que l'être divin en lui-même est « empathique » et « sympathique ». Ou'il est agapè ou caritas, c'est-à-dire don de soi mutuel et « contrôlé », si j'ose dire, par des relations de réciprocité. C'est à partir de cette idée qu'un théologien protestant du XXe siècle, Erik Peterson, a pu affirmer que la Trinité pouvait servir de modèle à une société démocratique. Mais revenons aux acquis des neurosciences sociales.

3. « Theory of mind » et foi-confiance

Dans le but d'étudier les relations intersubjectives où se joue la dialectique confiance/méfiance les neurosciences proposent - rappelons-le - deux méthodes complémentaires. La première que je viens d'esquisser est appelée à juste titre « approche par le corps ». Elle consiste à aborder la question par une analyse des signaux émis par notre cerveau lorsque nous avons des relations sociales à autrui, notamment dans un acte de confiance ou de méfiance. La seconde est nommée « approche mentaliste », parce qu'elle aborde la même question par une observation de notre comportement mental. C'est ici qu'on parle de « Theory of mind » ou « théorie de l'esprit ».

Le théologien a des raisons de préférer l'expression anglaise. Car la notion « mind » a une signification que ne recouvre guère le mot « esprit », surtout s'il ne se trouve précisé par aucun adjectif. « Esprit » reste opposable à « corps » et à « matière ». Il se prête donc à une interprétation dualiste ou

spiritualiste. En revanche « mind » non seulement ne s'oppose pas à ce qui est matériel, physique, biologique, neural, mais y est incarné, « embodied », comme le dit Gerald Edelman. Il est avec le corps l'un de ces deux « moments » de la structure humaine qui ne peuvent exister qu'ensemble et comme un tout. Même dans la mort, le « mind » ne doit pas être nécessairement pensé comme séparé du corps.

La condition « incarnée » du « mind » ne veut pas dire obligatoirement qu'il soit entièrement déterminé par des automatismes neuraux. Ou encore que la conscience soit une simple sécrétion du cerveau sur le modèle d'un autre organe comme le rein qui produit l'urine. Ici les opinions se séparent. Au réductionnisme matérialiste de Patricia Churchland s'oppose Antonio Damasio qui affirme que notre conscience possède « l'aptitude à s'élever au-dessus des diktats des dispositions neuronales »⁸.

Ceci dit, que signifie « Theory of mind » ? Contrairement aux apparences, il ne s'agit pas d'une thèse ou d'une doctrine sur le « mind », mais d'une action, d'un comportement du « mind ». L'expression rejoint le sens originel du verbe grec « théôrein » qui signifie observer, contempler, examiner, passer en revue⁹. Comment fonctionne la « Theory of mind » (en abrégé TOM) ? Quelle est sa visée ?

Se manifestant chez l'enfant dès sa troisième ou quatrième année, elle consiste dans un double regard qui va vers autrui et revient vers soi. En allant à l'autre, le sujet prend conscience de cet autre tout en devenant de plus en plus conscient de lui-même. Ce va-et-vient du regard, où le sujet observe autrui tout en s'observant lui-même, vise une connaissance, partant une compréhension d'un partenaire potentiel. Pour ce faire, il prend son temps. La compréhension de l'autre et du soi se réalise au moyen d'une longue série de comparaisons, d'attributions et de projections. Ce qui comporte aussi un risque d'erreur ou de partialité dans le jugement. Les informations recueillies par la TOM, de même que l'écoute qui accompagne leur perception visuelle, peuvent à tout moment s'avérer fausses ou inadéquates. C'est pourquoi une « vigilance épistémique »

⁸ A. Damasio, op. cit., p. 295.

⁹ Anatole Bailly, *Dictionnaire Grec-Français*, Hachette 1901.

est nécessaire pour que la rencontre entre la conscience d'autrui et la conscience de soi soit profitable aux deux.

Le danger que la TOM frappe dans le vide provient du fait que son auteur peut attribuer à autrui seulement les mêmes états mentaux que les siens propres. Qu'elle ne prend pas suffisamment conscience de la différence de l'autre. Qu'elle projette sur lui quelque chose qu'il n'est pas en réalité.

Cette faillibilité de la TOM se manifeste aussi lorsqu'elle s'actualise dans la création d'une relation de confiance. Voilà pourquoi il lui est utile d'intégrer d'emblée une part de méfiance et de doute servant à vérifier la fiabilité d'un partenaire. De cette règle critique la foi-confiance en Dieu n'est pas non plus exempte. Là aussi le croyant est invité à faire la part des choses, à distinguer entre ce qui est projection au sens de Feuerbach et ce qui est perception correcte du réel. Dans le cas contraire, on risque de tomber dans la crédulité ou de prendre ses propres désirs pour la réalité. Alors l'idée biblique de Dieu et du Christ se trouve déformée, altérée, livrée à l'arbitraire du sujet qui n'est que trop tenté de composer sa religion un peu à la manière d'une personne qui fait ses courses au supermarché.

À côté de leur souci de fiabilité et d'authenticité, ceux qui exercent la TOM en vue d'une communication fructueuse, notamment en vue de relations de confiance mutuelle, il faut relever le souci de contemporanéité entre l'observant et l'observé. Faut-il que l'objet de l'observation soit physiquement présent, qu'on puisse le voir, l'entendre ici et maintenant? Faut-il qu'il soit directement joignable? Ou bien peut-il être atteint aussi indirectement, grâce, par exemple, à une chaîne de médiations ou à une succession de témoins qui traverse les siècles? Le neuroscientifique répond par l'affirmative lorsqu'il développe sa théorie de la méta-représentation. Celle-ci n'est autre qu'une représentation de représentations. Il ne s'agit pas seulement de représentations d'états mentaux de congénères physiquement et temporellement présents, mais aussi de pensées et d'actions de personnes qui ont vécu dans le passé ou qui vivront dans l'avenir. On les re-présente au sens de les rendre présentes. Les moyens en sont multiples : commémoration, récit, ou anticipation d'un futur réellement possible. La TOM peut intégrer la réalité historique dans sa totalité,

y compris ce qu'on ne peut que prédire. La narrativité lui est connaturelle. C'est en cela que l'approche mentaliste complète utilement l'approche par l'analyse des réactions cérébrales : elle entre dans le monde des idées et des pensées par le biais de leur transmission historique.

Cela peut encore intéresser le théologien qui est responsable de la tradition véhiculant le contenu de la foi-confiance, autrement dit l'Évangile. En recourant à l'exégèse historico-critique, il se rapproche mentalement de ce que l'homme de Nazareth a véritablement dit, pensé, fait en le distinguant de ses revêtements légendaires et mythiques. Ainsi il a la possibilité de faire la jonction entre les états mentaux de Jésus et de ses disciples et les états mentaux des hommes d'aujourd'hui. Dans la mesure où le pouvoir méta-représentatif de la TOM est pris en considération dans une théologie de la foi-confiance il rend possible que des croyants du XXI^e siècle puissent rejoindre la foi des apôtres. L'approche mentaliste des origines chrétiennes permet aussi de ressourcer le message évangélique tout en l'incarnant dans les idées maîtresses de la modernité et d'échapper ainsi à une certaine anarchie intellectuelle et morale de la postmodernité.

La TOM bien comprise n'a pas seulement la fonction d'un « théôrein » informatif et théorique, elle a aussi des conséquences pratiques, éthiques et thérapeutiques.

4. Conclusions

4.1. La réflexion interdisciplinaire sur quelques aspects importants de l'humanité de l'homme que je viens de tenter s'inscrit dans une vaste entreprise de déconstruction de l'anthropologie dualiste, entreprise qui est en cours déjà depuis plusieurs décades en philosophie et sporadiquement même en théologie. La théologie chrétienne tant catholique que protestante a été imprégnée pendant des siècles par l'héritage de Platon et de Descartes, jusqu'à mettre en veilleuse la conception holistique de l'être humain que présentent les textes majeurs de la Bible. L'idée que l'homme est composé de deux substances séparables, l'âme et le corps, dont seule la première doit dominer et être vouée à l'immortalité, a pu générer certes une civilisation dont la grandeur reste

incontestable. Mais avec l'avènement de la modernité elle a contribué aussi à une rupture entre des courants de pensée matérialistes et spiritualistes, les premiers réduisant l'homme à un ensemble de fonctions biologiques, les seconds réduisant dans l'homme tout ce qui est animal, charnel, sexuel, à un niveau inférieur. La reconnaissance de la théorie de l'évolution comme « plus qu'une hypothèse » par Jean-Paul II aurait pu constituer un tournant antidualiste dans la doctrine officielle de l'Église catholique si elle n'avait pas exempté l'âme spirituelle de ses conditionnements évolutifs en attribuant son origine à un acte créateur direct de Dieu hors évolution¹⁰, ce qui encore une fois ne correspond pas au véritable sens du récit symbolique de la Genèse concernant la création d'Adam à partir de la glaise de la terre. Aujourd'hui la théologie, qui n'est une science que lorsqu'elle fait de la recherche, a une opportunité pour sortir du vieux dilemme, en partie grâce aux neurosciences qui ont salutairement bouleversé, non pas « l'humanité », mais une certaine conception de l'humanité de l'homme.

4.2. J'ai essayé de montrer que la notion clef permettant ce changement radical de perspective pourrait être celle de l'incarnation, concept d'origine théologique qu'emprunte, par exemple, Gerald Edelman pour parler en neuroscientifique de l' « embodiement of the mind ». Ce qui signifie que tout ce qui est en nous mental est entièrement incorporé et tout ce qui est en nous corporel ne cesse de se mentaliser à partir de notre cerveau. Cette constitution interactive de la personne ne justifie pas une réduction matérialiste de notre être, puisque, selon le schéma holistique, le moment mental et le moment physique de la structure que nous sommes ne peuvent exister qu'ensemble et

¹⁰ Aujourd'hui, près d'un demi-siècle après la parution de l'Encyclique, de nouvelles connaissances conduisent à reconnaître dans la théorie de l'évolution plus qu'une hypothèse. Il est en effet remarquable que cette théorie se soit progressivement imposée à l'esprit des chercheurs, à la suite d'une série de découvertes faites dans diverses disciplines du savoir. (...) C'est en vertu de son âme spirituelle que la personne tout entière jusque dans son corps possède une telle dignité. Pie XII avait souligné ce point essentiel : si le corps humain tient son origine de la matière vivante qui lui préexiste, l'âme spirituelle est immédiatement créée par Dieu Jean-Paul II Académie Pontificale des Sciences, le 22 octobre 1996, texte intégral sur

http://www.hominides.com/html/theories/jean_paul_evolution.php

comme un tout. On peut reprendre aussi l'affirmation de Damasio selon laquelle nos actes affectifs, cognitifs, volitifs libres gardent au sein de leur « embodiement » l'aptitude « à s'élever au-dessus des diktats des dispositions neuronales ». Je note enfin que la perspective holistique permet la dualité de notre approche par le corps et le mental sans tomber dans un dualisme méthodologique.

- **4.3.** C'est dans la logique de ce qui vient d'être dit que j'ai essayé d'amorcer un dialogue interdisciplinaire en empruntant des catégories familières aux neurosciences cognitives et sociales pour esquisser une nouvelle théologie de la foi-confiance en Dieu: communication directe et indirecte, empathie spontanée et contrôlée, theory of mind.
- **4.4.** J'ai essayé d'indiquer comment la dialectique confiance/méfiance se développe en passant par plusieurs niveaux avant de pouvoir devenir foiconfiance capable de transformer même la méfiance en facteur critique à l'intérieur d'un acte de confiance de haut niveau qui est porteur de sens.

Au premier niveau ce qui deviendra une attitude mentale confiante s'incarne dans la communication directe spontanée, irréfléchie, voire automatique. C'est ici que les neurones miroirs jouent leur rôle. Mais c'est sans doute aussi le lieu du « basic trust » qui se présente comme l'ébauche sensorielle d'un état de confiance en cours de devenir consciente. Le deuxième niveau est représenté par la communication directe où une série de médiations exerce une fonction décisive. Ici s'exerce la jonction entre mémoire reçue et mémoire personnelle et s'amorcent des relations de confiance ou de méfiance réciproques.

La réciprocité trouve son moteur dans l'affectivité de sujets concernés. Le cerveau empathique contribue au développement de la confiance. Moins au sens de la « projection affective » prônée par Lipps qu'au sens de la « participation affective » qui correspond à la sympathie selon Scheler. Là où règne cette sympathie entre des personnes en communication indirecte, ce sont des actes conscients qui contribuent à créer des états mentaux confiants.

Au niveau de la TOM où se produit une interaction fructueuse entre la conscience de soi propre et la conscience de soi d'autrui, la dialectique

confiance/méfiance a les meilleures chances de s'enrichir d'une nouvelle dimension : celle de la foi-confiance faite à Dieu reconnu simultanément comme le Tout Autre et comme le Fiable par excellence.

4.5. Selon la tradition chrétienne, la fiabilité de ce Dieu trouve sa vérification dans la vie, l'action et l'enseignement de l'homme de Nazareth, qu'il convient d'appeler dans un sens spécifique l'incarnation d'un Dieu analogiquement empathique, sympathique, à qui l'épître à Tite (3,4) applique à juste titre la notion de « philanthropie », amour inconditionnel pour tout ce qui est humain.

J'ai évoqué la possibilité de comprendre dans ce contexte la divinité de Dieu Trinité comme unité plurielle commandée par une relation de confiance mutuelle entre les Trois que nous appelons analogiquement personnes divines.

4.6. Plusieurs questions sont restées, par la force des choses, ouvertes, ou ont reçu seulement des amorces de réponse. Par exemple celle de l'absence spatiale et temporelle de Dieu par rapport au croyant, celle de la projection sur lui de nos états d'esprit spécifiquement humains, celle de la transmission de la foi-confiance à travers les âges, ou encore celle de l'interdisciplinarité difficile que nous avons tenté de mettre en œuvre. Mais qui ne risque rien ne fait rien. La recherche continue.

Neurosciences + *éthique* = *neuroéthique*

Bernard Baertschi¹

La naissance de la neuroéthique

Neil Levy a sous-titré son dernier ouvrage, Neuroethics, « *Défis pour le XXI*^e siècle »². C'est indiquer que l'essor récent des neurosciences se doit d'être accompagné par une réflexion éthique, et que celle-ci n'en est qu'à son commencement. En effet, c'est seulement depuis le début des années 2000 qu'on rencontre de manière insistante le vocable « neuroéthique ». Ce qui a déclenché cet usage et l'intérêt pour la discipline qu'il désigne, c'est un congrès fondateur tenu à San Francisco sous l'égide de la Fondation Dana en 2002, où William Safire, éditorialiste au *New York Times*, a employé ce terme³

Qu'est-ce que la neuroéthique ? Leon Kass, qui a présidé le *Conseil du président étasunien pour la bioéthique* entre 2002 et 2005, en parle comme de la discipline qui «embrasse les implications éthiques des avancées en neuroscience et en neuropsychiatrie»⁴. Cela reste assez général, d'autant qu'il est notoire que le terme « éthique » est souvent employé de manière large – chez certains auteurs même, tout ce qui n'est pas de la compétence du scientifique ressortit à l'éthique. Adina Roskies a tenté avec bonheur de mieux articuler le concept. Selon elle, la neuroéthique comprend deux domaines: l'éthique des neurosciences et la neuroscience de l'éthique, le premier divisé lui-même en éthique de la pratique des neurosciences et étude des implications éthiques des neurosciences⁵:

¹ Philosophe et éthicien, Université de Genève

² N. Levy, Neuroethics. Challenges for the 21st Century, Cambridge, CUP, 2007.

³ Les actes de ce Congrès ont été publiés par Steven Marcus, *Neuroethics. Mapping the Field*, New York, Dana Foundation, 2002.

⁴ President's Council on Bioethics, 15 janvier 2004,

http://bioethics.georgetown.edu/pcbe/transcripts/jan04/session3.html.

⁵ A. Roskies, «A Case Study of Neuroethics: the Nature of Moral Judgment», in J. Illes, dir., *Neuroethics*, Oxford, OUP, 2006, p. 18.

1° L'éthique de la pratique des neurosciences est simplement l'éthique de la pratique scientifique appliquée aux sciences du cerveau; elle consiste par exemple en la formulation de principes moraux pour guider la recherche et les traitements, afin d'assurer notamment le respect de l'autonomie des sujets de recherche et des patients, ainsi que la préservation de la confidentialité et de leur vie privée.

2° L'étude des implications éthiques des neurosciences concerne les effets des progrès de la connaissance du cerveau sur nos conceptions sociales, éthiques et philosophiques. Les neurosciences étudient le cerveau, c'est-à-dire un organe qui a ceci de particulier qu'il joue un rôle important pour notre identité personnelle; les implications éthiques de son étude débordent par conséquent largement les questions strictement normatives, s'étendant à des sujets comme l'existence de l'âme ou du libre arbitre.

3° La neuroscience de l'éthique enfin concerne l'approche scientifique de notre comportement moral, notamment par le moyen de l'imagerie cérébrale : que se passe-t-il dans notre cerveau lorsque nous raisonnons moralement et prenons des décisions ? Cette approche scientifique concerne en priorité l'être humain, mais elle est aussi utilisable pour l'animal, dont on peut ainsi étudier les attitudes mentales.

Dans cet article, je ne m'occuperai que des deux derniers sousdomaines, car ce sont eux qui ont un impact sur « l'humanité de l'homme », et sont à même, peut-être, de la « bouleverser », pour reprendre le titre du Colloque.

L'intérêt du public pour la neuroéthique – alors même que ce terme n'était pas encore utilisé – a été aiguillonné par le cas de Phineas Gage, raconté et étudié par Antonio Damasio dans son livre *L'Erreur de Descartes*, paru en 1994 et traduit en français l'année suivante⁶,^{7.} Gage, ouvrier des chemins de fer américains, spécialisé dans l'utilisation des explosifs pour faire sauter les rochers qui entravaient l'avancée de la pose des voies, a eu le crâne perforé de part en part par une barre de métal à la suite d'une erreur de

_

⁶ A. Damasio, L'erreur de Descartes, Paris, Odile Jacob, 2001.

⁷ Ce cas est également rapporté par J-P. Changeux, *L'homme neuronal*, chap. V, 1983.

manipulation. Il a perdu un œil, mais a survécu. Aucune de ses facultés ne semblait altérée, si ce n'est que son caractère était affecté: « Ces changements étaient devenus apparents dès la fin de la phase aiguë de la blessure à la tête. Il était à présent "d'humeur changeante ; irrévérencieux ; proférant parfois les plus grossiers jurons (ce qu'il ne faisait jamais auparavant) ; ne manifestant que peu de respect pour ses amis ; supportant difficilement les contraintes ou les conseils, lorsqu'ils venaient entraver ses désirs ; s'obstinant parfois de façon persistante ; cependant, capricieux, et inconstant" »⁸, rapporte son médecin, qui commente : « l'équilibre, pour ainsi dire, entre ses facultés intellectuelles et ses pulsions animales » avait été aboli.

D'où ce commentaire d'Albert Jonsen: « Gage vivait avec des capacités physiques intactes et des facultés cognitives en bon état - bien qu'avec une exception importante : il était devenu incapable de faire des choix moraux »⁹. Oue devient alors sa responsabilité? Phineas Gage est la parfaite illustration du fait qu'une lésion cérébrale peut modifier en profondeur notre caractère, au-delà de toutes nos intentions, de toute notre volonté, et avec d'importantes conséquences. Je viens de mentionner la responsabilité, mais ce n'est pas la seule question neuroéthique que le cas de Gage suscite. Le changement profond de son caractère oblige à se demander s'il est toujours la même personne et, au vu de la déconnexion de ses facultés cognitives et affectives relevées par son médecin, se pose à nouveau une vieille question de philosophie morale : quels sont les rôles respectifs de la raison et des émotions dans notre vie morale? Depuis le XVIIIe siècle au moins, à l'exception de certains philosophes comme David Hume, nous pensons qu'une décision morale doit être rationnelle et qu'il ne faut pas faire appel à nos émotions pour choisir bien. Kant notamment estime que nos émotions sont de très mauvaises conseillères, et qu'il faudrait tenter, si cela était possible, de les éradiquer : « Les inclinations mêmes [...] ont si peu une valeur absolue qui leur donne le droit d'être désirées pour elles-mêmes, que, bien plutôt, en être pleinement

⁸ A. Damasio, *L'erreur de Descartes*, p. 27.

⁹ In Neuroethics. Mapping the Field, p. 11.

affranchi doit être le souhait universel de tout être raisonnable. »¹⁰ Le cas de Gage ne nous oblige-t-il pas à revoir cette conception?¹¹

L'identité personnelle, le rôle des émotions dans l'éthique et la responsabilité morale sont trois suiets que les neurosciences nous demandent de reconsidérer; elles appartiennent à ce que Roskies nommait « l'étude des implications éthiques des neurosciences » Ce ne sont pas les seules. La question de l'existence de l'âme elle aussi refait surface: si, grâce à l'imagerie cérébrale, nous pouvons regarder ce qui se passe dans notre cerveau et comprendre son fonctionnement, reste-t-il encore quelque chose à expliquer qui demanderait le recours à une substance immatérielle? Si, au moyen de l'imagerie cérébrale, nous pouvons « lire dans l'esprit », comme disent certains, qu'allons-nous y découvrir de nous-mêmes? Les neuropsychologues utilisent déjà largement cette imagerie pour étudier comment notre mental fonctionne, aux niveaux cognitif, affectif et éthique (c'est « la neuroscience de l'éthique » dont parlait Roskies), progressant dans notre compréhension de l'humain. Ces recherches touchent aussi à la question des implications éthiques des neurosciences. Pensons à la médecine prédictive, à ce que certains nomment la phrénologie préventive (détecter les propensions à des activités criminelles chez certaines personnes et agir en conséquence – comment?), la détection des mensonges (qui intéresse fortement certains juristes et avocats) et l'évaluation des patients en état végétatif.

Je ne peux clore ce rapide survol des questions importantes qui se posent pour la neuroéthique sans mentionner la neuroamélioration, à savoir ce que les Anglo-saxons nomment *human enhancement*. En faisant progresser nos connaissances sur le cerveau, les neurosciences augmentent aussi notre pouvoir; nous voici alors en état d'améliorer les capacités de l'être humain. Est-ce quelque chose de désirable ou plutôt d'effrayant? Les désaccords sont profonds sur cette question, comme on le verra.

_

¹⁰ E. Kant, Fondements de la métaphysique des mœurs, Paris, Vrin, 1980, p. 104.

¹¹ Sur cette question, voir mon livre, *La Neuroéthique. Ce que les neurosciences font à nos conceptions morales*, Paris, La Découverte, 2009, ch. 1.

Quelques promesses philosophiques des neurosciences

Les problèmes relevant des implications éthiques des neurosciences ont longtemps constitué le domaine privilégié des philosophes et de ceux qui utilisaient la philosophie comme une servante (les théologiens). Ils ont ceci de particulier que les progrès des neurosciences nous obligent à les poser à nouveaux frais, vu les données dont nous disposons à présent. Certains auteurs sont très optimistes; d'autres sont sur la défensive ou, du moins, pensent que les neurosciences n'apportent rien de fondamentalement nouveau. C'est d'ailleurs bien la question qu'il faut poser, et pour y répondre, j'examinerai successivement quatre problématiques neuroéthiques de nature philosophique: 1° le dualisme psycho-physique, 2° la nature de l'amour, 3° la nature de l'éthique et 4° l'amélioration de l'être humain. Mais auparavant, et pour introduire le sujet, j'aimerais l'illustrer par quelques prises de position, généralement plutôt optimistes.

Dans un style simplement descriptif, Paul Root Wolpe relève que « l'affirmation que les neurosciences peuvent nous dire quelque chose concernant l'existence ou la non existence de l'âme est précisément le genre d'affirmation qui rend les auteurs religieux nerveux. »¹² Méfiance chez certains, enthousiasme chez d'autres. Sur la question du déterminisme, dans un article que j'ai écrit avec Alex Mauron, nous affirmons que « si tout événement qui a une influence sur notre comportement peut seulement l'avoir à travers quelque événement cérébral, si tout événement cérébral est déterminé par des chaînes causales d'événements, et si tout état mental est un état cérébral, alors le neurodéterminisme est identique au déterminisme *tout court* appliqué aux actions humaines. En conséquence, pour la première fois dans l'histoire, nous sommes à même d'observer le déterminisme *tout court* au travail dans notre vie mentale. »¹³ L'idée est la suivante : le déterminisme *tout court*, à savoir le déterminisme physique qui règne dans notre monde, doit être

.

¹² P. Wolpe, « Religious Responses to Neuroscientific Questions », in J. Illes, dir., *Neuroethics*, p. 292.

¹³ B. Baertschi & A. Mauron, « Genetic Determinism, Neuronal Determinism, and Determinism *tout court* », in J. Illes & B. Sahakian, dir., *Oxford Handbook of Neuroethics*, Oxford, OUP, 2011, ch. 9 (sous presse).

distingué des différents déterminismes partiels (génétique, économique, social, psychologique,...). Le neurodéterminisme, c'est-à-dire le déterminisme à l'œuvre dans notre cerveau et qui influence notre comportement et notre pensée, paraît être l'un de ces multiples déterminismes partiels qui modèlent notre conduite, un ensemble de facteurs nécessaires mais non suffisants pour l'expliquer. À y regarder de plus près, on se rend cependant compte qu'il n'est pas correct de voir les choses ainsi; en fait, vu que tout ce que nous subissons et faisons est inscrit dans notre cerveau et est médiatisé par lui, vu que par conséquent tous les déterminismes partiels s'v inscrivent, il faut affirmer que le neurodéterminisme n'est rien d'autre que le déterminisme physique luimême qui s'exerce dans cette partie du monde qu'est notre cerveau et qui façonne nos pensées et nos actions. Ainsi les neurosciences confirment la position déterministe, réfutent la croyance libertarienne dans le libre arbitre et ne laissent subsister que deux options concernant la liberté humaine : le déterminisme dur, qui exclut tout libre arbitre, et le compatibilisme ou déterminisme « soft », qui conserve le libre arbitre, mais en le redéfinissant : une action est libre lorqu'elle n'est pas contrainte et est motivée par des raisons

L'enthousiasme peut concerner les possibilités actuelles des neurosciences, comme on vient de le voir, mais aussi leurs promesses futures. Michael Gazzaniga espère qu'elles seront à même de nous doter de la morale universaliste à laquelle nous aspirons depuis Kant au moins: « Je suis convaincu que nous devons nous engager à nous persuader qu'une éthique universelle est possible et que nous devons chercher à la comprendre et à la définir. »¹⁴ Depuis la nuit des temps, nous nous sommes fiés à des contes et à des légendes pour nous comprendre et nous diriger; les neurosciences nous ont toutefois appris à quel point ces croyances n'étaient pas fiables. Mais nous avons besoin de directives pour vivre ; il nous faut donc les chercher ailleurs et les demander à la science contemporaine, qui devrait être apte à nous les fournir grâce aux progrès des connaissances sur notre cerveau.

¹⁴ M. Gazzaniga, *The Ethical Brain*, New York, Dana Press, 2005, p. 178.

De manière plus grandiose, le transhumanisme nous promet un avenir radieux grâce aux développements, présents mais surtout futurs, des neurosciences. Alors que nous pensions depuis l'aube des temps que la condition ou la nature humaine était immuable, qu'elle marquait les limites de toute action humaine possible, nous nous rendons compte maintenant que c'était une erreur. C'est du moins l'avis de Nick Bostrom, l'une des figures marquantes du mouvement transhumaniste : les sciences actuelles vont bientôt nous permettre d'agir sur les fondements mêmes de l'humain, bref, non seulement d'aménager la nature humaine, mais de la transformer ou plutôt, de la dépasser: « "Le transhumanisme" a acquis ses assises en établissant une facon de penser qui met au défi la prémisse suivante : la nature humaine est et devrait rester essentiellement inaltérable »¹⁵. Bien sûr, par le passé, certains déjà avaient pensé souhaitable et même nécessaire de changer l'être humain, mais personne ne disposait des movens de le faire; maintenant, cela devrait changer, grâce aux neurosciences ainsi qu'à l'informatique, au génie génétique et aux nanotechnologies (les technologies convergentes). On sent ici passer un souffle quasi-démiurgique.

Il est manifestement tentant et aisé de s'emballer quand on évoque les promesses des neurosciences; ce faisant, on tombe facilement dans des pièges, comme ceux qu'Eric Racine et Judy Illes ont appelé neuroréalisme et neuroessentialisme. Le neuroréalisme consiste à croire que les images délivrées par l'imagerie à résonnance magnétique fonctionnelle (IRMf) et d'autres scanners nous mettent en face du phénomène lui-même; par exemple que les neurobiologistes nous permettent non seulement de voir les émotions, mais encore de visualiser leur cause neuronale (alors qu'on voit des voxels colorés représentant des variations d'oxygénation du sang qui sont seulement corrélés à des états cérébraux, liés de manière souvent difficile à déterminer précisément à des états mentaux). Quant au neuroessentialisme, il consiste à identifier la personne et sa subjectivité à son cerveau, comme s'il allait de soi

¹⁵ http://www.transhumanism.org/index.php/WTA/more/147/

qu'elle n'était rien d'autre¹⁶. Ce n'est d'ailleurs pas que les enthousiastes aient tort, ou qu'une forme de neuroessentialisme ne soit pas correcte, mais si on veut l'établir, il y faut des arguments et non des actes de foi.

C'est justement aux arguments que je vais venir maintenant, en examinant les quatre sujets que j'ai annoncés plus haut. Je ne saurais évidemment les épuiser et n'ai d'ailleurs aucune intention de le tenter. Ce que je veux faire, c'est montrer et illustrer comment et en quoi les neurosciences ont fait évoluer les débats, quel est leur apport et leur fécondité.

La personne sans âme ou la réfutation du dualisme

Descartes est le père du dualisme moderne, c'est-à-dire de la thèse qui affirme que l'être humain est composé de deux substances distinctes et séparables, l'âme et le corps. Fondant cette thèse immédiatement sur le *cogito*, qui est « tout le meilleur que nous puissions choisir pour connaître la nature de l'âme, et qu'elle est une substance entièrement distincte du corps »¹⁷, il tente ensuite d'expliquer leur union, qui est pour lui un fait d'expérience, mais dont il veut rendre compte du mécanisme. « La nature m'enseigne aussi par ces sentiments de douleur, de faim, de soif, etc., que je ne suis pas seulement logé dans mon corps, ainsi qu'un pilote en son navire, mais outre cela que je lui suis conjoint très étroitement, et tellement confondu et mêlé, que je compose comme un seul tout avec lui »¹⁸, dit notre philosophe, qui place l'âme dans la glande pinéale (l'épiphyse), afin qu'elle puisse agir sur le cerveau et en recevoir les impressions par l'intermédiaire des esprits animaux circulant dans les nerfs.

Si le mécanisme de l'union psycho-physique reste délicat à déterminer précisément, c'est toutefois la possibilité même d'une union qui a été rapidement mise en doute. Thomas Hobbes a été le premier à exprimer sa perplexité, mais c'est la princesse Élisabeth qui a élevé l'objection la plus

¹⁶ A. Lavazza & M. de Caro, « Not so Fast. On Some Bold Neuroscientific Claims Concerning Human Agency », *Neuroethics*, 2010/1, p. 24.

¹⁷ Les Principes de la philosophie, in Œuvres philosophiques, éd. Alquié, Paris, Garnier, 1973, t. III, p. 95.

¹⁸ Méditations métaphysiques, éd. Alquié, t. II, p. 492.

forte : selon elle, l'interaction de l'âme et du corps est difficile à admettre, car elle viole le principe physique fondamental de la conservation de la quantité de mouvement $(m \times v)$. Lorsqu'un corps est mû, le mouvement qu'il acquiert provient du corps qui le meut, si bien que, en admettant l'inexistence de tout frottement, la quantité totale de mouvement reste constante. Mais l'âme est immobile ; elle ne peut donc transmettre son mouvement au corps ; elle doit donc créer du mouvement qui ne vient de nulle part, augmentant par là sa quantité totale, ce qui viole le principe.

Descartes sera dans l'impossibilité de répondre valablement à l'objection, et ses successeurs abandonneront soit l'interaction soit l'espoir de trouver une solution. John Locke adoptera cette dernière position en soutenant une forme d'agnosticisme : « Par rapport à nos notions il ne nous est pas plus malaisé de concevoir que Dieu peut, s'il lui plaît, ajouter à notre idée de la matérialité la faculté de penser, que de comprendre qu'il y joigne une autre substance avec la faculté de penser. »¹⁹. Monisme matérialiste et dualisme sont renvoyés dos à dos, vu que nous sommes dépourvus des capacités cognitives suffisantes pour les départager. Malebranche et Leibniz soutiendront, eux, l'impossibilité de l'interaction et lui substitueront le parallélisme psycho-physique, sous la forme de l'occasionnalisme et de l'harmonie préétablie : l'âme et le corps sont deux substances, mais elles n'interagissent pas ; elles se correspondent simplement. Leibniz illustre les trois positions dualistes – l'interactionnisme, l'occasionnalisme et l'harmonie préétablie – de la manière suivante:

« Figurez-vous deux horloges ou deux montres, qui s'accordent parfaitement. Or cela se peut faire de *trois façons*. La première consiste dans l'influence mutuelle d'une horloge sur l'autre; la seconde dans le soin d'un homme qui y prend garde; la troisième dans leur propre exactitude. La *première façon*, qui est celle de l'influence, a été expérimentée par feu Monsieur Huygens à son grand étonnement. Il avait deux grandes pendules attachées à une même pièce de bois; les battements continuels de ces pendules avaient communiqué des tremblements semblables aux particules du bois, mais

¹⁹ Essai philosophique sur l'entendement humain, Paris, Vrin Reprise, 1972, p. 443.

ces tremblements divers ne pouvant pas bien subsister dans leur ordre, et sans s'entr'empêcher, à moins que les pendules ne s'accordassent, il arrivait par une espèce de merveille, que lorsqu'on avait même troublé leur battement tout exprès, elles retournaient bientôt à battre ensemble, à peu près comme deux cordes qui sont à l'unisson. La *seconde manière* de faire toujours accorder deux horloges bien que mauvaises, pourra être d'y faire toujours prendre garde par un habile ouvrier, qui les mette d'accord à tout moment : et c'est ce que j'appelle la voie de l'assistance. Enfin la *troisième manière* sera de faire d'abord ces deux pendules avec tant d'art et de justesse, qu'on se puisse assurer de leur accord dans la suite; et c'est la voie du consentement préétabli. »²⁰

Bien des auteurs du XVIII^e siècle reprendront l'une ou l'autre de ces solutions, mais une manière différente de résoudre la difficulté va poindre, qui consiste à rejeter le dualisme au profit du monisme matérialiste. Locke, on l'a vu, avait envisagé cette possibilité, Spinoza en avait défendu une variante, mais ce sont La Mettrie et Diderot qui vont en donner une version spécifiquement moderne, fondée sur une propriété particulière du vivant, la sensibilité ou capacité de réagir à des stimuli, à laquelle ils estiment que la pensée est apparentée. Avant Lamarck, les matérialistes hésiteront pour savoir si la sensibilité, propriété proto-psychique du vivant, réside dans toute matière vivante ou survient à l'occasion de la complexification caractéristique de certains organismes, mais les principes du matérialisme sont clairement énoncés. Diderot, en effet, demande: « Qu'est-ce que la *sensibilité* ? Est-ce une propriété générale de la matière ? Est-ce une propriété résultante d'une certaine combinaison ? »²¹

Cabanis résumera la doctrine en employant une image célèbre, que Darwin reprendra: « Pour se faire une idée juste des opérations dont résulte la pensée, il faut considérer le cerveau comme un organe particulier, destiné

_

²⁰ G. W. Leibniz, Éclaircissements du Système nouveau de l'âme et du corps, in Die philosophischen Schriften von G. W. Leibniz, 7 vol., éd. Gerhardt, Reprint Olms, 1978, t. IV, p. 500.

D. Diderot, Commentaire, in F. Hemsterhuis, Lettre sur l'homme et ses rapports, éd. G. May, New Haven, Yale University Press, 1964, p. 109.

spécialement à la produire ; de même que l'estomac et les intestins à opérer la digestion, le foie à filtrer la bile, les parotides et les glandes maxillaires et sublinguales à préparer les sucs salivaires. Les impressions, en arrivant au cerveau, le font entrer en activité ; comme les aliments, en tombant dans l'estomac, l'excitent à la sécrétion plus abondante du suc gastrique et aux mouvements qui favorisent leur propre dissolution [...]. Nous concluons avec la même certitude que le cerveau digère en quelque sorte les impressions ; qu'il fait organiquement la sécrétion de la pensée. »²²

Cette conception va devenir dominante dans le monde scientifique, sous l'une ou l'autre de ses variantes contemporaines. Elle va être confortée par des arguments conceptuels et empiriques.

Parmi les arguments conceptuels, celui de l'exclusion causale, développé par Jaegwon Kim, est décisif à mes yeux. Soient deux événements cérébraux (C et C*) et deux événements mentaux corrélatifs (M et M*), unis par des relations causales, représentées par des flèches, pleines ou en pointillé (Figure 1). C et C* représentent donc l'activation de deux zones différentes du cerveau, M et M* les états mentaux correspondants. Les états cérébraux causent les états mentaux qui leur correspondent, mais les états mentaux aussi ont une influence causale, tant sur d'autres états mentaux que sur certains états cérébraux. Ainsi, par exemple, la perception d'un danger (M), causée par l'activation du cortex visuel (C), cause l'agitation de l'amygdale (C*), et donc un sentiment de peur (M*). Toutefois, il est facile de montrer que, dans une telle configuration, seules certaines de ces relations causales sont des relations réelles (représentées par les flèches pleines), les autres n'étant qu'apparentes (représentées par les flèches en pointillé). Il s'ensuit que le mental n'a aucune efficacité causale et donc qu'il ne peut exister indépendamment du cérébral : « Les relations causales M-à-M* et M-à-C* sont seulement apparentes, surgissant d'un processus causal authentique de C à C* »²³.

_

²² P. J. G. Cabanis, Rapports du physique et du moral de l'homme, Genève, Slatkine Reprints, 1980, p. 137-138. J'ai longuement étudié la question des rapports de l'âme et du corps pendant l'époque moderne dans mon livre Les Rapports de l'âme et du corps. Descartes, Diderot, Maine de Biran. Paris, Vrin. 1992.

²³ J. Kim, *Mind in a Physical World*, London, MIT Press, 1998, p. 45.

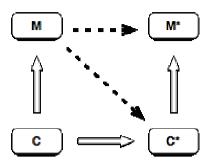


Figure 1

En effet, si l'événement cérébral (C) corrélat d'un état mental (M) cause l'événement cérébral subséquent (C*) correspondant à l'état mental subséquent (M*), la première pensée n'a plus rien à causer: elle est causalement superflue. Cela signifie qu'elle n'a pas de réel pouvoir causal, donc qu'elle ne fait pas partie des éléments irréductibles de la réalité (le mental est réductible au cérébral). Ainsi, les flèches en pointillé de la Figure 1 ne désignent pas de véritables relations causales, mais des relations apparentes: on a l'impression qu'une idée en produit une autre, mais *stricto sensu* c'est faux : il n'y a pas de causalité horizontale du mental au mental, ni de causalité descendante du mental au cérébral, du moins sur le plan ontologique²⁴. Il s'ensuit que le monde est causalement clos, qu'aucune entité non physique ne peut agir dans notre monde physique: c'est ce qu'on appelle la thèse de la clôture causale du monde.

Jusqu'ici, les neurosciences n'ont joué aucun rôle dans l'argumentation en faveur du matérialisme, même si La Mettrie et Cabanis s'intéressaient déjà au fonctionnement du cerveau et à l'effet de sa conformation sur notre comportement. Le matérialisme s'est établi indépendamment et antérieurement au développement contemporain de notre connaissance du cerveau. Par contre, ce développement a joué et continue à jouer un rôle important comme étai du

²⁴ Pour de plus amples développements, voir mon livre *La Neuroéthique*, p. 101-104.

matérialisme en montrant à quel point le mental est dépendant du cérébral. Pensons à Phineas Gage : la lésion cérébrale a profondément modifié sa psychologie, et on observe la même chose dans de très nombreuses situations.

Prenons, comme seul exemple, le syndrome de Capgras. Dans cette affection, le patient, tout en étant parfaitement capable d'identifier les visages des personnes avec lesquelles elle vit, affirme qu'elles ont été remplacées par des sosies²⁵ – c'est là une condition qui affecte la manière dont nous nous référons fondamentalement à autrui, pour ainsi dire à son « âme ». On observe un dysfonctionnement cérébral sous-jacent, mais comment expliquer l'illusion au niveau psychologique? Deux hypothèses ont été avancées. Pour la première, l'illusion se situe déjà au niveau de la perception, qui est incorrecte à cause du dysfonctionnement cérébral ; le jugement est donc faussé puisqu'il repose sur une base perceptuelle erronée. Pour la seconde, la perception est anormale en ce qu'elle n'est pas accompagnée des émotions habituelles, et l'erreur de jugement vient de l'esprit lui-même qui essaye de rationaliser sa perception, de lui donner un sens intellectuellement acceptable. C'est ainsi que Ramachandran l'explique: « Selon lui, il v a d'abord une perception anormale des parents du patients ou d'autres personnes aimées, puisqu'elle a lieu sans que soit simultanément expérimentée aucune réponse émotionnelle. Cela cause l'entrée en jeu d'un second facteur, la tentative de rationaliser cette expérience en regardant la personne aimée comme un imposteur. »²⁶

Comme ce cas l'illustre bien, le mental (l'âme) est chevillé au cérébral (le corps) : perception, émotion et jugement sont des événements mentaux qui sont inséparables de leurs corrélats cérébraux. Modifier ceux-ci, c'est modifier ceux-là ; ils n'ont aucune indépendance les uns par rapport aux autres. Les neurosciences actuelles nous offrent une myriade de situations de même nature. Certes, le XVIII^e siècle avait déjà fait ce genre d'observations et risqué

_

²⁵ Cf. W. Hirstein, « Confabulations About People and their Limbs, Present or Absent », in J. Bickle, dir., *The Oxford Handbook of Philosophy and Neuroscience*, Oxford, OUP, 2009, p. 482.

²⁶ J. Mundale & S. Gallagher, « Delusional Experience », in J. Bickle, dir., *The Oxford Handbook of Philosophy and Neuroscience*, p. 514.

quelques hypothèses, ainsi La Mettrie, lorsqu'il disait: « Que fallait-il à Canius Julius, à Sénèque, à Pétrone, pour changer leur intrépidité en pusillanimité, ou en poltronnerie? Une obstruction dans la rate, dans le foie, un embarras dans la veine porte. Pourquoi? Parce que l'imagination se bouche avec les viscères; et de là naissent tous ces singuliers phénomènes de l'affection hystérique et hypocondriaque. »²⁷ Or, le cerveau aussi est un viscère! Toutefois, on manquait alors de moyens pour explorer systématiquement l'effet de ces obstructions – ou lésions.

Le jugement que le même médecin portait sur l'être humain reste d'ailleurs à mon sens d'actualité: « Ces êtres fiers et vains, plus distingués par leur orgueil, que par le nom d'hommes, quelque envie qu'ils aient de s'élever, ne sont au fond que des animaux, et des machines perpendiculairement rampantes. »²⁸

L'amour neuronal

Le mental n'a pas d'efficace causale par lui-même, l'être humain n'est qu'un animal bipède; tout cela sonne comme du réductionnisme! À cette accusation – car cela en est une – je peux faire deux réponses, l'une courte, l'autre un peu plus longue. La réponse courte se résume en deux mots: et alors? En effet, si ce que j'ai dit est vrai et que c'est du réductionnisme, alors le réductionnisme est vrai et toute personne éclairée serait bien inspirée de faire profession de réductionnisme. Toutefois, et c'est la réponse plus longue, si tout ce que j'ai dit est compatible avec un réductionnisme, il n'y a aucune implication stricte. Pour le montrer, il faut d'abord dire en quoi le réductionnisme en jeu ici consiste, car c'est un concept notablement flou. Si, dans son sens strict et restreint, il concerne une relation entre des théories, dans le problème qui nous occupe il s'applique plutôt à des entités. Bref, c'est un réductionnisme ontologique, et plus précisément un réductionnisme du mental au cérébral (de l'âme au corps), qui pour certains mettrait en péril la dignité

²⁷ J. O. de la Mettrie, *L'homme-machine*, in *Œuvres philosophiques*, Paris, Fayard, 1987, t. I, p. 10.

²⁸ J. O. de la Mettrie, *L'homme-machine*, p. 110.

humaine. C'était déjà le souci de Bouvard et de Pécuchet, à qui Gustave Flaubert fait dire : «Ils éprouvaient une sorte d'humiliation à l'idée que leur individu contenait du phosphore comme les allumettes, de l'albumine comme les blancs d'œufs, du gaz hydrogène comme les réverbères»²⁹.

Le souci n'est toutefois pas fondé. D'abord parce qu'il n'existe aucune relation d'implication entre le matérialisme et la négation de la dignité humaine, et ensuite parce que l'argument de l'exclusion causale, s'il nie l'efficace causale du mental et implique donc un réductionnisme ontologique du mental au cérébral - réductionnisme qui n'est que la conséquence de la thèse moniste matérialiste –, reste compatible avec plusieurs conceptions qui ne sont pas toutes réductionnistes au sens phénoménologique ou épistémologique, c'est-à-dire qui n'affirment pas toutes que le mental n'est qu'une apparence (voire une illusion) ou que la neurophysiologie pourrait remplacer avantageusement la psychologie. Si le mental n'a pas d'efficace, cela peut certes déboucher sur son élimination (Patricia Churchland³⁰) ou sur l'épiphénoménalisme (Dieter Birnbacher³¹), mais cela peut aussi mener aux théories de l'identité (le mental et le cérébral sont deux aspects du même événement, ou deux propriétés de la même chose : David Armstrong³²), au fonctionnalisme (Daniel Dennett³³ et, par un anachronisme qui se justifie, Aristote³⁴) ou à l'émergentisme (le mental survient par rapport au cérébral), qui ne sont pas des théories qu'on peut taxer de réductionnistes autrement qu'au sens ontologique. L'argument de Kim se limite par conséquent à être un argument en faveur du matérialisme, sous bien de ses formes.

Expliquer le haut (le mental) par le bas (le cérébral) n'est donc pas ontologiquement problématique ; pour un partisan de la théorie de l'identité,

²⁹ G. Flaubert, *Bouvard et Pécuchet*, Paris, Gallimard-Folio, 1999, p. 117.

³⁰ Cf. P. Churchland, *Neurophilosophy*, Cambridge Mass., MIT Press, 1989, ch. 7-8.

³¹ Cf. D. Birnbacher, « The Mind-Body Issue », in J. Giordano & B. Gordjin, *Scientific and Philosophical Perspectives in Neuroethics*, Cambridge, CUP, 2010.

³² Cf. D. Armstrong, A Materialist Theory of the Mind, London, RKP, 1968.

³³ Cf. D. Dennett, *La Conscience expliquée*, Paris, Odile Jacob, 1993.

³⁴ Cf. M. Nussbaum & A. Rorty, dir., *Essays on Aristotle's De Anima*, Oxford, OUP, 1992.

c'est même deux manières de rendre compte d'un même phénomène qui s'exprime différemment au niveau de la neurophysiologie et à celui de la psychologie. Je vais l'illustrer par l'exemple de l'amour.

Dans *The Economist* du 12 février 2004, on lisait ceci: « Des scientifiques ont découvert que, en définitive, l'amour n'est qu'une addiction entre deux personnes. » L'explication de ce titre provocateur était fournie par un article sur deux populations de campagnols dont les mœurs amoureuses sont fort différentes, les campagnols des prairies et les campagnols des montagnes. Les premiers sont monogames, ce qui est rare chez les mammifères (seulement 3% des espèces le sont). Chez eux, la séduction dure 24 heures et est suivie par un lien qui n'est pas rompu pendant la durée de toute la vie. Le mâle et la femelle passent leurs journées ensemble, élèvent amoureusement leurs enfants, et le premier garde agressivement la seconde, éloignant tout rival potentiel. Ce n'est pas du tout le cas du campagnol des montagnes : chez lui l'amour, c'est-à-dire l'accouplement, ne s'étend pas sur une durée de plus de 24 heures, et il change régulièrement de partenaire.

Le comportement de ces deux populations de campagnols est très différent, pourtant il ne dépend que d'un petit nombre de gènes. Un d'entre eux semble jouer un rôle décisif (AVPR1A), en déterminant la place que les récepteurs de la vasopressine vont occuper dans le cerveau. Or, chez les campagnols des plaines, ces récepteurs se trouvent près des zones de l'addiction – et aussi de la récompense³⁵ –, d'où le titre de l'article de l'*Economist*

Ainsi en va-t-il chez les campagnols; mais à moins d'être un réductionniste borné – d'une autre sorte que celui qui réduit le mental au cérébral : il s'agirait ici de réduire l'être humain à l'animal –, personne ne voudrait en tirer des conclusions relativement à l'amour entre êtres humains, dont on connaît la richesse émotionnelle et spirituelle sans commune mesure avec ce que peuvent éprouver ces rats-des-champs! Et pourtant, on ne peut nier qu'il existe des points communs au niveau cérébral. Ainsi que le souligne Larry Young : « De la même manière, chez l'être humain, différentes formes

³⁵ L. Young, « Interview », AJOB Neuroscience, 2010/3, p. 51.

du gène AVPR1A sont associées avec des variations dans la formation des couples et dans la qualité de la relation. Une étude récente a montré que les hommes possédant une certaine variante de AVPR1A ont deux fois plus de chances de rester célibataires que les hommes qui n'ont pas cette variante ; ou, s'ils se marient, ils ont deux fois plus de chances de vivre une crise conjugale. »³⁶

Cette fois, penseront certains, on est bien de retour au réductionnisme, même s'il n'est pas borné : le mental (ici : l'émotion amoureuse) est réduite au cérébral (ici : la place des récepteurs de la vasopressine). Mais cette fois encore, c'est une erreur de le penser, ou du moins il faut distinguer et préciser. D'abord, tout ce que nous vivons psychologiquement est inscrit dans notre cerveau; il est donc normal que tout ce que nous vivons amoureusement le soit - le contraire serait même tout à fait inconcevable et se heurterait à l'impossibilité de l'action d'une entité immatérielle sur une entité matérielle. bref à l'impossibilité du dualisme cartésien. Ensuite, si ce que l'on observe au niveau neurophysiologique et ce qu'on observe au niveau de la psychologie. est ontologiquement identique, cela reste bien différent épistémologiquement : la neurophysiologie et la psychologie ont chacune et à bon droit leur vocabulaire et leurs conceptions propres - il suffit de lire un livre de neurophysiologie et un livre de psychologie pour s'en convaincre. Enfin, Young et ceux qui pensent comme lui n'affirment pas que l'amour est identique avec quelque chose comme la quantité de vasopressine présente dans notre cerveau : ils affirment seulement l'existence d'une relation de causalité entre la vasopressine et l'amour (ou son corrélat cérébral), et une relation de causalité qui n'est pas exclusive. Il existe une grande quantité d'hormones qui influencent la relation amoureuse et une grande quantité de facteurs non cérébraux (sociaux, culturels, éducationnels,...) qui, s'ils s'expriment évidemment dans le cerveau, puisque tout s'y exprime, ne sont pas tous liés à une cause génétique, mais souvent environnementale.

³⁶ L. Young, « Love: Neuroscience Reveals All », *Nature*, vol. 457, 8 janvier 2009, p. 148.

Bref, étudier la base neuronale de l'amour, comme celle de tous les sentiments, n'implique en aucune manière qu'on réduise la signification humaine de l'amour à des échanges hormonaux. Les Roméo et les Juliette n'ont rien à craindre des neurosciences – mais peut-être peuvent-ils en espérer quelque chose et qu'on pourra un jour, en agissant sur leurs émotions via la vasopressine et une autre hormone, l'ocytocine³⁷, empêcher qu'ils finissent tragiquement.

Les confabulations éthiques

Gazzaniga espère que les neurosciences nous aideront à construire une éthique universaliste, dont nous avons désespérément besoin. La neuroscience de l'éthique étudie notre comportement moral. Serait-il possible que cette étude nous mette sur la voie ? C'est ce que pense Joshua Greene³⁸. Selon lui, nous possédons d'ailleurs déjà cette éthique universaliste, c'est l'utilitarisme. Mais elle a de nombreux concurrents, et notamment le déontologisme. Ce que Greene attend de la neuroscience de l'éthique, c'est qu'elle nous permette de montrer que ces concurrents sont inappropriés.

Avant tout, je dois préciser ce qu'on entend par « utilitarisme » et par « déontologisme ». Ce sont les deux théories morales qui, actuellement, sont les plus répandues ; ce ne sont pas les seules (pensons à la théorie des droits, à l'éthique des vertus ou à l'eudémonisme), et donc le propos de Greene ne se veut pas exhaustif, mais seulement une contribution de poids au débat. L'utilitarisme est la théorie morale qui affirme qu'une action est moralement bonne si et si seulement elle maximise le bonheur de tous les individus concernés. Cette maximisation se mesure dans les conséquences de l'action, c'est pourquoi on appelle souvent l'utilitarisme un conséquentialisme. En fait, il existe de nombreuses théories conséquentialistes, selon la valeur qu'il s'agit

-

³⁷ Hormone qui joue aussi un rôle important dans la création du lien affectif entre la mère et son enfant; cf. L. Young, « Interview », *AJOB Neuroscience*, p. 52.

³⁸ Il n'est pas le seul. Neil Levy affirme dans un article récent où il critique le principe des actions à double effet: « We should use the neuroscience of ethics to illuminate the ethics of neuroscience. » (« Neuroethics: A New Way of Doing Ethics », *AJOB Neuroscience*, 2011/2, p. 8).

de maximiser ou de promouvoir; pour l'utilitarisme classique, c'est le bonheur entendu comme plaisir et absence de douleur, pour l'utilitarisme contemporain, c'est la satisfaction des préférences, et pour d'autres théories c'est la vie humaine ou la perfection. Le déontologisme, lui, soutient qu'une action est moralement bonne si elle est conforme à une norme, indépendamment de ses conséquences. La version du déontologisme la plus connue est celle de Kant, pour qui l'action humaine doit être conforme à l'impératif catégorique ; le *Décalogue* énonce aussi une morale déontologiste. Ce qui caractérise cette approche, c'est qu'elle promulgue en général des obligations et des interdictions qui se présentent sous une forme catégorique ou absolue, alors que pour l'utilitarisme les obligations et les interdictions sont toujours conditionnelles – Kant dirait qu'elles sont hypothétiques, c'est-à-dire soumises à un « si » : si telle action maximise le bonheur, alors fais-là ; sinon, ne la fais pas !

Greene pense que l'imagerie cérébrale nous permet de trancher entre ces deux approches morales, au profit de l'utilitarisme. Quelle est son argumentation ? Il la présente en s'appuyant sur ce qu'il considère comme une expérience de pensée cruciale, le *dilemme du wagon fou* (en anglais: *trolley problem*). Ce dilemme a été imaginé en 1967 par Philippa Foot dans un article célèbre à propos du problème de l'avortement³⁹. Voici la situation que le dilemme met en jeu – je m'appuie sur l'exposé qu'en proposent Marc Hauser et ses collègues⁴⁰. Un wagon dévale une pente à vive allure car ses freins sont hors-d'usage. Plus bas sur la voie travaillent cinq ouvriers qui vont être écrasés, d'autant qu'il n'y a aucun moyen de les prévenir. Toutefois, un aiguillage permettrait de faire aller le wagon sur une autre voie où un seul ouvrier travaille. Quelqu'un, qu'on appellera Denise, a la possibilité d'actionner cet aiguillage. A-t-elle le droit de le faire ? 85% des personnes

³⁹ P. Foot, « The Problem of Abortion and the Doctrine of the Double Effect » in *Virtues and Vices and Other Essays in Moral Philosophy*, Oxford, Blackwell, 1978.

⁴⁰ M. Hauser, F. Cushman, L. Young, R. Kang-Xing & J. Mikhail, « A Dissociation Between Moral Judgments and Justification », *Mind and Language*, 2007, vol. 22.

interrogées affirment qu'elle en a le droit. Cela ne surprend pas vraiment : ne vaut-il vaut pas mieux qu'une seule personne meure plutôt que cinq?⁴¹

On change maintenant le scénario (*le gros homme* ou *la passerelle*). Le wagon dévale la pente, mais il n'y a aucun aiguillage. Toutefois, un individu nommé Frank se trouve sur une passerelle qui enjambe la voie, à côté d'un homme suffisamment gros pour que, si Frank le pousse et qu'il tombe sur la voie, son corps arrête le wagon et l'empêche de poursuivre sa route meurtrière. A-t-il le droit de le faire? Cette fois, 88% des mêmes personnes interrogées nient qu'il en ait le droit, alors même qu'ainsi une seule personne mourrait plutôt que cinq, tout comme dans le cas précédent.

Cette histoire est déroutante, car les mêmes personnes répondent différemment alors que la situation paraît moralement analogue et que le résultat est identique : un mort au lieu de cinq. Comment expliquer cette différence? Pour le savoir, Greene a utilisé l'imagerie cérébrale (IRMf) afin de voir ce qui se passait dans leur cerveau lorsqu'on leur présentait le dilemme. Il a constaté que ce ne sont pas les mêmes zones qui sont activées dans les deux cas. Plus précisément, il a observé que plus la question morale qui se pose est personnelle, c'est-à-dire plus l'agent est impliqué activement dans ce qu'il faut faire, et plus il peut le faire directement et simplement en étendant le bras, plus les zones émotionnelles cérébrales de l'observateur ou de l'agent sont actives⁴²; or dans le gros homme, l'agent est bien plus impliqué personnellement et plus proche de l'action. En généralisant, on dira que nos décisions morales et nos jugements moraux sont influencés par la place que nous occupons dans l'action, et plus nous sommes concernés, plus nos émotions sont vives, inhibant l'action lorsqu'elle implique un dommage sérieux fait à autrui

⁴¹ Ces résultats ont été obtenus par sondage sur internet, ce qui a permis de toucher énormément de personnes, de toute origine, même s'il y a un léger biais, vu que seules les personnes ayant accès à internet ont pu voter.

⁴² Hauser présente le dilemme du point de vue de l'observateur; Greene le fait du point de vue de l'agent. Les résultats sont un peu différents selon le mode de présentation, mais pas de manière à infirmer l'argument présenté, qui reste valable dans les deux cas. Voir T. Nadelhoffer & A. Feltz, « The Actor-Observer Bias and Moral Intuitions: Adding Fuel to Sinnott-Armstrong's Fire », *Neuroethics*, 2008/2, p. 139-140.

C'est là une constatation psychologique : on observe ce qui se passe dans notre cerveau lorsque nous prenons des décisions morales ou portons des jugements moraux. Cela ne nous dit pas quelle est la bonne réaction du point de vue normatif. Toutefois. Greene pense qu'il est possible de sauter le pas, car ce qu'il a observé manifeste une irrationalité que nous devons tenter de surmonter, pour aller vers plus de rationalité et donc vers plus de moralité. Il élève ainsi la portée de l'étude du dilemme au statut d'expérience cruciale permettant de trancher entre la pertinence morale du déontologisme et celle de l'utilitarisme: « Nous faisons l'hypothèse que la pensée de causer la mort de quelqu'un en le poussant personnellement (comme dans le dilemme du gros *homme*) est plus forte émotionnellement que la pensée de causer les mêmes conséquences de manière plus impersonnelle (par exemple en manipulant un aiguillage, comme dans le dilemme du wagon fou) [...]. C'est-à-dire que les personnes tendent vers le conséquentialisme dans les cas où la réponse émotionnelle est faible, et tendent vers le déontologisme dans les cas où la réponse émotionnelle est forte »⁴³. Nous sommes tous tantôt déontologistes, tantôt conséquentialistes, selon que nous nous trouvons dans la situation d'une Denise ou d'un Frank. Autrement dit, tantôt nous plaçons notre comportement sous la férule de commandement et d'interdits absolus, lorsque notre réponse émotionnelle est forte, tantôt sous l'égide de l'évaluation des conséquences de nos actes, lorsque notre réponse émotionnelle est faible – ce disant et de manière tout à fait paradoxale, Greene range le déontologisme, et donc Kant, du côté de l'irrationnalisme

Mais il n'y a là rien d'immuable. En effet, il nous est loisible de devenir plus utilitaristes ou, plus généralement, conséquentialistes, c'est-à-dire de faire preuve de plus de rationalité; c'est-là une attitude que nous pouvons décider d'adopter, sinon en faisant taire nos émotions, du moins en ne les écoutant pas, bref en ne leur conférant pas de poids moral. C'est que les réactions émotionnelles sont évolutionnairement archaïques: « La tension entre les perspectives utilitaristes et déontologistes en philosophie morale reflète une

.

⁴³ J. Greene, «The Secret Joke of Kant's Soul», in W. Sinnott-Armstrong, *Moral Psychology*, Cambridge Mass., MIT Press, vol. 3, 2008, p. 43.

tension plus fondamentale qui vient de la structure du cerveau humain. Les réponses socio-émotionnelles que nous avons héritées de nos ancêtres primates (dues sans doute à quelque avantage d'adaptation qu'elles conféraient), structurées et affinées par la culture, sous-tendent les interdictions absolues qui sont centrales au déontologisme. Par contraste, le "calcul moral" qui définit l'utilitarisme est rendu possible par des structures apparues plus récemment dans les lobes frontaux, soutenant la pensée abstraite et le contrôle cognitif de haut niveau »⁴⁴

Il existe donc deux types de théories morales, le conséquentialisme et le déontologisme, où dominent respectivement la raison et les émotions, dont la pertinence morale n'est pas du tout la même. Seule celle qui s'appuie sur la raison est authentiquement morale. Pourtant, les déontologistes eux aussi invoquent la raison – j'ai mentionné à quel point il était paradoxal de faire de Kant un champion, même inconscient, des émotions, et j'ai cité plus haut un passage où il affirme même que les émotions sont moralement non fiables. En outre, les kantiens ont développé une théorie morale sophistiquée où de multiples *raisons* sont alléguées pour justifier ou condamner nos actes. Par exemple, un kantien dira que Frank n'a pas le droit de pousser le gros homme, parce que cela viole l'interdiction de l'instrumentalisation de l'être humain, instrumentalisation qui n'a pas lieu lors de l'action de Denise (la mort de l'ouvrier n'est pas un moyen pour sauver les cinq autres).

Que répond Greene à cette objection ? Selon lui, le fait que les sondés, représentant la morale commune, jugent d'une certaine manière et qu'on peut trouver des raisons morales à leurs jugements n'implique pas que leurs raisons soient d'authentiques raisons morales. Il les taxe en effet de « rationalisations » et même de « confabulations ». On pourrait penser qu'il s'agit là d'une manœuvre rhétorique – déconsidérer l'adversaire –, mais ce n'est pas le cas. Comme on l'a vu, Greene développe une approche de la décision morale dans laquelle les interdits absolus – et celui de l'instrumentalisation en est un – ressortissent à une structure cérébrale

_

⁴⁴ J. Greene, L. Nystrom, A. Engell & J. Cohen, «The Neural Bases of Cognitive Conflict and Control in Moral Judgment », *Neuron*, 2004, n° 44, p. 389.

archaïque, émotionnellement très chargée, par opposition à la pesée rationnelle des intérêts dépendant de structures plus évoluées et donc plus récentes. Richard Dean développe ce point de la manière suivante: « Les êtres humains ont développé de fortes aversions émotionnelles contre l'acte d'infliger des dommages personnels à autrui, mais aucune réaction émotionnelle similaire dans les situations morales "impersonnelles". La force de ces aversions vient du fait qu'elles "aident les individus à disséminer leurs gènes dans un contexte social", et non parce qu'elles "reflètent des vérités morales profondes que l'on pourrait découvrir rationnellement". Ainsi, cet "aspect contingent et non moral de notre histoire" est "moralement non pertinent" [...]. Si la théorie déontologiste est fondée sur ces réactions, alors elle est fondée sur des facteurs moralement non pertinents et est par conséquent non fiable. »⁴⁵

On le voit, l'argumentation de Greene en faveur de l'utilitarisme ou du conséquentialisme dépasse largement le cadre des neurosciences. L'imagerie cérébrale est utilisée pour mieux comprendre les réactions des sujets de l'expérience, mais les conséquences qui en sont tirées reposent encore et surtout sur une certaine conception de la rationalité, du rôle des émotions et sur des considérations évolutionnaires. On n'est donc pas étonné qu'elles aient été largement critiquées, à l'instar de toute thèse philosophique⁴⁶. Notamment, la manière dont Greene argumente paraît s'appuyer sur une distinction très classique du rôle de la raison et des émotions en morale, qui a justement été rejetée à la suite des travaux de Damasio. On voit ainsi que les neurosciences peuvent alimenter des positions différentes, voire antagonistes. Et ce n'est pas surprenant, puisque ce qu'elles fournissent, ce sont des données qu'il s'agit

-

⁴⁵ R. Dean, « Does Neuroscience Undermine Deontological Theory? », *Neuroethics*, 2010/3, p. 46.

⁴⁶ Voir mon article « Neurosciences et éthique: que nous apprend le dilemme du wagon fou? », *IGITUR*, 2011 (accessible sur le site http://www.igitur.org/). Très récemment, les observations de Greene ont aussi été contestées: ce qu'il aurait testé ne serait pas l'opposition entre les réactions utilitaristes et les réactions déontologistes, mais l'opposition entre les réactions à des dilemmes dont la solution est soit intuitive soit contre-intuitive; cf. G. Kahane, K. Wiech, N. Schakel & al., « The Neural Basis of Intuitive and Counterintuitive Moral Judgment », *Social Cognitive and Affective Neuroscience*, advance access published March 18, 2011.

d'interpréter et que toute interprétation se fait à partir de certaines conceptions qui, elles, sont d'ordre théorique et qui, souvent, n'admettent pas d'expérience cruciale, comme c'est le cas des théories philosophiques et éthiques qui nous occupent. Ainsi, si les neurosciences sont un auxiliaire précieux pour la philosophie, ce n'est toutefois qu'un auxiliaire qui ne saurait se substituer à la réflexion et à l'argumentation conceptuelles⁴⁷.

L'humain amélioré

L'être humain n'est pas satisfait de lui-même et de ses capacités ; de tout temps, il a cherché à les améliorer, même s'il y a quelque paradoxe en cela puisque, comme le relève le personnage d'un roman de Voltaire: « Malgré l'opiniâtreté des hommes à louer l'antique aux dépens du moderne, il faut avouer qu'en tout genre les premiers essais sont toujours grossiers »⁴⁸. C'est que, renchérit Alphonse Allais: « L'homme est imparfait, mais ce n'est pas étonnant quand on pense à l'époque où il a été créé »⁴⁹. Éducation, apprentissage, médecine, mise au point d'outils et de machines, invention de l'écriture ont été les vecteurs principaux de l'amélioration de l'être humain. Depuis peu, il est question de neuroamélioration, c'est-à-dire d'améliorations que les neurosciences nous apportent déjà et nous promettent pour le futur. Ces améliorations ciblent essentiellement nos états mentaux, tant cognitifs qu'affectifs. Cette cible non plus n'est pas nouvelle. Pensons à la mémoire. On l'améliore depuis longtemps par l'apprentissage (réciter des poésies, jouer aux échecs,...) et par des outils (agendas, calculatrices, ordinateurs,...). L'attention est aussi aidée par l'ingestion de substances excitantes naturelles, comme le café, et l'humeur s'améliore chez certaines personnes grâce à l'alcool. Qu'est-

_

⁴⁷ Certains auteurs contestent même l'utilité d'expériences de pensée comme celles du wagon fou pour la réflexion éthique; cf. B. Smith, « Can We Test the Experience Machine », *Ethical Perspectives*, 2011/1, p. 44-45. Je n'ai pas la place ici pour en discuter.

⁴⁸ Voltaire, *Le Monde comme il va*, in *Romans de Voltaire*, Livre de poche, 1961, p. 92.

⁴⁹ http://www.proverbes-citations.com/conf6.htm.

ce que les neurosciences nous amènent de plus et que penser des interventions amélioratrices sur le cerveau humain?

Avant tout, il est nécessaire d'introduire une précision concernant l'objet de l'amélioration. Celle-ci peut porter sur la *personne* elle-même, sur ses *capacités*, sur ses *états* ou sur ses *réalisations*. Si un individu prend des médicaments pour mieux réussir à un examen, c'est qu'il veut améliorer une réalisation (une performance); s'il les prend régulièrement pour améliorer sa concentration, ce sera une de ses capacités (une compétence) qui sera modifiée. Il peut aussi vouloir améliorer son humeur (un état), ou lui-même (sa propre personne). Comme une personne est un ensemble de capacités, d'états et d'actes, c'est nécessairement par l'intermédiaire de l'amélioration de ces derniers qu'elle s'améliorera elle-même. L'important est de relever que toute amélioration n'a pas pour objet un changement de la personne elle-même, ni même d'une de ses capacités; elle peut viser une performance transitoire seulement, sans grand impact sur ce qu'elle est, comme une mémorisation ou une attention accrue de manière ponctuelle⁵⁰.

Parmi les améliorants que proposent les neurosciences, il y a des neuromédicaments. Je vais y limiter mon propos. Diverses substances psychoactives ont été mises au point assez récemment; il s'agit de médicaments, dont seul l'usage élargi, hors-indication, est amélioratif. Les principales sont la fluoxétine (tel le Prozac®), un antidépresseur qui peut améliorer la confiance en soi et le bien-être, le méthylphénidate (tel la Ritaline®), prescrite à des millions d'enfants pour traiter l'hyperactivité, et qui a un effet amélioratif sur l'attention, le modafinil (tel le Provigil®), mis sur le marché à l'origine pour soigner la narcolepsie, mais qui permet de rester éveillé et alerte de longues heures, le propanolol, contre le trac et susceptible d'effacer les souvenirs traumatisants, et le donépézil (ou donézépil), qui améliore la mémoire.

L'usage élargi de ces substances est déjà répandu, ainsi que diverses études l'ont montré. Par exemple, aux États-Unis, entre 6,9 et 16,2% des

_

⁵⁰ Cf. B. Baertschi & A. Mauron, « Qu'est-ce qu'une véritable amélioration? », *Bioethica Forum*, n° 4, 2011/1.

étudiants disent avoir utilisé des psychostimulants; les chiffres sont un peu plus bas en Europe. Une enquête menée en 2008 par la revue Nature a montré que 20% des scientifiques avaient eu recours à l'une ou l'autre de ces substances⁵¹. Cependant. l'efficacité de ces dernières en tant que neuroaméliorants n'est pas très spectaculaire, sauf chez certains individus. enthousiasmés par leurs performances. On a certes mesuré un effet sur la mémoire, l'attention et les fonctions exécutives (comme la capacité de planification), mais c'est parfois au détriment d'autres fonctions mentales. Par ailleurs, en ce qui concerne la mémoire, il semble que si les individus qui ont une faible mémoire en bénéficient, ceux qui en ont déjà une bonne voient une détérioration de leurs performances mnésiques⁵². Certains auteurs sont encore plus pessimistes et comparent l'effet des antidépresseurs sur des individus en bonne santé à celui du millepertuis et celui du modafinil à celui du café⁵³. Toutefois, ces résultats ne concernent que les quelques molécules actuellement à notre disposition; il est fort probable que, dans un avenir assez proche et grâce aux progrès des neurosciences, nous disposions d'autres substances, bien plus efficaces.

Comment cet usage élargi est-il perçu dans la société? De manière fort différente selon les milieux, nous apprend une enquête menée par Éric Racine et Cynthia Forlini⁵⁴. Dans les études en santé publique, on parle d'abus de médicament, un jugement très négatif; dans les revues de bioéthique, il est question d'amélioration des capacités, ce qui est parfois purement descriptif, parfois positif, l'amélioration désignant un bénéfice et faisant référence à un idéal de perfection. Dans les médias enfin, on souligne qu'il s'agit du choix d'un style de vie, mettant l'accent sur la liberté de vivre selon l'idéal de la

-

⁵¹ M. Schermer & I. Bolt, « The Future of Psychopharmacological Enhancements: Expectations and Policies », *Neuroethics*, 2009/2, p. 79.

⁵² M. Schermer & I. Bolt, « The Future of Psychopharmacological Enhancements: Expectations and Policies », p. 77-78.

⁵³ M. Synofzik, « Ethically Justified, Clinically Applicable Criteria for Physician Decision-Making in Psychopharmacological Enhancement », *Neuroethics*, 2009/2, p. 97.

⁵⁴ É. Racine & C. Forlini, « Cognitive Enhancement, Lifestyle Choice or Misuse of Prescription Drugs? », *Neuroethics*, 2010/1.

personne qu'on entretient, bref, quelque chose de positif encore. N'y a-t-il donc que des médecins pour s'inquiéter ?

Non, bien des auteurs éprouvent des soucis moraux. Anjan Chatterjee en relève quatre : 1° un souci concernant la sûreté (safety) : 2° un souci concernant le caractère : 3° un souci concernant la justice distributive : 4° un souci concernant la coercition⁵⁵. La question de la sûreté concerne toute nouvelle substance mise sur le marché et ne pose donc pas de problème spécifique pour notre sujet, si ce n'est que les substances psychoactives disponibles ont été testées comme médicaments seulement et non comme améliorants. Le souci à propos de la justice distributive lui non plus n'est pas spécifique, puisqu'il concerne aussi l'accès aux médicaments proprement dits et à tous les produits des biotechnologies nouvelles, le danger étant que seuls les riches puissent se les offrir. Le souci concernant le caractère consiste en ceci : en agissant pharmacologiquement sur nos états mentaux cognitifs et surtout affectifs et émotionnels, nous risquons d'éroder notre force morale. Comme le relève le Comité du Président étasunien, « des personnes en bonne santé dont le comportement perturbé est "remédié" par des calmants plutôt que par leurs propres efforts ne sont pas en train d'apprendre le contrôle de soi »⁵⁶. Quant au souci lié à la coercition sociale, il consiste en ce que, si la prise de neuromédicaments pour améliorer nos performances devenait banal et même une exigence de la vie que nous menons, alors ceux qui décideraient de ne pas en prendre seraient désavantagés.

Dans cette situation, comme toujours, il faut faire un calcul coûts / bénéfices, car ainsi que le disait William Safire: « Supposez que nous puissions développer un médicament qui nous rende moins timide, plus honnête ou intellectuellement plus séduisant, avec un bon sens de l'humour. Qu'est-ce qui nous empêcherait d'utiliser un tel "Botox pour le cerveau" ? »⁵⁷ C'est que, lorsqu'on est moins timide, plus honnête, intellectuellement plus

⁵⁵ A. Chatterjee, « Cosmetic Neurology and Cosmetic Surgery: Parallels, Predictions, and Challenges », *Cambridge Quarterly of Healthcare Ethics*, n° 16, 2007, p. 131.

⁵⁶ The President's Council on Bioethics, *Beyond Therapy*, Washington DC, 2003, p. 291.

⁵⁷ In Neuroethics. Mapping the Field, p. 8.

séduisant, le tout avec un bon sens de l'humour, on est plus heureux : notre bien-être, et celui de notre entourage s'en ressentent de manière significative. Dans le même ordre d'idée, Carl Elliott demande : «Si vous êtes anxieux et vous sentez seul, qu'un médicament peut régler le problème, pourquoi rester anxieux et seul ?»⁵⁸ Pour certaines personnes, je l'ai rapidement mentionné, les substances psychoactives sont réellement des améliorants impressionnants. Ainsi Johann Hari, un journaliste américain, rapporte:

« Je m'assis et pris une tablette de 200 mg de modafinil avec un verre d'eau. Je pris un livre sur la mécanique quantique et la théorie des cordes que j'avais l'intention de lire depuis longtemps. Cinq heures plus tard, je réalisai que j'avais atteint la dernière page. Je n'avais rien noté de particulier, excepté les mots que je lisais, et ils coulaient sans difficulté; je ne m'étais jamais arrêté et n'avais butté sur aucune expression. Perplexe, je me levai, me fis un sandwich – et je fus submergé par un désir violent d'écrire un article qui tournait dans ma tête depuis des mois. Cela fut fait en quelques heures, et cet article fut meilleur que ceux que je rédigeais habituellement. J'étais capable de planer dans un état de concentration profonde, calme et sans effort. C'était comme si j'avais ouvert une fenêtre dans mon cerveau et que tout l'air pesant en était sorti, pour être remplacé par une brise calme. Une fois que j'eus fini cet article, je voulus en faire plus. J'écrivis un autre article, et tout sortait de mon esprit sans effort »⁵⁹. On comprend que John Harris puisse dire: « L'amélioration (enhancement) de l'être humain est par définition bonne, tout comme un bienfait doit évidemment être bénéfique. C'est vrai de manière triviale, mais les améliorations sont aussi bonnes parce que les choses que nous appelons "améliorations" font du bien. »⁶⁰

Alors, la neuroamélioration, rêve ou cauchemar?⁶¹

⁵⁸ C. Elliott, *Better than Well*, New York, Norton & Company, Inc., 2003 p. 298.

⁵⁹ Cité in R. Goodman, « Cognitive Enhancement, Cheating, and Accomplisment », *Kennedy Institute of Ethics Journal*, 2010/2, p. 147.

⁶⁰ J. Harris, *Enhancing Evolution*, Princeton, Princeton UP, 2007, p. 185.

⁶¹ C'est la question à laquelle j'ai essayé de répondre dans « Devenir un être humain accompli: idéal ou cauchemar ? », in J.-N. Missa & L. Perbal, dir., « *Enhancement »*.

Conclusion

J'ai essayé de montrer et surtout d'illustrer ce qu'était la neuroéthique en développant quatre questions: le dualisme psycho-physique, la nature de l'amour, la nature de l'éthique et l'amélioration de l'être humain. Ces développements montrent que, à l'évidence, cette discipline concerne « l'humanité de l'homme ». Cette humanité, les neurosciences la mettent en question, mais en s'inscrivant dans une longue tradition, celle de la philosophie. Ce qu'elles ajoutent à cette dernière, c'est d'une part la possibilité d'un éclairage expérimental sur le fonctionnement moral de l'être humain et d'autre part une action transformatrice / amélioratrice de cette humanité sur laquelle l'éthique doit se prononcer. La neuroéthique a maintenant quelques années derrière elle et les réflexions qu'elle mène sont riches et établies institutionnellement. Elle est très probablement promise à un bel avenir, puisque les neurosciences le sont.

L'humanité de l'homme bouleversée par les neurosciences : Table ronde finale

avec : Bernard Baertschi, Alexandre Ganoczy, Bernard Feltz, Jean-Michel Maldamé, Michel Simon, Eric Charmetant, Thierry Magnin, Jean-Paul Banquet. Modérateur : Philippe Deterre¹

Philippe Deterre:

A partir des questions que vous avez posées par écrit et que j'ai la charge de reprendre dans cette table-ronde, nous allons d'abord poser deux questions aux trois conférenciers sur

- 1. Les différences entre sujet, personne et âme
- 2. Qu'est ce que les neurosciences apportent vraiment à la question de la morale et à la question de la responsabilité. Avons-nous vraiment besoin des neurosciences?

Les différences entre sujet, personne et âme

Bernard Feltz: Pour faire rapide, j'aurais tendance à dire que le concept de sujet et celui de personne sont très proches. Par contre le concept d'âme me parait un peu périphérique car il est chargé d'histoire et de complexité. Il est utilisé pour caractériser l'humain dans notre culture, en partant de Platon et de Descartes. C'est donc un concept qui doit être travaillé parce que dans la pensée aristotélicienne le concept d'âme veut dire vide. En bref sujet et personne sont à peu près équivalents mais la signification du concept d'âme doit se négocier.

¹ Notes prises et rédigées par Jean Leroy, et relues par Philippe Deterre.

Alexandre Ganoczy: Nous avons discuté d'âme et de sujet dans le carrefour avec Jean-Michel Maldamé et nous avons constaté que ce terme pouvait avoir beaucoup de sens différents. Nous avons esquissé trois approches selon qu'elles sont plus intellectualistes, plus axées sur l'expérience religieuse, plus théologiques. Ce que je voudrais rajouter en tant que théologien c'est que nous aurions dû parler beaucoup plus de la notion de personne, d'autant plus que la tradition philosophique nous présente deux grandes lignes, l'une qui met l'accent sur l'individualité et l'autre qui met en valeur la relation. En théologie, la conception relationnelle de la personne est à privilégier depuis que Tertullien a développé une doctrine de la Trinité en se basant justement sur une notion relationnelle de la personne. A partir du Père du Fils et du Saint Esprit il a dit analogiquement on pouvait les appeler personnes à cause de leurs relations d'amour mutuel. Dans les temps modernes il v a eu une période où on a accentué l'individualité et j'ai l'impression que maintenant on revient à une synthèse entre ces deux tendances. Le concept de sujet me paraît assez connoté par l'existentialisme et j'aurais tendance à le laisser tomber; ce qui m'intéresse, c'est la notion d'agent capable d'agir. La notion d'âme, il faut peut être la laisser tomber et la remplacer par les états psychologiques, états mentaux ou les états cérébraux et leurs relations. Cela me paraît plus fructueux. Pour le concept de personne, il y a une tradition théologique et une tradition philosophique qui remonte à Boèce où la personne n'est pas définie par référence aux trois hypostases de la Trinité mais comme un personnage doué de raison, et ce concept est très important dans la philosophie occidentale parce que, si on est une personne, on ne peut pas être traité n'importe comment, contrairement aux choses et aux animaux. Il y a une dignité attachée à la notion de personne qu'il est important de conserver dans notre époque où on a de plus en plus de mal à distinguer l'homme de l'animal. Les capacités qui font la personne sont la raison, la conscience de soi, la conscience morale qui sont des propriétés que l'on valorise, depuis les Grecs en tous cas. Et il me paraît important de conserver cela au centre d'une réflexion éthique.

Jean-Michel Maldamé: Je suis évidement concerné par la question puisque j'avais proposé un groupe de travail sur l'âme. La relation sujet/âme relève de trois écoles de pensée très différentes qui se recoupent malgré des approches différentes. J'ai pensé qu'il fallait reprendre la notion d'âme parce qu'elle permet de dire ce que le sujet ou la personne ne peuvent pas dire aussi bien. Cette option a donné lieu dans notre groupe de travail à des débats vifs et animés, contrastés, ce qui prouve qu'il était intéressant d'en débattre. Mais nous n'avons pas conclu parce que personne ne pouvait avoir le dernier mot. Nous parlons des sciences et des neurosciences, mais il n'y a pas de travail scientifique qui ne repose sur des options de méthode et aussi sur une vision du monde. Dans cette vision il y a forcément de la philosophie de la nature mais aussi quelque chose qui est déjà de la métaphysique. Souvent ces options sont implicites, non nommées en raison, ce qui nuit à l'objectivité requise par présentation des résultats. Il n'empêche que ces options philosophiques et métaphysiques sont présentes. La réflexion de notre petit groupe sur la notion d'âme a été une incitation à dévoiler ce qui était caché. Les notions de sujet, personne et âme sont utiles pour le dévoilement des présupposés implicites qui sont souvent niés, ce qui est une façon d'affirmer la négation de ce qu'on rejette. La difficulté qui est apparue dans notre groupe est que s'il y a des différences d'appréciation entre ces termes, on doit se demander si elles sont une cacophonie qui renvoie à un non sens ou à un arbitraire, ou bien si ces divergences manifestent que le langage ne peut pas maîtriser totalement ce dont il parle. Il faut alors adopter une attitude différente proche de la méthode expérimentale.

Michel Simon : Si on pense à la vie psychique de l'homme, il y a avec la notion de personne la question de l'individualité et de la « relationalité », et il y a aussi (ce qui a émergé du carrefour proposé pendant ce colloque) le point commun entre neurosciences et psychanalyse : la notion de singularité qui est reprise par Isabelle Stengers et Julia Kristeva et c'est peut-être cela qu'il est intéressant d'introduire dans la trilogie sujet – personne - âme.

Les neurosciences, la morale et la responsabilité

Philippe Deterre: Nous venons de terminer un colloque sur les neurosciences et leurs conséquences sur la manière de penser l'humanité. La question qui se pose est la suivante: « Qu'est-ce que nous apportent les neurosciences sur les questions de la morale et de la responsabilité humaine? »

Alexandre Ganoczy: Je voudrais d'abord poser une distinction entre morale et éthique. Pour moi l'éthique est plus associée à la responsabilité des personnes dans leurs fonctions, tandis que la notion de morale est associée à un certain code qui est moins existentiel que la notion d'éthique. L'apport des neurosciences dans ces domaines serait de nous apporter une nouvelle compréhension de l'engagement éthique du sujet en soulignant les conditions neurobiologiques de la prise de conscience et de la responsabilité.

Bernard Feltz: Dans la question de la responsabilité, je trouve que l'apport des neurosciences est soit problématique soit positif sans qu'il soit facile de distinguer entre les deux. Mais ma réticence par rapport à la position de Spinoza est que si on construit une anthropologie basée sur le déterminisme intégral où on prétend qu'il n'y a pas de libre arbitre, on doit organiser la société sur la base d'un autre concept que celui de la responsabilité. On peut fermer les prisons et ouvrir de nouveaux hôpitaux psychiatriques. La délinquance serait alors un problème de fonctionnement et pas de responsabilité. S'il y a un enjeu au niveau philosophique, c'est qu'on paie trois cents ans plus tard. Appliquer à l'homme le concept cartésien d'animal machine revient à laisser tomber le concept de liberté, et à terme le concept de responsabilité sur le plan éthique (sans parler du plan juridique). Et il me semble que cela amène à une modification profonde de l'homme et des relations sociales.

Alexandre Ganoczy: Je crois beaucoup à la liberté, et en lisant la littérature neuroscientifique j'ai très rarement rencontré la notion de liberté. J'ai lu aussi un livre de 350 pages sur la décision et je n'ai pas trouvé une seule fois le mot liberté. Il y a là un problème, pourquoi cette espèce de réserve des neurobiologistes sur la notion de liberté alors qu'on parle de décision. Pourquoi la notion de liberté est-elle devenue très discutable? Est-ce un nouveau déterminisme ou une mauvaise compréhension de ce qu'est la liberté? La liberté ne consiste pas seulement dans la possibilité de choisir entre ceci et cela, il y a aussi l'acte libre qui consiste à ne pas être dépendant d'une contrainte extérieure, surtout dans la perspective théologique où on est libre pour faire quelque chose pour quelqu'un, pour les autres.

Bernard Baertschi: Trois remarques sur la question des apports possibles des neurosciences. Une étude par imagerie nous dit éventuellement ce que les gens pensent mais pas ce que l'on doit faire, il faut bien distinguer entre les deux. Ce n'est pas un chapitre de psychologie morale. Par exemple dans le test où des ouvriers peuvent se faire écraser par un train incontrôlable. ce n'est pas parce que 12% de gens interrogés disent qu'il faut pousser le gros homme sous le train pour sauver les 5 autres qu'on peut en déduire si c'est moral ou non de le pousser. Il faut donc bien distinguer la question normative et la question empirique, les deux sont reliés bien sûr parce que ce qui nous intéresse est de construire une éthique qui soit appropriée pour les êtres humains, mais on ne va pas lire l'éthique par les réactions qu'ont les gens. La deuxième remarque concerne une étude qui a été faite : on a donné des vignettes aux étudiants, l'une avec un déterminisme complet et avec l'autre on a accepté qu'ils pouvaient tricher. On a constaté que ceux qui avaient des vignettes déterministes trichaient plus facilement que les autres. Donc même si la liberté est une illusion, la croyance dans sa réalité est utile. Quelqu'un a pu dire au 19^{ème} siècle, « si le libre arbitre n'existe pas, il faut surtout ne pas le dire sinon les gens deviendraient tous des crapules ». Ce genre de réflexion est ancien car Diogène, 300 ans avant J.C., le disait à propos d'un esclave : celuici, ayant volé se faisait fouetter : il protestait en disant que c'était son destin de

voler; on lui répondait que c'était aussi son destin de se faire fouetter. On peut aussi citer Luther répondant aux critiques du libre arbitre d'Erasme, en parlant du serf-arbitre. Calvin a aussi des propos sur la prédestination, qui ne laisse pas beaucoup de latitude à la liberté humaine. Cela n'empêche pas les tribunaux de fonctionner car ils utilisent un droit, du point de vue juridique, qui juge pour savoir si quelqu'un est coupable et s'il est capable de comprendre ce qu'il avait fait, s'il avait la capacité de diriger ses actions au sens où il n'était pas sous une contrainte ou dans un délire. Si on croit au libre arbitre au sens métaphysique du terme tout le monde est capable de distinguer une action addictive d'une action qui vient de nous-mêmes. A mon sens on peut conserver toutes nos institutions morales et juridiques et être pour ou contre le libre arbitre au niveau métaphysique, ce sont deux problèmes qui sont différents

Éric Charmetant: À mon sens les neurosciences pourraient éclairer quelque chose qui est mal compris, c'est la question de la psychologie morale et du développement moral chez l'enfant. On pourrait ainsi évaluer de manière plus précise le fonctionnement de l'empathie au moment de la mise en place de la moralité chez l'enfant. On a envisagé que c'est ce qui manquerait chez les psychopathes. Pour rendre compte de ce qui est à l'œuvre dans les paradoxes entre les kantiens et les utilitaristes, il faut considérer que le kantisme insiste beaucoup sur la raison dans des situations à forte teneur émotionnelle, alors que les utilitaristes pensent que cette teneur émotionnelle était sous estimée par le kantisme. Au contraire, l'utilitarisme qui insiste beaucoup sur la place du plaisir et du déplaisir, interviendrait davantage dans les situations où on est moins engagé émotionnellement. Cette bipolarité émotion - raison, qui est une vieille question en philosophie, pourrait peut-être avancer davantage grâce aux neurosciences. Le deuxième point qui pourrait être éclairé concerne le niveau du degré de responsabilité des personnes, ce qui n'est pas sans poser de questions sur leur utilisation dans certains procès aux USA. Peut-on faire appel à de l'imagerie cérébrale pour dire par exemple que s'il y a des lésions, la responsabilité est atténuée? Ce serait à mon sens un

apport possible. Mais il reste que, du point de vue de l'éthique évolutionniste, on ne peut pas attendre des neurosciences une révolution de l'éthique ou l'émergence d'une nouvelle théorie morale.

Philippe Deterre : Une nouvelle question sur le monisme : sachant qu'à la mort le cerveau est entièrement détruit, comment la liberté peut-elle exister dans une conception moniste ?

Bernard Feltz: Le monisme s'oppose au dualisme à la manière de Descartes, Platon et Augustin qui pensent qu'il existe une âme qui est une substance différente du corps. Le monisme est une conception qui donne au contraire le moyen de penser l'humain sans recourir à une âme qui serait une substance différente du corps. Pour le chrétien qui est moniste la question est : que se passe-t-il au moment de la mort? Je ne suis pas théologien, mais comme intellectuel je dirai simplement que l'immortalité de l'âme ne se trouve pas dans les grands textes fondateurs du christianisme. Ce qui se trouve dans ces textes, c'est la résurrection des morts, ce qui renvoie à la restauration de la personne, et non pas à une instance qui ne serait pas la totalité de la personne qui a vécu sur terre. D'ailleurs je dis souvent aux étudiants que l'immortalité de l'âme était une évidence chez les Grecs aux premiers siècles, ce qui annulait pour eux la nouvelle apportée par la résurrection du Christ. La problématique anthropologique des premiers siècles était précisément de lutter contre le gnosticisme et le dualisme, et ce n'est que vers le 4ème siècle, notamment avec Augustin, que ces conceptions ont été reprises. Donc je pense que le fond de la culture chrétienne est moniste, et il me semble que le fond de la pensée de Thomas d'Aquin est clairement moniste. Il fait une concession à l'immortalité de l'âme dans la ligne de pensée d'Aristote pour lequel l'âme est la forme du corps; et il y a différents types d'âme, végétative pour les végétaux, sensitive et motrice pour les animaux, l'âme intellective pour les humains qui permet le langage la pensée etc. Elle est mortelle comme le reste. Pour Aristote le seul problème est que l'âme intellective donnant accès à la connaissance donne accès à une vérité universelle, alors comment l'homme peut-il avoir accès à

une vérité universelle si son âme n'est que la forme d'un corps particulier? Précisément pour avoir accès à l'universel il est amené à distinguer « l'intellect agent » qui est un principe universel immortel et impersonnel, et « l'intellect patient » qui est la forme de chaque individu, donc personnel et mortel, mais lui permet d'avoir un accès à l'universel par l'intermédiaire de l'intellect agent. Mais chez Aristote l'intellect patient est parfaitement mortel. Au 13ème siècle, Thomas d'Aquin réintègre la problématique aristotélicienne en changeant une seule chose : il reprend la tradition augustinienne qui affirme que l'âme, l'âme intellective et l'âme personnelle font l'objet d'une création spéciale par Dieu. Mais pour Aristote, l'âme est la forme du corps. La substance d'un être, c'est-à-dire sa réalité, est l'union d'une forme et d'une matière. Donc si pour Thomas l'âme est immortelle, sans corps elle est une modalité complètement inférieure de l'être humain. On peut dire les choses autrement: dans le concept de formation, il v a celui de forme donc l'immortalité, c'est que Dieu a dans sa mémoire toutes les informations concernant tout le monde, c'est donc l'immortalité de la forme, mais cela ne suffit pas ; il faut aussi le concept de résurrection qui représente l'idée d'une restauration de la personne dans sa totalité y compris dans sa matérialité parce qu'un être humain ne peut pas exister sans matérialité. C'est cela la conception moniste, juive au point de départ et c'est cela que Thomas d'Aguin a essayé de restaurer

Troisième question : celle de la liberté. C'est une question importante bien que paradoxale parce qu'elle se construit. Je pense que dans la problématique qui consiste à dire ou bien c'est le cerveau qui est la cause, ou bien ce sont les émotions qui sont la cause, on a l'alternative suivante. Si on prend le point de vue synchronique on aurait tendance à dire que c'est le cerveau qui est la cause, mais on doit prendre en compte le fait qu'un cerveau n'existe pas par lui-même, il est dans un corps, il a toute une histoire, les émotions sont aussi le produit de notre culture. Même la vie amoureuse est conditionnée par notre vécu dans une famille depuis la naissance qui ellemême fait partie d'un clan et tout cela structure nos émotions. Autrement dit l'idée qui est associée à la problématique ci-dessus est qu'un cerveau est

toujours conditionné par une histoire y compris dans les émotions et l'éducation à l'éthique. Un autre aspect qui est souvent sous estimé, c'est l'importance du langage ; ce n'est pas par hasard qu'Aristote dit que l'humain est un animal doué du langage, qui est précisément produit par le cerveau mais qui cependant échappe complètement à son déterminisme. Tout au moins au sens du déterminisme d'un réseau d'automates booléens qui sont censés représenter le cerveau. Il me semble que ce qui est important dans cette problématique c'est la liberté comme logique de l'autodétermination; c'est par exemple la possibilité qu'avait un être humain de ne pas faire l'erreur qu'il vient de faire. L'idée c'est de se dire que le langage constitue une sphère qui a son autonomie propre et qui confère à l'humain une sorte de distanciation par rapport aux contraintes de ses réseaux neuronaux. Alexandre Ganoczy a cité plusieurs fois cette idée de Damasio. Les contraintes du langage se superposent aux contraintes des réseaux neuronaux ce qui laisse un espace pour une pluralité de choix en face d'une décision à prendre, c'est cela le concept de liberté. Il me semble qu'il y a deux médiations nécessaires : c'est une complexité supérieure et l'accès au langage qui donnent accès à une autre sphère qui a sa logique et en retour peut structurer le cerveau. On est alors dans un espace moniste, car c'est effectivement le cerveau et le corps qui ont un impact décisif, mais puisque c'est le corps lui-même qui est marqué par toute la problématique langagière, on est dans un espace moniste mais qui échappe au déterminisme neurobiologique.

Alexandre Ganoczy: J'adhère à ce qui vient d'être dit mais je préfère le terme holistique à moniste parce que dans celui-ci il y a monos, ce qui renvoie à l'un isolé, sans relation alors que holos réfère à un ensemble dans lequel on peut voir tous les aspects qui concourent pour gouverner nos décisions. L'anthropologie biblique a cet avantage que chaque aspect de l'homme est désigné par des termes qui ont plusieurs sens ce qui permet de signifier la complexité et la diversité de ce qui constitue la personne. C'est pourquoi je préfère holistique à moniste. En ce qui concerne la liberté je salue la perspective historique et autobiographique de la décision libre. Je crois qu'il ne

faut pas exagérer la tendance qui consiste à trouver la liberté dans la décision instantanée qui occulte tout le processus qui conditionne cette décision.

Michel Simon: Une petite question sur le monisme. Mes souvenirs sur Aristote sont un peu lointains, mais il me semble qu'il y a un texte qui parle du noos venu d'ailleurs.

Bernard Feltz; Il v a effectivement des "intellects agents" dont le principe est universel. Mais ce qui me paraît important c'est que l'âme "intellective agent", étant un principe d'universalité, n'est pas une propriété du corps. Le dualisme d'Aristote n'est pas anthropologique. L'intellect patient fait partie de l'humain ; l'âme intellective patiente de l'humain a une connaturalité avec l'intellect agent qui lui permet de percevoir des lumières venant de l'intellect agent. Le *noos* effectivement n'est pas anthropologique mais il est ce qui permet à chaque individu corporel d'avoir accès à l'universel. Aristote est le fondateur de la métaphysique et par son raisonnement il accède à une vérité universelle bien qu'il soit corporel donc mortel et contingent. Voilà comment Aristote arrive à concilier la finitude humaine et son accès possible à l'universel, ce n'est pas parfait mais c'était un progrès pour l'époque. Pour les Grecs, la question épistémologique est beaucoup plus centrale que la question anthropologique. L'anthropologie d'Aristote est celle d'un biologiste qui a fondé la métaphysique et qui de plus est un disciple de Platon et veut intégrer la possibilité qu'a l'homme de penser les mathématiques, lesquelles relèvent du monde des idées non contingentes.

JP Banquet: Je voudrais savoir quel est l'apport des neurosciences dans ce débat. D'autre part je voudrais savoir pourquoi on a glissé progressivement du monisme spiritualiste au monisme matérialiste.

Bernard Feltz: Un dernier commentaire sur le monisme. Je préfère parler non pas d'holisme mais de conception unitaire de l'humain, ce qui est la même chose. Le monisme est très chargé historiquement et je ne l'utilise pas

au sens philosophique mais simplement pour caractériser une position anti dualiste

François Douchin: Je voudrais profiter de ce qu'on a dit sur Aristote pour souligner la culbute supplémentaire faite par Thomas d'Aquin parce que pour lui l'universel doit être personnalisé, si bien que l'intelligence humaine est attachée à la personne et à l'unité de l'âme. Pour cela il analyse le mécanisme de la connaissance qui fonctionne en intellectualisant des données qui proviennent de la sensibilité. Donc la raison n'existe pas en dehors de liaison intrinsèque à la sensibilité. Comme la sensibilité est le côté personnel et individuel, du coup la raison est embarquée dans ce qu'on appellera la personne. En cela il s'oppose à Averroès qui, lui, est resté fidèle à Aristote avec son noos flottant.

Neurosciences, déontologie, éthique

Philippe Deterre : Je passe à la question suivante qui revient sur l'aspect déontologique des neurosciences. N'y a-t-il pas en neurobiologie des interventions qui sont proches de la manipulation mentale ?

Bernard Baertschi: Dans les travaux de recherche sur le cerveau il existe des règles qui encadrent les actions sur le cerveau humain. Mais pour certaines expériences de comportement psychologique, il est nécessaire de tromper le sujet faute de quoi l'expérience est impossible. En revanche il n'y a pas de limitations pour les recherches sur le cerveau des animaux ce qui permet déjà de faire avancer la compréhension des mécanismes cérébraux. Par exemple on a implanté des électrodes dans le cerveau d'un taureau et on pouvait lui envoyer par télécommande un signal qui amenait l'animal à stopper immédiatement une charge qu'il avait entreprise. On a fait aussi des expériences sur les rats équipés avec des électrodes placées judicieusement dans le cerveau; avec des signaux envoyés par radio fréquence à ces électrodes il était possible de diriger le rat comme une petite voiture

téléguidée. On peut imaginer beaucoup d'applications utiles pour ce genre de technique. Avec des substances psycho actives on peut aussi obtenir des modifications de comportement chez les humains. Par exemple pour calmer les enfants hyper actifs: il y en a beaucoup mais sont-ils tous malades? La question de la manipulation des personnes se pose légitimement. De même les militaires: en Irak, on les a obligés à absorber toutes sortes de substances sans leur demander leur avis ni leur donner d'explications, ce qui pourrait donner lieu à toutes sortes de dérapages. Le meilleur rempart est une démocratie pluraliste et libérale où les gens ont accès à l'information qui leur permet de réagir éventuellement. Mais les manipulations de l'opinion peuvent aussi être faites par le moyen de la propagande dans un contexte où la critique est impossible.

Philippe Deterre : Je passe à une autre question qui a été posée. Est-ce qu'on ne tire pas des conclusions hâtives de certaines expériences de neurobiologie en généralisant imprudemment sur la question de la liberté par exemple ?

Bernard Baertschi: C'est une tentation inévitable car une expérience est en général faite pour vérifier une certaine hypothèse et on veut élargir le plus possible la connaissance ainsi acquise. La discussion et la critique sont les moyens pour limiter les risques d'interprétations indues. Par exemple de l'expérience de Benjamin Libet, qui est intéressante, certains ont conclu que le libre arbitre n'existe pas. La conclusion de Libet lui-même était au contraire que le libre arbitre existe. En fait l'interprétation n'est pas donnée avec l'expérience, il peut y avoir des débats contradictoires sur l'interprétation. Mais l'utilisation de l'IRM dans les tribunaux me semble être un danger parce qu'on utilise une technique qui n'est pas encore au point pour tirer des conclusions qui peuvent être erronées ce qui peut être grave sur le plan judiciaire.

Maurice Sadoulet: Il me semble que, dans ce domaine, l'étendue de notre ignorance étant très grande, il faut exercer la plus grande prudence dans les applications.

Jean-Paul Banquet : J'ai l'impression qu'il y a une tendance à profiter des opérations chirurgicales légitimes sur le cerveau pour greffer dessus des protocoles expérimentaux qui vont bien au-delà de la thérapeutique. Est-ce justifié éthiquement et quels bénéfices en tire-t-on? Les patients sont-ils informés? Je pense notamment aux interventions sur les malades de Parkinson.

Bernard Baertschi: Il y a des connaissances qu'il est impossible d'acquérir autrement, mais en principe on ne peut pas faire cela sans le consentement du patient. Cela est interdit, mais il n'est pas impossible que cet interdit ait été transgressé.

Philippe Deterre : Peut-on revenir sur la question que nous posions au début : Peut-on espérer que les neurosciences puissent un jour nous fournir les bases d'une morale ?

Bernard Feltz: J'ai déjà un peu répondu en citant Auguste Comte, mais on peut aussi dire que les neurosciences nous fournissent des connaissances sur la manière dont les gens décident mais ne donnent pas d'indications sur les bonnes normes. Pour cela d'autres considérations philosophiques, politiques ou autres qui interviennent dans le débat éthique et les neurosciences peuvent donner, dans certains cas, des éléments d'appréciation supplémentaires. D'ailleurs, les normes peuvent-elles être universelles? Elles dépendent de présupposés philosophiques ou religieux, par exemple la casuistique ne donne pas de norme universelle.

Des questions théologiques

Philippe Deterre: Nous passons maintenant à la question théologique. J'ai retenu de l'exposé de Marc Jeannerod, qu'il y a une presque égalité entre sympathie et empathie contrôlée. Il y a ici la controverse ou le dialogue entre Lipps et Scheler. Il me semble que c'est plutôt le point de vue de Scheler qu'il faut retenir². Alexandre pourra compléter s'il le souhaite. Je lui pose déjà la question: Peut-on préciser ce qu'est un Dieu empathique dans une conception trinitaire? Par exemple, un Dieu empathique peut-il laisser sans rétribution ceux qui meurent pour lui?

Alexandre Ganoczy: Il n'est pas très usuel de parler du Dieu empathique, c'est de l'opportunisme qui se présente au moment où un théologien entame un dialogue avec un neurobiologiste. C'est un emprunt du langage de l'interlocuteur pour tenter d'exprimer ce qu'on veut lui proposer. En tous cas je ne tiens pas absolument à ce mot empathique; cependant on pourrait lui donner une signification théologique, car qu'est-ce que l'empathie si ce n'est une traduction de l'amour de Dieu pour les hommes, la philanthropie de Dieu qui se déplace en quelque sorte vers l'homme, entre dans l'homme? On parle de l'inhabitation dans l'homme etc. Dans ce sens il y a une possibilité de donner une signification acceptable à l'expression « Dieu empathique ». Mais pour moi, l'intuition de Scheler est théoriquement impensable parce que, avec la notion d'empathie, je me projette dans l'autre, j'essaye de me mettre à sa place. Avec la notion de sympathie par contre, il y a rencontre, il y a échange, cela correspond beaucoup plus avec ce qu'il y avait autrefois et encore aujourd'hui et que les théologiens expriment par le mot agapê. L'agapê désigne l'amour qui n'est pas d'abord désir, amour captatif, mais l'amour don de soi ou désir d'être avec l'autre pour partager l'existence et la vie de l'autre. D'ailleurs Scheler lui-même dit que dans la sympathie on partage affectivement la vie de l'autre : sentir avec l'autre, sentir ce que l'autre est supposé ressentir c'est quand même autre chose que ce désir de partager, de

² cf l'exposé de Marc Jeannerod dans ce colloque

participer. Si maintenant on pense aux Trois qu'on appelle analogiquement « personnes » avec toutes les précautions qui conviennent pour un tel mystère, on considère qu'il y a une relation mutuelle et réciproque entre les trois personnes divines, la notion de sympathie dit pratiquement la même chose, il y a "naître avec", il y a un partage.

Philippe Deterre : Autre question posée à Alexandre Ganoczy. Quel est le statut de la prière dans une relation telle que vous avez essayé de nous la décrire ? Concerne-t-elle la communication directe corporelle ou la communication indirecte, mentale, ou les deux ?

Alexandre Ganoczy: Je dirais les deux, mais la communication directe est tellement automatique, spontanée que je ne voudrais pas utiliser cette notion pour décrire la prière. Parce que la prière c'est quand même un choix, un acte conscient, elle suppose un minimum de connaissance de qui est en face, on s'adresse à quelqu'un, c'est un acte relationnel. C'est pourquoi je dirais plutôt communication indirecte. En outre, dans cette relation je rentre dans une tradition de croyants priants. Je ne suis pas en tant que personne isolée, je suis toujours directement ou indirectement avec les autres. La prière est une concélébration.

Philippe Deterre: On a bien vu comment vous élaborez une relation entre l'homme et Dieu à partir des neurosciences sociales et des neurosciences psychologiques, mais que devient alors la foi comme grâce ou comme don de Dieu?

Alexandre Ganoczy: Là il y a une analogie anthropologique très facile. Qu'est-ce qui suscite en moi la confiance vis-à-vis de toi? C'est que je t'ai reconnu comme fiable. Cela peut aussi s'appliquer pour la foi confiance. Qu'est-ce que la foi? La foi peut être définie comme tenir pour vrai tel ou tel énoncé, tel ou tel message ou tel ou tel témoignage, je le tiens pour vrai, donc j'y crois, j'accepte mentalement, intellectuellement la réalité de ces faits. Mais

il y a autre chose qui est essentiellement inhumaine, qui est l'acceptation de se mettre en dépendance, s'en remettre à Dieu totalement. C'est une attitude holistique qui prend ma raison, mon affectivité, ma volonté.

Question : Mais la prière n'est-elle pas une vraie connaissance de Dieu ?

Alexandre Ganoczy: Comme pasteur on est mené à valoriser cette relation à Dieu qui est très peu conceptualisée et qu'on observe chez beaucoup de fidèles. C'est pourquoi je suis un peu gêné par votre présentation de la prière beaucoup plus axée sur la connaissance et je trouve important de valoriser cette autre relation qui est plutôt cheminement et qui peut aboutir à une découverte d'une relation plus complète.

Philippe Deterre: Si j'ai bien compris, Alexandre dit qu'il ne peut pas y avoir de prière si on n'a pas une connaissance minimale, mais quel est ce minimum?

Alexandre Ganoczy: Je parle de mon expérience personnelle. J'ai lu l'évangile et je suis aussi théologien en asseyant de différencier ce qui est historique et ce qui est légendaire. Dans ce cas là je découvre Jésus, et c'est lui qui me fait prier. Et sa prière est le modèle de notre prière, le Notre Père. Sommes-nous d'accord sur cette façon de voir?

Thierry Magnin: D'un point de vue anthropologique ternaire, qui garde une unité sans découper en trois morceaux (corps âme esprit), la fine pointe de l'homme c'est l'âme. Alors est-ce que cette anthropologie qui est importante dans le domaine de la santé, pourrait-on l'intégrer dans ta vision?

Alexandre Ganoczy: Mais bien sûr. Quand j'étais au séminaire, pendant toute ma formation cléricale j'ai été abreuvé de toute la spiritualité de l'école française selon laquelle la fine pointe de l'âme rejoignait l'aspect

mystique de la prière. Cette spiritualité tripartite ne m'est pas du tout étrangère et c'est une possibilité pour fonder une théologie de la prière, je ne la récuse pas du tout. Seulement il y a des alternatives qui correspondent mieux par exemple à la mentalité de Jean-Pierre Changeux qui me demande « Comment te débrouilles—tu quand tu pries ? Pourquoi pries-tu ? » Il faut alors une autre conceptualité que celle de la fine point de l'âme.

Philippe Deterre : On n'a pas répondu à la question de Michel Simon sur l'individualité singulière. Je pose la question au théologien.

Alexandre Ganoczy Pouvez-vous définir ce que vous entendez par singularité ?

Michel Simon: Dans tout ce que nous avons entendu on ne fait jamais référence à quelque chose d'immuable et permanent qu'on appelle le soi.

Alexandre Ganoczy: Ma réponse est la suivante: le soi ne tombe pas du ciel, il n'est pas une substance isolable. Le soi est essentiellement historique et autobiographique comme l'a montré de façon crédible Damasio. Le soi devient et est en devenir. Je crois que la singularité de la personne humaine est quelque chose qui est le fruit d'une évolution qui est historiquement descriptible. Si la singularité est ce que je suis devenu comme soi, je dis que c'est très important et c'est le fondement de la personnalité.

A-t-on besoin d'un Dieu pour expliquer le comportement humain ?

Philippe Deterre: Maintenant, une question pour Jean-Michel Maldamé: a-t-on besoin d'un Dieu pour expliquer le comportement humain? Y a-t-il un lien entre la création et l'évolution?

Jean-Michel Maldamé: Alors distinguons deux parties dans la question : d'une part, a-t-on besoin de Dieu ? Et d'autre part, quel est le lien entre création et évolution? Je dirais que la deuxième question, répond à la première. Dans la mesure où Dieu crée un être dans son autonomie cela implique que l'homme peut être compris pour lui-même dans son ordre propre. Il peut donc y avoir une anthropologie, une morale, une éthique, une science qui ne fait pas référence à Dieu dans son propre développement. Ainsi, le comportement humain peut s'étudier dans l'objectivité de l'observation qui n'est pas directement théologique. Ceci étant dit, le comportement humain renvoie à la personne dans sa singularité, à un sujet, à un être humain dans son unité et cela renvoie évidemment à la question de son origine et de sa fin et par conséquent à la création. On peut donc en parler, mais cela ne doit pas être le bouche trou des comportements humains. Sur la guestion création-évolution, la création est une notion spécifiquement théologique qu'il faut entendre comme telle. C'est pourquoi dans la Bible, le verbe bara, qui signifie créer, n'a que Dieu pour sujet. Le mot évolution désigne bien des choses, le processus, la théorie qui propose une explication scientifique de l'univers et de la diversité des vivants tout en donnant des pistes de recherche pour éclairer les zones d'ombre qu'elle contient. Je viens d'écrire un livre intitulé « Création par évolution »³. Son titre indique que les processus mis en lumière par la théorie de l'évolution montrent un processus d'apparition des êtres vivants, un processus d'apparition de l'homme dans l'arbre des vivants, et que l'action créatrice de Dieu est coextensive de tout ce qu'il y a de bien dans l'histoire des êtres vivants que nous pouvons retracer d'une manière incontestable bien qu'imparfaite. La création n'est pas l'acte qui vient boucher les trous de l'ignorance mais ce qui vient porter le dynamisme du vivant.

Alexandre Ganoczy: Un petit complément: je suis d'accord mais je signale que Darwin lui même a fait la distinction entre création et évolution. Il parle à plusieurs endroits dans ses ouvrages de l'action du créateur et il dit:

³ Création par évolution, Jean-Michel Maldamé, 2011, Editions du Cerf

n'est-il pas plus beau que le créateur ait commencé à faire le monde à partir du très petit et très simple pour que cela se développe et devienne ce que nous voyons aujourd'hui? Darwin a donc eu le soin de rattacher le processus d'évolution au créateur. Darwin était monothéiste mais pas forcément trinitaire. Je voudrais ajouter que la modernité avait raison, je crois, d'inventer la notion de créativité humaine parce que c'est une analogie de ce que Dieu fait en laissant le monde devenir tel qu'il est par l'évolution, c'est à dire rendre possible du nouveau. La création est la nouveauté absolue. Le théologien Pannenberg développe toute une théorie des possibilités qui se réalisent au cours des temps, et du moment que nous inventons du nouveau dans quelque domaine que ce soit, nous sommes analogues au Dieu créateur.

Neurosciences et âme humaine

Jean-Michel Maldamé

Comment ne pas se réjouir des progrès accomplis par les sciences de la vie et la médecine à propos de la connaissance du cerveau et de son rôle dans la vie humaine tant au plan personnel que relationnel? En effet, les neurosciences donnent des movens pour soigner, diagnostiquer et traiter des pathologies graves : elles permettent de rendre à des personnes souffrantes l'opportunité de mieux vivre ; elles ouvrent un large champ de promesse en matière de thérapie. La démarche éthique humaniste est attentive à la mise en œuvre de la déontologie médicale au service des patients. Outre ce savoir dont tous espèrent les bienfaits, un point mérite grande attention : le développement des neurosciences donne une meilleure connaissance de la vie; ainsi, elles entrent dans des questions anthropologiques étudiées habituellement en philosophie ou en psychologie. Les neurosciences bouleversent cette approche et servent de fondement à une nouvelle anthropologie. Elles se présentent comme le fondement sûr de l'anthropologie fondamentale et entrent de plein pied dans les débats métaphysiques et tranchent les questions, comme le dit clairement Francisco Varela: « Pendant des millénaires, les hommes ont eu d'eux-mêmes une compréhension spontanée, dépendante de la culture de leur époque; maintenant, pour la première fois, cette vision populaire de l'esprit entre en contact avec la science et en est transformée. 1 ». C'est sur ce point que le recours à la notion d'âme me semble opportun, voire nécessaire.

Une nouvelle philosophie de l'esprit

Aujourd'hui, on parle en effet d'une nouvelle philosophie de l'esprit fondée sur les sciences cognitives et tout particulièrement les neurosciences².

_

¹ Francisco VARELA, Connaître. Les sciences cognitives : tendances et perspectives, Paris, Édit. du Seuil, 1989.

² Une introduction à cette nouvelle perspective philosophique est donnée dans une anthologie : Denis FISETTE et Pierre POIRIER, *Philosophie de l'esprit*. Textes clés de la

Le discours qui est tenu est très développé dans le monde nord-américain; il entre en Europe où il bouleverse le champ des études anthropologiques, tant en métaphysique, qu'en psychologie et dans les sciences humaines³. Cette nouvelle philosophie s'oppose à la représentation commune au titre d'une exigence de scientificité. Cette philosophie récuse la notion traditionnelle d'esprit⁴, liée à l'immatérialité. Elle s'oppose aussi à la psychologie qui place au principe de l'action des intentions qui sont d'un autre ordre que la mise en œuvre des « états mentaux » pour les remplacer par des états physiologiques,

philosophie de l'esprit, t. I, *Psychologie du sens commun et sciences de l'esprit*; t. II, *Problèmes et perspectives*, Paris, Vrin, 2002 et 2003.

Le premier volume est divisé en deux parties : « La psychologie populaire : mythe ou réalité ? » et « Conceptions de l'esprit ». Dans la première partie, la parole est donnée à des philosophes, Wilfrid Sellars, Paul Churchland et Daniel Denett ; la notion de « psychologie populaire » désigne les conceptions les plus représentatives de la psychologie anglo-saxonne au milieu du XX^e siècle qui parle d'états mentaux pour comprendre le comportement humain. La deuxième partie donne la traduction des textes de quelques pionniers de la discipline que sont Williard Quine, Donald Davidson, Hilary Putnam, David Lewis, Jerry Fodor et Carl Hempel. Avec ces auteurs on voit paraître une grande diversité des approches : béhaviorisme, fonctionnalisme, théorie représentationnelle de l'esprit et des problèmes auxquels ils se sont confrontés, en particulier la question de l'identité.

Le second volume porte sur les problèmes fondamentaux de la philosophie de l'esprit et sur les perspectives actuelles liés aux modèles de cognition. La troisième section étudie les traits traditionnellement attribués à la pensée « intentionalité, rationalité et conscience ». Les textes cités sont ceux d'auteurs « classiques » : Hilary Putnam et Fred Drestke sur l'intentionnalité, Stephen Stich sur la rationalité et enfin sur la conscience David Rosenthal et Joseph Levine. La quatrième section a pour thème les débats en cours sur les modèles appropriés à l'étude de l'esprit, à savoir le computationnalisme ou fonctionnalisme computationnel, le connexionisme ou les modèles dynamiques avec Jerry Fodor sur le computationnalisme, puis Zenon Pylyshyn et Paul Smolensky sur le connexionisme et les propositions de Timothy van Gelder.

³ Sur la situation actuelle des diverses anthropologies concurrentes, voir Francis Wolff, *Notre humanité*. *D'Aristote aux neurosciences*, Paris, Fayard, 2010. L'ouvrage présente quatre modèles anthropologiques: celui de Platon et l'école dualiste ou spiritualiste, celui d'Aristote où l'âme est forme du corps, celui du structuralisme où le sujet s'efface et enfin celui des neurosciences où la différence entre l'homme et l'animal s'efface. L'ouvrage montre l'évolution des idées en la matière.

⁴ Notons que la traduction française est délicate : le mot esprit traduit l'anglais *mind*! ce qui est loin d'être évident, tant les cultures sont divergentes entre la tradition philosophique continentale et l'approche empirique ou analytique de la tradition des penseurs d'outre-Manche.

objet de science. Celle-ci repose sur des observations, des mesures de grandeur (entités ou flux d'énergie) et des mises en forme mathématique tant pour les lois biologiques que pour les règles du comportement.

Nous n'examinerons pas ici l'immense littérature consacrée à cette nouvelle philosophie de l'esprit, pour aller au cœur de ce qui fonde l'éthique en posant la question : qu'est-ce que l'homme ?

Une philosophie de l'homme

La philosophie de l'esprit liée aux neurosciences doit être située par rapport à d'autres qu'il serait superficiel de disqualifier parce qu'elles n'utilisent pas les capacités données par la science.

En simplifiant, on distingue dans la pensée humaine plusieurs approches. Une première considère que les réalités qui relèvent du domaine de l'esprit (pensée, raisonnement, désir, motivation...) ne sont pas matérielles. Certes, elles reconnaissent l'importance du corps, mais il n'a qu'un statut d'instrument. Cette perspective, présente chez Platon et dans le christianisme marqué par la tradition platonicienne avec saint Augustin, a été reprise par Descartes lors de la fondation des sciences modernes. Une deuxième perspective considère que l'approche de l'humanité et la détermination du propre de l'homme doit partir de la réalité biologique et donc prendre en compte le corps humain, cet organisme enraciné dans le cosmos, la biosphère, et inscrit dans l'arbre des vivants retracé par la théorie de l'évolution. Cette perspective, rattachée emblématiquement à Aristote, reprise par saint Thomas d'Aquin et aujourd'hui par Teilhard de Chardin, n'a pas de difficulté à parler de l'animal humain⁵. Cette tradition accueille ce qui est fait aujourd'hui comme un progrès incontestable; pour elle, c'est une bonne nouvelle que le développement des sciences du cerveau⁶. Cet accueil n'est cependant pas

⁵ Cf. Jean-Michel MALDAMÉ, *Création par évolution*, Paris, édit. du Cerf, 2011.

⁶ Nous avons consulté les ouvrages de la collection « neurosciences et cognition » éditée chez DeBoeck-Université à Bruxelles, Jean DELACOUR, *Une introduction aux sciences cognitives*, 1998 et *Conscience et cerveau. La nouvelle frontière des neurosciences*, 20001 et Michael S. GAZZANIGA, Richard B. IVRY et George R. MANGUN, *Neurosciences cognitives*. La biologie de l'esprit, 2001.

aveugle et il y a place pour une lecture critique des propos des spécialistes de neurosciences. En effet, l'anthropologie de la tradition aristotélicienne, dans laquelle je m'inscris, ne saurait être confondue avec celle du structuralisme qui efface le sujet humain, ni avec la naturalisation opérée par les sciences cognitives revendiquant le matérialisme comme horizon du savoir⁷. Ceci conduit à une évaluation de la notion de science comme activité humaine et pas seulement comme ensemble de résultats.

L'épistémologie la plus commune montre qu'une science n'est pas une accumulation de résultats, d'observations, de mesures et de mises en forme. Une science suppose une méthode et repose sur une conceptualisation. Quel est le statut de la conceptualisation? Il est clair qu'elle advient comme un prolongement de l'observation ou du calcul, mais ce n'est qu'un aspect, car en réalité, elle joue un rôle fondateur. C'est en effet en fonction d'une certaine conceptualisation que se construisent les processus expérimentaux⁸. Il y a « dialogue expérimental », puisque l'observation et la conceptualisation se bouclent dans un mouvement incessant. Ce trait général vaut pour les neurosciences qui se développent grâce à des concepts fondateurs⁹. Ceux-ci sont intégrés dans des systèmes philosophiques construits pour eux-mêmes¹⁰ et

.

⁷ Le concept de naturalisation est caractéristique de la méthode des neurosciences dans la mesure où elles considèrent que les états mentaux se réduisent à leur inscription physiologique décrite par les neurosciences.

⁸ Sur le statut des concepts, voir Isabelle STENGERS et Judith SCHLANGER, *Les concepts scientifiques*. Invention et pouvoir, Paris, édit. de la Découverte, 1989.

⁹ Il est important de noter que le terme « cerveau » n'a rien de naïf. C'est un concept construit. Pour une mise en perspective générale, voir Bernard Andrieu, *L'Invention du cerveau*. Anthologie des neurosciences, Paris, Pocket, 2002, et le gros volume: François CLARAC et Jean-Pierre TERNAUX, *Encyclopédie historique des neurosciences*, Bruxelles, De Boeck, 2008. On en voit la mise en œuvre dans l'ouvrage collectif de Olivier HOUDÉ, Bernard Mazoyer et Nathalie TZOURIO-MAZOYER, *Cerveau et psychologie*, Paris, PUF, 2002.

Ainsi la philosophie de l'esprit liée aux neurosciences doit beaucoup à l'école néopositiviste du cercle de Vienne et la sortie de cette école est marquée par l'importance de Wittgenstein et de sa philosophie qui donne au langage un statut bien différent que celui d'être une simple représentation de la réalité. La pensée se pense ; elle n'est pas seulement une aptitude à communiquer ; elle est réalisation humaine de soi.

qui contiennent une anthropologie. On ne peut éluder l'examen de ces concepts pour voir les implications éthiques de l'essor des neurosciences.

Le choix des concepts

Je constate que l'anthropologie immanente aux neurosciences, voire fondatrice, repose sur deux types de concepts qu'il importe de distinguer pour notre réflexion sur l'éthique des neurosciences.

Une première famille de concepts s'inscrit dans la perspective de naturalisation évoquée. Seuls les concepts opératoires peuvent y être reçus. Les concepts traditionnels sont transposés dans l'ordre opératoire des sciences expérimentales — il y a de ce fait un obstacle pour un dialogue entre philosophes et scientifiques. La difficulté est de respecter la spécificité de la vie, car il est admis que le vivant ne peut se réduire à la chimie qui est à la base de tout fonctionnement¹¹.

Une autre famille de concepts repose sur l'appartenance de l'humanité au monde des vivants¹². Ils sont bienvenus en neurosciences, tant pour l'observation du cerveau que des comportements selon les procédures de la recherche médicale et pharmaceutique où l'on teste les médicaments sur les animaux avant de le faire pour les êtres humains. La conceptualisation qui en résulte englobe tous les vivants. Ainsi l'on utilise dans les neurosciences des termes comme perception, sensation ou sensibilité, émotion ou réactivité, ou encore conscience. Ces termes traditionnels sont étendus à tout domaine d'observation tant pour passer de l'humanité aux autres animaux qu'inversement. Ainsi on parle en éthologie d'apprentissage ou même d'intelligence. Grâce à ces concepts, l'étude scientifique se développe dans

.

¹¹ Les fondateurs de la biologie moderne le soulignent avec force, Ernst MAYR, *Histoire de la biologie*, Paris, Fayard, 1989.

¹² Qu'est-ce que la vie ? Pour répondre, je propose un chemin qui va de la fonction à l'être. 1° Éléments spécifiques du vivant : atomes, molécules, cellules, organes... 2° Activités caractéristiques du vivant : les verbes se mouvoir, se nourrir, se reproduire. 3° Capacités spécifiques : communiquer, souffrir, connaître, vouloir... 4° Structure : les organes, les systèmes, l'enracinement écologique... 5° Organisation : le plan d'organisation, l'organisation. 6° L'unité et la permanence pour surmonter le temps : mémoire et projet. 7° Principe d'unité : l'âme.

toutes ses dimensions : études médicales, imagerie cérébrale, étude du comportement...

Une telle méthodologie a pour fondement une anthropologie très spécifique. La différence entre l'homme et les animaux (surtout ceux qui sont dits supérieurs) est reconnue; mais cette différence est une différence de degré et non de nature¹³. Les termes comme sensation, émotion, conscience, intelligence... ont le même champ de signification. Cette option mène à penser que l'homme, soumis aux mêmes règles et aux mêmes exigences biologiques et sociales que les autres vivants, n'est pas vraiment différent. Le choix de concepts qui expriment la continuité est donc gros d'une option implicite dont on voit qu'elle modifie les fondements de l'éthique privilégiant la personne dans sa dimension irréductible aux déterminismes somatiques.

Être attentif à l'unité du vivant

Face à cette manière de faire, je propose de recourir à une conceptualisation différente. Il ne s'agit pas d'entrer dans la perspective dualiste, où le choix des concepts exclut *a priori* la pertinence des sciences modernes pour parler du propre de l'homme, sa pensée. Je me situe dans une perspective de type aristotélicien, où les concepts philosophiques sont enracinés dans l'expérience sensible et donc valorisent les termes employés dans les sciences. L'homme est placé dans le monde des vivants; il est reconnu comme « animal », mais en précisant comment cette animalité est spécifiée. On dit de lui qu'il est un animal politique, un animal doué de raison¹⁴. Je constate que cette qualification est philosophique et donc se situe explicitement sur un autre ordre que celui des concepts des neurosciences. C'est là un renversement de la perspective! Il n'y a pas un simple prolongement, mais une création de sens. Quel est ce renversement?

-

¹³ Cf. le colloque *Intelligence animale, intelligence humaine*, Paris, Vrin, 2006; le comble de cette mise en perspective est le fait de ceux qui disent, comme Pascal Picq: « Il est faux de dire que l'homme descend du singe; l'homme est un singe » ; ce trait d'humour est fort significatif d'une mise en parfaite continuité.

¹⁴ Le terme grec est *logos* qui a une grande richesse de significations.

La démarche scientifique procède par analyse. Qui dit analyse dit séparation, distinction, mise en ordre en quête d'une raison des liens entre phénomènes. L'analyse peut porter sur des éléments, mais aussi sur des structures et sur des fonctions, selon une hiérarchie explicative qui est achevée par la considération de l'organisme. Mais c'est toujours une approche par analyse. La démarche philosophique ne part pas des éléments¹⁵; elle part de la totalité en tant qu'elle est unifiée. La philosophie est attentive à l'unité du vivant. C'est à partir de la prise en compte de l'unité du vivant que la démarche philosophique se situe autrement que la méthode des travaux fondateurs des neurosciences. Elle en fait la critique.

Une première critique relève que les neurosciences procèdent à une étude scientifique du cerveau et du système nerveux, mais qu'elles méconnaissent le corps¹⁶. Le philosophe parle du corps et considère que l'approche des neurosciences se limite à une seule dimension de l'activité corporelle, fût-ce la plus noble. Parce que l'observation est de plus en plus précise, elle s'éloigne de la vision d'ensemble en se limitant au cerveau dans ses dimensions sensori-motrices ; elle oublie que le corps est l'unité concrète du vivant.

Une autre manière de faire – dans laquelle je me situe – consiste à voir que les activités étudiées par les neurosciences à propos de la connaissance ou des émotions (perception, affect, sensation, imagination ou conscience) doivent être étudiées en considérant l'unité de l'acte posé (percevoir, être ému, sentir, imaginer ou être conscient). L'acte est saisi pour lui-même et il apparaît qu'il transcende l'analyse des processus étudiés par la science¹⁷. Ce n'est pas seulement un organisme, mais un être vivant.

¹⁵ Sachant qu'il y a divers niveaux pour le choix des éléments : molécule, cellule, tissu, organe, système, organisme...

¹⁶ Cette critique est habituellement développée dans le cadre de la phénoménologie moderne.

¹⁷ On ne parle pas de la vie sans considérer l'apport de la théorie de l'évolution, par exemple John L. BRADSCHAW, Évolution humaine. Une perspective neuropsychologique, Bruxelles, DeBoeck-Université, 2003. La théorie de l'évolution postule la continuité de la croissance de l'arbre des vivants. Le terme qui est alors utilisé est celui d'émergence. Le terme est nouveau dans le monde de la philosophie et

En prenant acte de la dimension unifiée de l'acte humain, cette démarche se réfère à un responsable (un acteur ou un agent) de cet acte. Trois termes sont possibles et ils rencontrent les recherches de neurosciences. On parle souvent de personne ; d'autres parlent de sujet ; il me semble important de ne pas méconnaître l'importance du mot âme. La tradition aristotélicienne a choisi de donner le nom de *psychè* au principe de la vie. Le latin a traduit *psychè* par *anima*, transcrit en français par âme. Le langage commun s'inscrit dans cette filiation quand il emploie le terme « animal » pour dire un vivant. Tout vivant a une âme ; il est vivant par l'âme qui est le principe immanent. Le mot principe est à entendre comme acte unifiant un donné qui est un potentiel. Comme hélas le terme « âme » évoque nécessairement des perspectives dualistes étrangères à la biologie actuelle, précisons que la notion d'âme ici introduite désigne le principe qui unifie le vivant ; ce n'est qu'un principe ; ce n'est pas une substance, ni une entité séparée. L'âme est le principe actif qui fait que le vivant est vivant ; c'est le principe ontologique de son unité¹⁸.

Valeur du terme âme

La décision de recourir à un principe d'unité pour comprendre le vivant permet de surmonter les difficultés de situer l'homme dans le monde des vivants, en respectant les résultats avérés de la théorie de l'évolution. Les notions de personne, de sujet ou d'âme le font quand elles se réfèrent à l'unité. Elles permettent d'introduire une reconnaissance de la spécificité de l'homme.

des sciences. Il est apparu à la fin du XIX^e siècle. Il est très commode, car il permet de dire qu'il y a apparition d'une nouveauté et en même temps respecte la continuité spatio-temporelle et énergétique du phénomène. Le terme est présent dans les neurosciences à propos des activités comme la conscience dont on dit qu'elle est le fruit d'une émergence. Ainsi à un certain seuil de complexité, il y a conscience. Avant, il n'y a pas suffisamment d'éléments ni d'organisation; ensuite c'est un effet optimal qui est mis en œuvre grâce à un principe de sélection. Quel est le statut de ce concept? Il me semble utile dans une description scientifique, mais équivoque pour une détermination de ce qui est spécifique. C'est là la difficulté de la notion de continuité souvent réduite à la contiguïté. Cf. Nancy MURPHY et William STOEGER, Evolution and Emergence. Systems, Organisms, Persons, Oxford, Oxford University Press, 2007.

¹⁸ La notion d'âme est revisitée dans le contexte actuel par Laura Bossi, *Histoire naturelle de l'âme*, Paris, PUF, 2003. Cet ouvrage, très remarquable dans le contexte actuel, montre les soubassements philosophiques des matérialismes contemporains.

L'animal humain comme tout vivant est vivant par le principe de la vie qui fait la singularité du vivant. Or ce principe est irréductible et il est qualifié par les réalisations.

La notion de sujet est employée pour dire la situation qui fonde l'exception humaine. Le terme a un sens juridique ; il fonde le droit où l'on considère qu'un être humain est sujet de droits et de devoirs et donc responsable de lui-même. Dans le même sens et par un tout autre chemin, la notion de sujet est revenue en force dans l'école de pensée phénoménologique pour dire que l'objectivation faite par les sciences humaines ne faisait de lui qu'un agent et dissolvait toute responsabilité¹⁹.

La notion de personne est employée habituellement dans le monde chrétien; elle est d'origine théologique, puisque formée pour dire l'irréductibilité des trois qui sont un seul Dieu. La notion a une dimension politique pour récuser le totalitarisme moderne dans ses diverses formes²⁰.

Nous pensons pourtant que la notion d'âme, entendue dans la tradition aristotélico-thomiste, permet de mieux situer l'humanité dans le monde des vivants. En effet, elle est enracinée dans la biologie où elle dit la singularité d'un être vivant et sa transcendance. Elle est bienvenue dans le contexte d'une pensée qui définit la vie comme auto-constitution d'un vivant. Cette notion est enracinée dans une démarche scientifique qui se veut objective, en ce sens qu'elle prend en compte les aptitudes et les capacités. La notion d'âme comme unité du vivant situe le « propre de l'homme », défini comme « zoon logikon ». L'homme est ici un « animal », mais il l'est non seulement selon mais par le logos. Ce terme désignant la pensée, la raison, la parole et la logique. Le logos est pensée, comme acte mis en œuvre non seulement dans des perceptions, mais dans des jugements et dans le déploiement de la richesse de la parole.

de personne », Bruxelles, DeBoeck-Université, 1999.

¹⁹ Le débat est ouvert dans l'ouvrage de Jean PETITOT, Francisco VARELA, Bernard PACHOUD et Jean-Michel ROY, *Naturaliser la phénoménologie*. Essais sur la phénoménologie contemporaine et les sciences cognitives, Paris, CNRS-éditions, 2002.
²⁰ Le débat est abordé par Stefen REED, *Cognition. Théories et applications*, « Question

L'animal humain est qualifié par le fait qu'il y a en lui un logos, non seulement une langue, voire un langage, mais une parole²¹. Or la parole est un acte ; elle est présence à soi et à autrui. La vie humaine est fondée sur le fait que le sujet humain puisse dire à l'autre : « je te promets... » ou à lui-même : « je m'engage à... ». La parole (logos) comme acte! Telle est la racine de l'humanité de l'homme. La parole humaine n'est pas que communication, elle est relation, présence et reconnaissance. L'enfant qui vient de naître crie et pleure ; la mère entend quelqu'un qui l'appelle et elle répond par les gestes et par la parole. Par les paroles de mère et de père l'enfant advient à son identité et il accède à la parole qui s'adresse à un autre reconnu pour lui-même. Combien d'adolescents ne peuvent qu'user de violence, faute d'avoir entendu une parole dite en vérité? Que de souffrances dans les couples quand la parole fait défaut : la violence advient – tant de femmes battues – et c'est un désert affectif. Que de souffrances dans nos repas de famille lorsque le mutisme s'instaure entre les générations! Or ceci ne s'observe pas ailleurs dans le monde des vivants.

Une approche non dualiste de l'être humain

Attention! Parler d'âme risque d'être source de contresens, si on entend ce mot selon la tradition dualiste, qui pose un esprit immatériel dans un corps qui serait un instrument. Non! La notion aristotélicienne d'âme est le fruit d'une considération globale. Ce n'est pas celle du dualisme où une seule partie de l'être humain est envisagée. Tout en l'être humain participe de ce qui lui est propre, le logos, l'aptitude à la parole. Ainsi l'intelligence humaine est-elle dans tout ce qui est: le cerveau, mais aussi la main, la stature, les yeux... Ou encore le plus humble morceau de chair ou d'os! L'homme a une main, parce qu'il est intelligent et il est intelligent parce qu'il a une main. De même pour la station debout. De même encore pour la forme du crâne, la position des organes de perception sur le visage... Tout ceci mène à une anthropologie qui

²¹ Il y a sur ce point une vive tension entre la psychanalyse et les neurosciences. Sur ce point voir Pierre MAGISTRETTI et François ANSERMET, *Neurosciences et psychanalyse*, Paris, Odile Jacob, 2010.

voit clairement la différence entre l'homme et ses proches sur l'arbre généalogique retracé par l'évolution.

Selon cette notion d'âme, l'être humain est humain en tout ce qu'il est. Pour cette raison, nous récusons la démarche souvent reprise par les chrétiens qui s'appuient sur un texte de saint Paul parlant d'une trilogie : « corps, âme et esprit » pour mettre une stricte hiérarchie entre les trois niveaux et y placer le « surnaturel » théologique. Cette hiérarchie ne doit pas conduire à méconnaître que l'être humain et un, unifié par un principe actif, son âme qui fait de tout ce qu'il est un vivant.

À partir de l'unité de l'être humain due à son principe, il faut manifester ce qui fait l'homme différent des autres animaux. Pour cela, nous devons considérer deux ordres de la vie : connaître et vouloir.

Le propre de l'homme

La connaissance est présente dans le monde animal. En effet, tout animal a besoin de connaissances sur le monde. Certains ont une connaissance consciente – au sens premier du terme. Comme telle elle n'est ni vraie ni fausse; elle est factuelle. Un deuxième degré de connaissance est la connaissance de sa connaissance qui vient du fait qu'il y a non seulement langue ou langage, mais pensée et jugement qui lie entre eux un sujet et un prédicat: le ciel est bleu, le printemps est arrivé, les oiseaux chantent. Il y a alors vérité, au sens premier du terme : ce qui est dit doit être conforme à ce qui est. Mais ceci ne suffit pas. En effet, le jugement posé est vrai ou faux, juste ou injuste... mais pour que ce jugement soit pleinement fondé, l'intelligence doit accéder à la raison objective pour laquelle le jugement est vrai ou faux. C'est le propre du *logos* qui permet de dire « ceci est vrai parce que cela... » ou encore « ceci advient à cause de cela... ». C'est cela qui fait l'homme humain et atteste sa spécificité.

L'homme est animal rationnel, il est aussi animal politique ou sujet éthique. Cela ne signifie pas qu'il soit bon ou mauvais par nature, mais qu'il est appelé à être ce qu'il est par un chemin dont il a la responsabilité. Là encore il faut situer l'homme parmi les vivants.

Une première condition est l'action – par rapport à ce que l'on subit. Là un acteur ou agent pose un acte ; il est la cause d'un effet. Le mot agir est à entendre au sens vital. Un vivant agit, une machine fonctionne. Une action est une conduite intentionnelle, où l'intention est plus ou moins consciente. Les degrés de conscience définissent une hiérarchie dans le monde animal. Il y a donc des espèces animales plus ou moins capables d'action.

Mais une deuxième condition est présente dans l'agir humain. Il ne suffit pas d'être agent ; il faut être sujet, c'est-à-dire, pouvoir agir librement, par soi-même. Il faut pouvoir désirer ou ne pas désirer d'agir. Un agent est libre s'il agit comme il veut, c'est-à-dire s'il agit comme il veut désirer d'agir. On considère qu'il pourrait agir autrement qu'il n'a fait. C'est cette capacité à être le sujet de ses actions qui fait le sujet. Comme il y a jugement, il y a ici désir de désir, ce que la philosophie appelle « volition ».

Il est enfin une troisième condition pour qu'il y ait action humaine. Vouloir, vouloir vouloir, mais encore, vouloir selon ce que le sujet considère comme ce qu'il doit faire. Il doit agir selon des raisons d'agir, des motifs motivés, des désirs de désir désirant. Il doit pouvoir dire de lui-même : « J'agis ainsi parce que c'est bien » ou « il est juste pour telle ou telle raison d'agir ainsi ».

Cette attitude est le propre de l'homme qui se libère de la pesanteur du groupe auquel il appartient. Le fondement de la vie humaine se dévoile avec Antigone ou avec Socrate en prison; leur action est au-delà des intérêts, de la satisfaction des besoins fondamentaux du vivant, de l'obéissance à des règles sociales ou à des lois, mais dans la relation à un au-delà du visible.

Pourquoi cela est-il possible, sinon parce que le principe qui fait l'homme humain transcende l'ordre de l'utile, pour accéder à l'infini de la valeur et de l'esprit.

Conclusion

Le maître-mot d'une anthropologie revitalisée par les apports des neurosciences me semble être la notion de responsabilité. Les neurosciences comme travail scientifique sont dans un souci d'objectivité. Elles se situent donc dans une dynamique de relation. Cette dynamique est née dans le contexte médical du souci de l'autre. Son objectivité est le corrélat de la reconnaissance de l'autre comme autre en son altérité.

Ceci est manifeste dans ce qui est dit de la parole : une parole adressée à un tiers reconnu pour ce qu'il est dans une irréductible altérité²². Le « grand autre » de la psychanalyse est au fondement de l'identité humaine. L'être humain s'adresse à un autre homme en le considérant autrement que comme le pourvoyeur de ses besoins²³.

La dimension chrétienne de l'existence s'inscrit dans cette perspective²⁴. Pour les chrétiens, le lieu où Dieu se rencontre est indiqué par Jésus déclarant : « Ce que vous avez fait au plus petit d'entre les miens, c'est à moi que vous l'avez fait ».

-

²² Telle est la raison de notre malaise avec la démarche réductrice des fondateurs des neurosciences comme Churchland ou Fodor.

²³ Nous voyons un piège narcissique dans la remarque suivante : « J'assimile la notion de spirituel à une expérience intense de l'harmonie, au sentiment que l'organisme fonctionne avec la perfection la plus grande possible. Cette expérience s'associe au désir d'agir à l'égard des autres avec bienveillance et générosité. Ainsi avoir une expérience spirituelle, c'est éprouver durablement des sentiments dominés par telle ou telle variante de la joie, même si elle est sereine », Antonio DAMASIO, *Spinoza avait raison. Joie et tristesse, le cerveau des émotions*, Paris, Odile Jacob, 2003, p. 280.

²⁴ Pour une rencontre entre théologie chrétienne et neurosciences, voir Alexandre GANOCZY, *Christianisme et neurosciences*. Pour une théologie de l'animal humain, Paris, Odile Jacob, 2008; voir aussi l'ouvrage collectif John RUSSEL, Nancy MURPHY, Theo C. MEYERING, Michael A. ARBIB, *Neuroscience and the Person*. « Scientific Perspectives on Divine Action », Observatoire du Vatican et Center for Theology and the Natural Sciences, Vatican et Berkeley, California, 1999.

Compte-rendu du carrefour "Neurosciences et âme humaine"

Brice de Malherbe

Jean-Michel Maldamé introduit le carrefour en appelant à se réjouir de l'apport des neurosciences. Sur le plan thérapeutique d'abord, où elles donnent des moyens pour diagnostiquer et traiter des pathologies graves. En anthropologie ensuite, par la stimulation d'une nouvelle « philosophie de l'esprit » qui aide à reposer la question anthropologique fondamentale : qu'estce que l'homme ?

Face à cette question, une première approche, héritière de Platon, Augustin, Descartes, insiste sur l'immatérialité de l'esprit. Le corps est réduit au statut d'instrument. Une deuxième considère que tout est matériel, tant les objets que leurs relations — légitimement quantifiées dans des modèles mathématiques. Une troisième, à la suite d'Aristote, Thomas d'Aquin et Teilhard de Chardin, enracine l'esprit dans le donné corporel qui rattache l'homme à l'ensemble des vivants. Cette tradition accueille les résultats des neurosciences aujourd'hui comme un progrès heureux, mais invite à une lucidité sur les concepts anthropologiques qui nourrissent l'interprétation des phénomènes observés.

Dans notre domaine, deux familles conceptuelles peuvent être distinguées. La première exprime la continuité entre les vivants. Ainsi les neurosciences utilisent des termes comme perception, sensation, émotion, conscience, intelligence. Grâce à ces concepts, l'étude scientifique s'appuie sur l'éthologie et sur l'étude de l'évolution du cerveau. Une telle philosophie expérimentale a pour effet de présenter une anthropologie où la différence entre l'homme et l'animal est considérée comme une simple différence de degré puisque les termes utilisés ont le même sens chez l'homme comme chez tout autre animal.

À l'inverse, une deuxième famille de pensée voit la différence entre l'homme et l'animal comme une différence de nature. Les termes employés

dans les sciences ne sont pas exclus, mais situés dans une perspective proprement philosophique. La démarche analytique des sciences, attentive aux divers éléments constituant les vivants est replacée dans l'attention à l'unité du vivant. Ainsi le philosophe parle du corps et considère que les neurosciences se limitent à une seule dimension de l'activité vitale à partir du cerveau. Ou encore, quand les neurosciences parlent de connaissance (perception) le philosophe parle de réalités vécues (unité de l'acte de percevoir). Dans cette perspective, ce qui fait l'homme humain n'est pas identique à ce qui fait les autres animaux selon leur espèce. Il y a une différence de nature.

Pour exprimer ce qui fait le propre de l'homme, plusieurs termes sont employés. Le plus fréquent est celui de personne; on parle aussi de sujet. Pourquoi ne pas oser parler de l'âme? Dans la tradition aristotélicienne la notion d'âme désigne non pas une substance séparée mais le principe d'unification du vivant. La notion a le mérite de dire aussi la singularité et l'irréductible individualité de chaque être humain. Entendu dans cette tradition, le terme peut donc être repris pour assumer les résultats des neurosciences et la dimension provocatrice de cet usage invite à sortir des équivoques du monisme matérialiste.

Le débat qui a suivi cette introduction a pris trois orientations :

La première sur la pertinence de l'utilisation du mot « âme ». Pour certains, si l'insistance sur l'unité du vivant est opportune, la notion d'âme est à éviter dans le dialogue avec les scientifiques. L'unité de l'être humain en tant que structure où le tout est plus que la somme des parties est un donné observable. L'âme quant à elle renvoie notamment à l'expérience religieuse. Il conviendrait peut-être de se rapprocher de la notion anglaise de *mind*, qui désigne une réalité incarnée. L'on pourrait parler du « mental » qui rend compte de l'enracinement sensible de l'activité de l'intelligence comme de la volonté.

D'autres répondent qu'avant d'écarter le mot « âme » il faudrait en retrouver la richesse de sens chez Aristote qui prend en compte les dimensions végétative, sensible et spirituelle chez l'être humain. Edith Stein quant à elle

insiste sur la compénétration de l'âme et du corps. Un rapprochement pourrait être fait avec les notions bibliques de *lev* – le « cœur » –, de *nepesh* – « l'âme » comme principe vital –, de *ruah* – le « souffle » qui n'est pas propre à l'homme. En hébreu toute cette terminologie renvoie à l'expérience concrète.

La deuxième piste est revenue sur la distinction méthodologique entre science et philosophie ou théologie. La connaissance scientifique se fonde sur des faits vérifiables; la connaissance philosophique n'a-t-elle pas tendance à avancer des affirmations à partir d'expériences invérifiables? Sans doute faut-il reconnaître la diversité épistémologique. En tout cas, philosophes et théologiens avancent scientifiquement dans leurs domaines en avançant d'hypothèses en hypothèses avant de poser des affirmations.

Enfin, le débat est revenu sur la continuité animal-homme à partir de l'évolution. Les frontières communément admises entre le singe et l'homme ne sont pas si facile à distinguer. Des travaux sur les singes ont pu déceler une certaine intelligence ou une certaine conscience. En revanche nous parlons de l'homme avant l'apparition de l'art pariétal ou de la pratique des sépultures. Le langage est-il si discriminant? Au fond, l'évolution est sans doute le fruit à la fois de fonctions continues et de fonctions discontinues.

Neurosciences et Psychanalyse : le cas Thérèse d'Avila

Michel Simon¹

Neurosciences et Psychanalyse ont, sinon le même objet, un terrain commun : le cerveau/esprit et la vie psychique des humains, la pathologie, la maladie mentale, la folie, les handicaps et la souffrance. Elles ont aussi un horizon commun : la connaissance, l'explication ou l'interprétation, le soin, la thérapie, une meilleure aptitude, grâce à elles, à vivre l'extravagante aventure humaine. Entre les deux, ni arbitre ni otage, singulière, Thérèse d'Avila.

Pour les uns : une névrosée, hystérique, hallucinée ; pour les autres : une victime d'un dysfonctionnement électrique ou hormonal du cerveau, une épileptique. Pour d'autres encore : une croyante, sainte et « mystique ».

"Pour la première fois une personne – et c'est une femme – décrit avec une lucidité clinique sans précédent les états de dépersonnalisation dus aux rafales épileptiques et leur dépassement dans la foi et dans l'amour de l'Autre."

Aux diagnostics des neurologues sur « le cas Thérèse », aux lectures plurielles, aux diverses interprétations avancées par les psychanalystes, Thérèse répond par le récit de sa vie, de ses "crises", de ses "ravissements", de ses "visions", de ses "extases"; par la suractivité de ses fondations; par ses "fictions", par la "métapsychologie" qu'est son *Château intérieur*.

Bref, elle aggrave son cas, mais comme en se jouant et en déjouant diagnostics et interprétations, auxquels elle oppose tranquillement l'assurance de sa singularité et de son expérience quand elle affirme " *Par expérience, nous avons vu que cette pierre précieuse* [les « visions imaginaires » de la présence du Christ] *nous a délivré de plusieurs maladies qu'elle a la propriété de guérir* ". Y aurait-il aussi une possibilité de transformations et de

¹ Centre théologique de Meylan-Grenoble

métamorphoses par la capacité de faire sens de l'animal humain, de la chair désirante, parlante, pensante, aimante ?

Voyons cela ensemble.

Introduction

Pierre Magistretti et François Ansermet² ont réussi un tour de force en réunissant pour un colloque au Collège de France en 2008 neuroscientifiques et psychanalystes : deux disciplines entre lesquelles le dialogue est pour le moins difficile. Une des idées directrices qui émergent de ce colloque est que les deux disciplines se trouvent confrontées à "l'émergence de la singularité, à la création de l'unique et au devenir imprédictible du sujet³". Qu'en est-il lorsque le « sujet » Thérèse d'Avila devient « objet » de l'attention de ces deux disciplines ?

Que nous disent les neurosciences sur Thérèse?

Avant les neurologues contemporains

Charcot (1825-1893) neurologue officiant à la Salpêtrière⁴, a lu la *Vie* de Thérèse et devant le récit de son coma de quatre jours, il porte un diagnostic d'hystéro-épilepsie, à un moment où l'épilepsie n'était pas aussi connue qu'actuellement. Dans son livre de 1897 *La foi qui guérit*, publié dans la « bibliothèque diabolique», il renouvelle son diagnostic : "François d'Assise⁵ et sainte Thérèse étaient eux-mêmes des hystériques indéniables." Il assimile

-

² Pierre Magistretti et François Ansermet, *Neurosciences et psychanalyse*, *Une rencontre autour de la singularité*, Odile jacob, 2010. P. Magistretti est directeur du *Brain Mind Institute* de l'école polytechnique de Lausanne et du centre de neurosciences psychiatriques de l'Université de Lausanne. Il a été professeur associé au Collège de France en 2007-2008. François Ansermet est psychanalyste, professeur de pédopsychiatrie à l'Université de Genève et chef du service de psychiatrie d'enfants et d'adolescents des hôpitaux universitaires de Genève. Ils ont publié ensemble *A chacun son cerveau. Plasticité neuronale et inconscient.*

³ Op.cit. p.12.

⁴ cf. le célèbre tableau par Brouillet de Charcot présentant une hystérique à la Salpêtrière.

⁵ François en raison des stigmates.

Thérèse d'Avila à Sœur Jeanne des Anges, supérieure du couvent où s'est déroulée l'affaire des possédées de Loudun⁶ étudiée par Michel de Certeau.

Concernant Thérèse d'Avila, on connaît aussi le jugement de Pierre Janet⁷ (1859-1947) "La patronne des hystériques". Par contre, Joseph Breuer, qui a cosigné avec Freud les *Etudes sur l'hystérie*, (1895) refuse d'identifier hystérie et maladie mentale : il interroge : "« La patronne des hystériques », sainte Thérèse, ne fut-elle pas une femme géniale, ayant le sens pratique le plus développé ?⁸"

Freud, à Rome, s'est arrêté longuement devant le Moïse de Michel-Ange. Il écrira *Moïse et le monothéisme*. Jacques Lacan, lui, dans l'église Santa Maria della Vittoria s'arrêtera longuement devant la statue du Bernin représentant la *transverbération* de Thérèse. Il mettra cette représentation sur la couverture de son Séminaire *Encore*.

Thérèse a écrit, sur commande, le récit de sa vie, récit qui contient les deux scènes retenues l'une par les neurologues (un long coma de 4 jours où on l'a crue morte), l'autre, la célèbre vision de la transverbération, par les psychanalystes et particulièrement Lacan. Voilà donc Thérèse ballottée pour longtemps au gré des diagnostics des neurosciences et des interprétations des disciples de Freud. Thérèse d'Avila attend ses lecteurs au carrefour tridimensionnel où s'affrontent et se rencontrent neurosciences, psychanalyse et... expérience mystique. Un beau thème pour le Réseau Blaise Pascal réunissant les groupes francophones travaillant sur le thème « sciences et foi ».

Le « cas Thérèse » aujourd'hui

• Avec les neurologues ayant étudié l'épilepsie

Théophile Alajouanine (1890-1980)⁹, C et H. Gastaut ont étudié de manière approfondie l'épilepsie de Dostoïevski. Sur la base de leurs travaux,

⁶ Sœur Jeanne des Anges, née en 1615 –Urbain Grandier brûlé en 1634. Michel de Certeau, *La possession de Loudun*, folio, 2005.

⁷ Philosophe, psychologue et médecin.

⁸ Sigmund Freud et J.Breuer, *Etudes sur l'hystérie*, (1895) PUF, 1967, p. 187.

⁹ Théophile Alajouanine a occupé à la Salpêtrière la chaire de Clinique des maladies du système nerveux dont Charcot fut le premier titulaire.

Estraban Garcia-Albea en Espagne et Pierre Vercelletto¹⁰ en France ont revisité le « cas Thérèse »

• Le livre de Pierre Vercelletto (PV) neurologue, épileptologue est celui qui fait actuellement le point sur « la maladie de Thérèse », versant neurosciences. Michel Bonduelle, le préfacier de *Epilepsie et état mystique* écrit "Son analyse sera précieuse au théologien. Elle donnera à réfléchir au médecin. Elle donnera à l'honnête homme à méditer sur les mystères de l'âme, de l'esprit et du corps."

Le livre de PV est un essai d'observation clinique à partir des *écrits* de Thérèse et des *témoignages* recueillis lors des procès de béatification de Saragosse (1595), Villa Nueva de Jara (1596) et Avila (1610).

• Vercelletto défend la thèse que la « maladie de Thérèse » est de nature épileptique. Sur ce diagnostic il rapproche les « comas », les « ravissements » et les « extases » de Thérèse des « crises extatiques » de l'épilepsie temporale. Galien dès le 2ème siècle avait qualifié d' « aura » le début conscient de la crise épileptique. Vercelletto avec les épileptologues reprend ce terme pour désigner "tous les débuts conscients des crises qui se manifestent dans les zones cérébrales permettant la conscience, donc la fixation des souvenirs." Ces auras peuvent être accompagnés de peur et d'angoisse ou, plus rarement, de plaisir et de joie qui peut aller jusqu'à un état de béatitude avec une exaltation de la compréhension intellectuelle ; on les qualifie alors de « crises extatiques 11 ». L'existence de ces crises était mise en doute jusqu' aux travaux sur l'épilepsie de Dostoïevski (que Freud avait diagnostiqué à tort comme une névrose). Vercelletto, dans son livre, cite et résume les études sur Dostoïevski.

_

¹⁰ Pierre Vercelletto, Epilepsie et état mystique. La Maladie de Saint Thérèse d'Avila, Les éditions de la Bruyère, 2000.

¹¹ Vercelletto distingue *crise épileptique* ("manifestation clinique d'une hyperactivité anormale et brusque d'un groupe de neurones cérébraux" (13), *crise extatique* (modalité plaisante des crises épileptiques) (30) et *extase mystique* qui ne relève plus exclusivement de l'épileptologue. Comparant Dostoïevski et les poèmes de Jean de la Croix, Th. Alajouanine est frappé par la même difficulté à exprimer l'inexprimable et la similitude des modifications affectives: joie, béatitude, sentiment d'universelle harmonie (33).

Puis, il rassemble dans la littérature médicale actuelle (depuis 1950) quinze observations d' « **auras plaisantes** » avec "sensation d'intense bien être", voire d'extase dont le récit dépend de la personnalité du malade. Le départ de ces crises implique souvent le lobe temporal.

Après l'étude des témoignages sur la « maladie de Thérèse », recueillis lors des procès de béatification après sa mort, PV estime que l' "on peut retenir de tout cela des arguments en faveur d'une maladie épileptique compliquée à l'âge de 23 ans par un état de mal gravissime. (...) Après un état de mal de cette importance, l'ædème cérébral peut laisser des traces, notamment sur les lobes temporaux et modifier l'aspect clinique des crises. Une possible épilepsie temporale constituera un argument en faveur de la réalité de crises extatiques..."

En ce qui concerne les **paroles** entendues et les « **visions** » de Thérèse. PV estime que "la plupart ne peuvent entrer dans le cadre de l'épilepsie temporale". Thérèse elle-même distingue trois sortes de vision. Les visions corporelles, visions d'une scène mystique avec les veux du corps. Pour PV, "véritable perception sans objet", ce sont des hallucinations visuelles. Les visions imaginaires, (au sens précis de différentes des visions corporelles), selon Thérèse, sont vues « avec les yeux de l'âme ». Par exemple, lors d'une vision qu'elle raconte ainsi "Je vis Notre Seigneur dans Sa Sainte Humanité toute entière, tel qu'on le peint ressuscité. Il m'apparu dans une beauté et majesté incomparable..." elle ajoute en insistant "Je n'ai jamais contemplé cette vision, quoiqu'imaginaire, ni aucune autre, des yeux du corps, mais seulement des yeux de l'âme"12. Enfin, les visions intellectuelles sont sans images, c'est un "sentiment de présence ou perception intime" : Thérèse ne voit rien mais ne peut douter de la présence. Selon Jean Lhermitte¹³ pour Thérèse, visions imaginaires et visions intellectuelles sont très proches comme elle le dit elle-même. Vercelletto, lui, confie aux théologiens le soin d'en donner une explication!

¹² Thérèse d' Avila, Œuvres complètes, Vie,27, 4-5.

¹³ Les hallucinations, 1951.

Au final, PV s'efforce d'évaluer ce que son hypothèse d'une épilepsie temporale permet de mieux comprendre chez Thérèse. A son avis, "il est beaucoup plus facile d'admettre l'authenticité de l'expérience mystique de sainte Thérèse d'Avila si l'on extrait, pour les rattacher à sa maladie épileptique, les phénomènes extraordinaires qu'ont été les ravissements et certaines visions corporelles telles que la transverbération ou la vision de la colombe couverte d'écailles qui jetaient un vif éclat et dont elle entendait le bruissement des ailes pendant la durée d'un Ave Maria. 14 Vie, trad. P.Marcel Bouix, s.j. 1867, p. 509.

Le diagnostic d'épilepsie temporale permet de donner une explication :

- aux "évanouissements", aux « maux de cœur » [mal de Corazon , en fait l'expression des castillans de l'époque pour désigner le *morbo sacro*, "mal sacré" ou épilepsie] qui, apparus à l'adolescence, se sont ensuite renouvelés.
- à la grave crise à 23 ans (août 1538) dont elle a failli mourir et qui était à n'en pas douter un mal épileptique.
- aux séquelles de ce dernier qui se traduisirent par des accès de "perlesia" [Paralysie cérébrale ; maladie qui prive du sentiment et du mouvement], de tremblements, de troubles transitoires du langage, d'absences.
- Enfin aux « ravissements » qui selon PV, peuvent être interprétés comme des "crises extatiques" qui malgré leur rareté sont parfaitement connues dans l'épilepsie temporale.

Son diagnostic final : "Thérèse d'Avila, épileptique célèbre et mystique authentique qui a pu «avoir fait bon usage de sa maladie », comme le dit Alajouanine de Dostoïevski."

Il n'est pas inutile d'entendre l'éloge que fait Julia Kristeva du travail du docteur Vercelletto : " On ne saura jamais assez gré au docteur Vercelleto d'avoir, en précisant la valeur du terme « aura » en épileptologie, suggéré avec prudence qu'à la crise motrice, (caractérisée par des secousses cloniques et diverses paralysies des membres ou de l'élocution – ce dont Thérèse se plaint à maintes reprises) se relient des acquisitions neuropsychologiques « spécifiques à chaque personne » ; ses singularités sensorielles,

¹⁴ Op.cit. p. 77.

intellectuelles, voire métaphysiques, sans oublier ses tendances... à l'hypergraphie. Or, en matière de singularité, Thérèse n'a rien à envier à personne, le clinicien en convient. Et, en toute honnêteté, il souligne qu'il ne s'agit nullement de réduire la sainteté de Thérèse à son lobe temporal. Pas plus qu'il n'est possible de l'ignorer."

passe alors diagnostics portés d'autres aux par neurologues/psychiatres qui voient chez Thérèse un terrain hypersensible propice aux régressions, aux exaltations, voire à l'alexithymie¹⁵. JK ajoute à propos de ces états limites (borderline) que "si la science accroît de plus en plus sa connaissance (...) elle a toujours du mal à saisir comment un sujet s'en sert pour s'en sortir..."16.

Les interprétations psychanalytiques

Denis Vasse : L'Autre du désir et le Dieu de Thérèse

Denis Vasse (né en 1941) jésuite, est docteur en médecine et psychanalyste. Il a écrit près d'une quinzaine d'ouvrages, parmi eux celui qui nous intéresse tout particulièrement aujourd'hui, L'Autre du désir et le Dieu de la foi.17

1. - Psychanalyse et neurosciences

Au seuil de son ouvrage, DV nous avertit que les textes de Thérèse nous resteront étranges et fermés si nous les séparons de son expérience de foi. Si nous abordons Thérèse d'Avila "au seul empan de notre sagesse psychiâtriquement humaine, nous la classerons dans les productions

¹⁵ L'alexithvmie est un "état limite" (borderline) "l'incapacité de certains sujets de traduire en langage leurs surexcitations, ce qui entraîne diverses manifestations: dégoût, anorexie, parfois même des épilepsies atypiques." Pour JK Thérèse serait bien plutôt une hyperlexithymique verbalisant avec virtuosité les moindres frémissements de ses affects ou perceptions.

¹⁶ Julia Kristeva, *Thérèse mon amour*, p. 187

¹⁷ Denis Vasse, L'autre du désir et le dieu de la foi. Lire aujourd'hui Thérèse d'Avila, Seuil, 1991.

« hystériques » ou « psychotiques »... ¹⁸ "Pourtant, "A lire et à relire son œuvre, l'enseignement de ce docteur de l'Église éclaire la structure de l'appareil psychique de l'homme. Dans l'exploration de la connaissance de soi, Thérèse se laisse conduire jusqu'à ce point d'échappement ou d'ouverture qui est le point aveugle¹⁹ du désir."

2.- Psychanalyse et anthropologie

Pour ce jésuite et psychanalyste, proche de Lacan et de Françoise Dolto, l'être humain est aveugle et inconnu. Il ne sait pas qui il est : "l'homme n'est pas l'idée qu'il a de lui-même; Et Dieu n'est pas davantage la représentation qu'il s'en fait. Dans l'exaltation projective de ce que nous pensons, nous n'accédons pas à la vérité que nous cherchons."

Son premier ouvrage nous parlait du temps du désir. Désir dont il faut discerner la marque au cœur de l'homme, car il est chemin de vérité. "La marque du chemin de la vérité est bien le désir sans limite et sans retenue, infini. Ce désir ne saurait être limité par rien ni retenu par quelque chose: il n'a pas d'objet ou, plutôt, son Objet n'est pas un objet, il est l'Autre." Le désir, selon la formule connue de Lacan, est "désir de l'Autre". Pour que cela puisse apparaître, encore faut-il se défaire de la confusion et du mensonge qui hantent notre imaginaire et notre volonté et qui nous trompent sur l'objet du désir. En effet, "j'imagine que l'objet de ma volonté est identique à l'objet du désir. Évitant, par là, de m'offrir au travail du discernement qui les distingue, en croyant vouloir ce que je désire, je ne fais que désirer ce que je veux, et le rapport à l'Altérité qui fonde la demande et la prière dans le désir se dissocie dans la perplexité mortelle du doute. Si je m'obstine dans cette confusion en voulant en sortir par moi-même, comme l'on dit, me voilà livré à la préoccupation harassante de me construire une idole que j'adore et de la

-

¹⁸ Denis Vasse, op. cité, p. 20. Au moment où il écrit son livre, DV ignore la thèse de l'épilepsie temporale.

¹⁹ Ce point aveugle où les fibres de la rétine se rassemblent pour devenir le nerf optique, ou, inversement, les neurones venus du cerveau s'épanouissent en plage rétinienne, nous permet d'être des "voyants": c'est par lui - si l'on peut dire - que ce que nous voyons parle à notre esprit.

brûler ensuite puisqu'elle ne fait que me révéler l'impuissance de ma propre volonté à vouloir ce qui est désiré en vérité dans mon corps." "Ne sachant pas en vérité ce qu'il désire, l'homme fait l'expérience du vide quand il obtient ce qu'il imaginait vouloir en le prenant pour l'objet du désir.(...) Ce qui le fait vivre (...) c'est un Autre, impossible à objectiver dans une image. (...) L'être de l'homme est fondé dans le désir de l'Autre. (MS). (....) L'Autre du désir échappe à la compréhension de l'imaginaire. Il est hors de portée de la volonté propre qui ne peut l'atteindre. L'objet du désir est, à l'instar du sujet de la parole, inconscient. A ne vouloir que ce qu'il imagine consciemment, l'homme se méconnaît comme vivant. Il juge de ce qui le fait vivre selon ce qu'il pense. Il ne croit qu'en ce qu'il sait ou en ce qu'il voit. C'est-à-dire que ni il ne croit, ni il ne désire. Une telle "créance" est source d'erreur et, dans l'obstination du vouloir, elle est la marque assurée de l'aveuglement.".

L'anthropologie de Denis Vasse est une structure ouverte qui questionne sur l'objet et la vérité du désir. "Avec son élaboration théorique, Lacan nous introduit à la manière dont l'écoute des signifiants qui représentent le sujet ouvre le discours de l'homme à ce qui parle en lui, même et surtout à son insu. Non sans raison, il situe la découverte freudienne de l'inconscient dans le champ de la parole et du langage. Dire cela revient à dire que, lorsque l'homme s'ouvre à ce qui parle en lui, il prend corps en vérité"

3. - Le chemin de Thérèse

"L'Autre n'est pas Dieu mais ouvre et maintient ouverte la question de l'homme sur lui-même."

• Pour le disciple de Lacan qu'est Denis Vasse, « L'Autre n'est pas Dieu. Pourtant, sa nécessaire position dans l'ordre logique de la structure maintient ouverte la question de l'homme sur lui-même là où elle risque de se clore : dans la prison du Même indéfiniment dédoublé où il serait assigné à résidence forcée - jusqu'au suicide - par l'horreur d'une souffrance sans espérance. Là où un Autre Vivant ne pourrait plus prendre Corps. Cette

ouverture nécessairement maintenue autorise, dans l'ordre historique, la position de la question de Dieu.²⁰ »

• DV voit la clef de la « théologie » de Thérèse dans le chapitre XXII de sa Vie. "[Thérèse] conteste la médiation du savoir et lui oppose le corps, l'humanité, la connaissance expérimentale." "L'histoire de Jésus confirme que le don de Dieu passe par la vie du corps et non par l'idée que l'homme aurait de Dieu."

"Aux yeux de la croyante qu'est Thérèse, il n'y a que par, avec et en l'humanité du Christ que ce rapport du sujet et de l'Autre se donne comme le chemin, la vérité et la vie de l'homme. Mais l'Autre alors n'est plus l'Autre abstrait, gnostique ou psychanalytique, il est le Dieu de la foi : renonçant par amour à l'altérité absolue, il se fait le prochain. Il est la prévenance qui se révèle comme la vérité et la vie dans le chemin qui conduit à une rencontre d'où naît le sujet humain. Il se donne à contempler dans la métaphore du père qui, n'ayant pas retenu jalousement le Fils premier-né, prévient toujours déjà son retour et ne peut qu'accueillir tous ceux qui le suivent."

• Les impasses du désir. Sur son chemin, Thérèse a dû combattre obstinément ce qu'elle appelle le « point d'honneur », cette "manière que nous avons de mettre notre honneur [et notre bonheur] dans un « objet » auquel nous tenons plus qu'à tout, comme si de lui dépendait notre identité véritable et que, sans lui, la mort serait préférable à la vie. Cet honneur là — celui du monde — est détournement du désir de l'Autre qui est le désir de l'homme. Ce Tout-Autre, c'est Dieu pour Thérèse ; Et le mot de Dieu ne signifie en effet rien d'autre."

"Thérèse met Dieu, dans son discours, à l'endroit où Lacan met l'Autre, mais pour elle (...) Il est la Présence dans la nuit et dans l'aveuglement. Elle nous montre la voie où vivre vraiment : croire en la vie qui est donnée. Après avoir tenté de se confier au compromis de ses charmes et de son intelligence, elle se donne tout entière à Celui qui se donne tout entier à elle et qui est le chemin, la vérité et la vie d'une Parole faite chair. Il n'est pas mensonger d'y

²⁰ Denis Vasse, Le poids du réel, la souffrance. Seuil, 1983.

croire quand il est donné d'en discerner les effets dans la vie et jusque dans la mort."

- Le chemin de Thérèse conduit l'être humain au centre de lui-même : "Au centre de lui-même, l'homme passe un seuil : celui de son âme. Là, il ne subsiste dans l'être que de ce qu'un Autre l'y précède. Dans la perception des choses créées, il croyait dire la vérité en devenant conforme à l'exactitude de son savoir. Et voilà qu'en ce seuil où il est conduit par la connaissance des choses et des êtres jusqu'à la méconnaissance de lui-même, c'est la vérité qui parle. Le désir de connaître et de comprendre, de se connaître et de se comprendre, l'entraîne vers un détachement ou une dépossession de la maîtrise intellectuelle ou affective. Au plus intime de l'homme, le Tout-Autre l'attend. Sa Parole silencieuse l'y appelle, le nomme de son Nom et l'épouse."
- Le terme du voyage "La paix et la joie éprouvées dans le corps, près de la fontaine d'eau vive, marquent le terme du voyage, celui d'une rencontre qui commence sur le chemin de l'oraison : comme il en est lorsque celui ou celle vers qui l'on va nous laisse entendre qu'il vient à notre rencontre."

Julia Kristeva: Thérèse mon amour

Et puis vient Julia Kristeva. Sémioticienne, psychanalyste, écrivaine. Exploratrice du génie féminin : Hannah Arendt, Mélanie Klein, Colette... Après cette trilogie, *Thérèse mon amour*, récit...^{21,22}.

Psychanalyste, c'est avec ce référentiel qu'elle aborde la vie et les textes de Thérèse d'Avila. Comme le dit J. Hochmann: "l'âme, pour un psychanalyste, est une métaphore du corps sexué et agressif, une mosaïque de complexes et non le reflet monochromatique de la lumière divine"²³. Littéralement fascinée par son « sujet », c'est "au galop" que Julia Kristeva mène son enquête et sa réflexion dans un texte foisonnant de plus de 700 pages. On peut s'y perdre. J'essaie d'en tirer quelques fils conducteurs.

²² Dans cet ouvrage il faut notamment lire son « Je vous salue Thérèse » (p. 40-41)

_

²¹ Julia Kristeva, *Thérèse mon amour, récit*. Fayard, 2008.

²³ J. Hochmann *Le cerveau dans tous ses états*, Presses du CNRS, 1991, p. 89.

- 1 D'abord, pour elle, le « croire » n'est pas une attitude réservée aux croyants mais une donnée anthropologique fondamentale, préreligieuse, sans laquelle il n'y aurait pas de vie psychique ni de culture. La "psychisation" de l'être humain se noue autour de la *fonction paternelle*. C'est l'identification au père qui détache l'enfant de la seule satisfaction sensorielle vécue avec la mère. Cet investissement sur le père, ce désir d'unification avec lui est porté par un besoin de croire et de reconnaissance qui est à la source du langage, de la pensée et de la culture. Elle cite le *Psaume 116* « J'ai cru et j'ai parlé »... L'importance de la fonction paternelle, du « Nom du Père » (Lacan) pour l'avènement du sujet (séparation de la mère, accès au langage...) est la poutre maîtresse sur laquelle s'appuie JK pour parler de "père-vers" ou de "père-version" mais aussi de "mère-version" pour désigner la manière dont un sujet a traversé le défilé pulsionnel de l'Œdipe.
- 2 Or, pour le christianisme il existe **un Père idéal**²⁶. La foi est un amour pour et de ce père. Le syllogisme du croyant est « J'aime parce que je suis aimé/e, donc je suis », syllogisme qui va nourrir la vie mystique. "La Foi chrétienne est une confiance inébranlable en l'existence d'un Père Idéal, et un amour absolu pour ce Père aimant, qui serait tout simplement le fondement du sujet parlant, lequel n'est autre que le sujet de la parole amoureuse.²⁷"
- 3 Malheureusement, la foi chrétienne méconnaît l'Œdipe et la vie pulsionnelle découverte par Freud. Elle ne retient de l'amour qu'agapè en occultant éros. Impossible aujourd'hui, après Freud, de lire Thérèse et ses symptômes en ignorant les mouvements de la vie pulsionnelle chez les humains (les pulsions et le destin des pulsions).

²⁴ « Oser l'humanisme » - Parvis des Gentils - (Paris, mars 2011).

_

²⁵ Thérèse mon amour, p. 217s.

²⁶ Pas seulement pour le christianisme, à vrai dire, car le Père idéal est "le fantasme, car il est une représentation sexuée qui décolle de la sexualité. Il est « père », donc géniteur, mais « idéal » parce que défini par sa fonction symbolique. Un être de carrefour, en somme, entre désir et sens, entre passion et pensée." *Thérèse mon amour*, p. 210

p. 210.

²⁷ J.Kristeva, « La Passion selon Thérèse d'Avila » conférence prononcée au colloque «Vers les monothéismes», organisé par le Réseau de Recherche Psychanalyse et religions. Publication n° 96, 2006.

4 • Kristeva réintroduit alors sous la face solaire des extases de Thérèse, une face plus étrange. Elle fait remarquer à Thérèse "celui qui vous aime et que vous aimez est un fils et non moins un père souffrant, battu et mis à mort". A l'exemple de Freud qui a attiré l'attention sur le fantasme « un enfant est battu », Kristeva soutient qu'il y a au fondement de la foi chrétienne un fantasme du « père battu à mort ». A l'analyste, le fantasme du « père battu à mort » apparaît comme une nécessité logique dans la construction chrétienne du sujet du désir. Il déculpabilise l'amour incestuel du père et pour le père, par son déplacement sur la souffrance-passion comme voie obligée de la sublimation. Cette nécessité logique commence par déplacer *l'interdit* ou *l'abandon* sous la forme d'une *punition* de et par le père, vécue comme une souffrance passionnelle; avant d'autoriser enfin *l'amour* du et pour le père dans la « réconciliation » par l' « amour intellectuel infini » (Spinoza), par la sublimation. Un itinéraire où se mêle tour à tour masochisme, sadisme et jouissance. [cf. la père-version]

5 • Au sujet des extases, des ravissements de Thérèse et du récit qu'elle en fait, JK propose une interprétation originale qui a sa source dans l'œuvre du psychanalyste britannique Donald Winnicott. Ses extases seraient une régression jusqu'à ce que Winnicott appelle un « psyché-soma » ; un état de régression où l'individu perd ses contours identitaires et, au dessous du seuil de la conscience, devient un « psyché-soma ». Devant certains états régressifs de ses patients caractérisés par une perte des facultés jugeantes et par une grande acuité sensorielle et imaginaire, Winnicott à l'intuition que tous les sens et tous les organes participent à la perception de soi et du monde, comme si la psyché était corps et le corps psyché, donnant lieu à un touché, un senti, un imaginaire sensible. Kristeva estime que cela convient particulièrement bien aux récits de Thérèse décrivant la perte de ce qu'elle appelle les puissances supérieures. Elle prête aussi grande attention dans les textes de Thérèse à la métaphore de l'eau, élément vital, fiction, représentation sensible de la rencontre, du contact mutuel de Dieu et de sa créature. "Thérèse s'immerge, au-dessus des mots-signes, dans le « psyché-soma ». C'est par sa

²⁸ « Un père est battu à mort », *Thérèse mon amour*, pp.453-475.

fiction (mieux et autrement que par l'épilepsie) qu'elle se soustrait aux « puissances » (entendement, mémoire, imagination)²⁹". (...) Le soi « se perd », se « liquéfie », « délire » : ...le soi se perd et se retrouve, catastrophé et jubilant, entre deux eaux. L'eau, fiction absolue, inévitable, du toucher amoureux.

- 6 On peut alors approcher ce que sont les textes de Thérèse et tout particulièrement son *Château de l'âme* ou *Livre des demeures*. Plus qu'une métaphore, ils sont le récit d'une métamorphose. "A notre regard d'incrédules du troisième millénaire, l'expérience mystique se confond avec cette recomposition de l'être parlant par l'écriture métaphorique. Thérèse écrit une décomposition de son identité intellectuelle-physique-psychique dans et par le transfert amoureux avec l'Etre Tout Autre : Dieu, figure paternelle de nos rêves infantiles, insaisissable Époux de la Sulamite [du *Cantique des cantiques*.]³⁰"
- 7 JK ajoute un Post-scriptum à son étude sur Thérèse. Pendant toute la rédaction de son livre, elle a pensé constamment au livre de Denis Diderot *La Religieuse*. Ce livre est le fruit d'un pari entre joyeux drilles pour démontrer, dans un récit, à quel point la claustration religieuse est un refoulement féroce des instincts par "ces vœux qui heurtent la pente générale de la nature", une forme de « bestialité ». Or, voilà que Diderot en écrivant ce réquisitoire, se met à un moment à fondre en larmes et que l'ouvrage est resté inachevé comme l'atteste encore la conclusion bâclée. "Denis Diderot pleurait sur le manuscrit de sa *Religieuse* qu'il n'arrivait pas à finir, incapable de donner du sens à la vie de son héroïne après l'avoir sortie du couvent. "Pleurez-vous, interroge J.Kristeva, votre sœur cadette Marie-Angélique, morte folle à vingt-huit ans chez les ursulines et dont le souvenir ne vous quitte pas, puisque votre fille bien-aimée se prénommera aussi Marie-Angélique ? Pleurez-vous parce que vous êtes scandalisé [...] par l'obscurantisme religieux qui opprime la « pente

_

²⁹ J. Kristeva, « La Passion selon Thérèse d'Avila ».

³⁰ Thérèse mon amour, p.122.

³¹ Julia Kristeva « La séduction mystique », M, le supplément au journal "Le Monde" du 2 décembre 2010.

générale de la nature »? Ou bien vos pleurs vous surprennent-ils parce que vous n'êtes pas sans savoir que, chez la bête humaine – pour autant qu'elle est un être parlant – cette « pente générale de la nature » est déjà de toute façon « heurtée » par la capacité de faire sens. Une capacité elle aussi spécifiquement - donc naturellement – humaine de heurter la nature par le langage et la pensée ?³²² "Dans les plis subtils de la culture, la chair brûle le verbe et le verbe brûle la chair. (...) Le besoin de croire serait inopérable du désir, qui est désir de sens.³³³"

Je termine par le jugement de Michel de Certeau sur le *Château intérieur* de Thérèse d'Avila³⁴. Ce livre dont on a pu dire qu'il est "pour la mystique ce que la *Logique* d'Aristote est pour la philosophie traditionnelle."³⁵ M. de Certeau écrit : "Ce traité de l'âme, de l'oraison et du discours (ou itinéraire) mystique s'inscrit sans nul doute dans une longue tradition, socratique et spirituelle, du « connais-toi toi-même »³⁶, mais il la déplace dès le départ en la traduisant par deux autres questions : « Qui d'autre habite en toi ?» et « À qui parles-tu ? ». Une problématique de l'être et de la conscience est d'emblée déportée vers l'énonciation, c'est-à-dire vers une structure dialogale de l'altération - « Tu es l'autre de toi-même ». L'âme devient le lieu où cette *séparation d'avec soi* est le ressort d'une *hospitalité*, tour à tour « ascétique » et « mystique », qui *fait place* à l'autre. Et parce que cet « autre » est infini, l'âme est un espace infini où entrer et recevoir -, « les Indes de Dieu », disait, au XVIème siècle, [le poète] Francisco de Aldana.³⁷".

³² Thérèse mon amour, p. 673.

³³ Ibidem, 674 et 679.

³⁴ Michel de Certeau, La fable mystique, XVI^e – XVII ^e siècle, Gallimard, 1982.

³⁵ Melquiades ANDRES, La Teologia españiola en el siglo XVI, Madrid, BAC, 1976,t.1,p.166.

³⁶ Roberto Ricard, « Notes et matériaux pour l'étude du "socratisme chrétien chez sainte Thérèse et les spirituels espagnols », in Bulletin hispanique, t. 50, 1948.

 $^{^{37}}$ cité par Michel de Certeau, *La fable mystique*, op. cité, p. 267. Francisco de Aldana ($\approx 1537-1578$), était un militaire et l'un des plus grands poètes de la Renaissance espagnole.

Compte-rendu du carrefour "Neurosciences et psychanalyse"

Michel Dion

Thérèse d'Avila a marqué par son parcours spirituel. Agée d'une vingtaine d'année, elle subit un coma de quatre jours, si bien que ses proches la croient morte. Elle fait également l'expérience de la « transverbération » (elle est virtuellement transpercée par un dard d'or). Lors de ces crises extatiques, elle ressent plaisir, béatitude et sa compréhension intellectuelle est exaltée. Mais cela s'accompagne aussi d'évanouissements et de maux de cœurs. Les neurobiologistes et psychanalystes se sont intéressés à son cas sur la base de son propre témoignage. Le docteur Charcot pense qu'elle a fait un coma hystérique mais depuis que l'on connaît mieux les symptômes de l'épilepsie temporale, on pense plutôt qu'elle était atteinte de cette maladie.

Julia Kristeva, psychanalyste vient de lui consacrer un ouvrage important : « Thérèse mon amour » (2008). Pour elle, il est impossible de lire le récit de la vie de Thérèse sans le mettre en relation avec les pulsions humaines fondamentales (*Eros/Thanatos*); sans suivre à la trace leurs métamorphoses dans une existence humaine (jouissance, masochisme, sadisme même...); sans examiner comment la figure chrétienne de Dieu (le Dieu Père, la Passion/résurrection du Christ...) sollicite le désir humain.

Pour Denis Vasse, jésuite et psychanalyste, Thérèse conteste la médiation du seul savoir ; « ce qui la meut sur le chemin de la vérité, c'est le désir de l' Autre ».

Quoiqu'il en soit, ses crises extatiques ont pu avoir une influence sur sa vie mystique. Elle les a cependant bien gérées. Son témoignage indique son chemin, ses progrès vers la rencontre de l'Autre.

Parmi les participants au carrefour, les réactions sont diverses, mais globalement, son cas laisse un peu perplexe. Même si l'épilepsie temporale est plus subtile que l'hystérie, qu'est ce que cela nous apprend? Que fait-on avec cette expérience? Peut-être faut-il entendre les deux questions que nous pose son expérience mystique, au-delà du "Connais-toi toi-même" de Socrate : « Qui d'autre habite en toi ?» et « À qui parles-tu ? ».

ABONNEMENT ET COMMANDE D'ANCIENS NUMÉROS

Abonnement (deux numéros) : 18 €; abonnement de soutien 25 €.				
L'Association Foi et Culture Scientifique peut envoyer par courrier les N° parus :				
N° 6 à 32 : 7 €, le N° simple; 14 € le N° double ; N° 33 à 34 : 8,5 € ; N° 35 : 9 € ; N° 36-37 : 18 €.				
Les N° 1 à 34 sont téléchargeables à : evry.catholique.fr/IMG/pdf/AFCS_connaitre.pdf				
Si vous souhaitez soutenir l' Association Foi et Culture Scientifique , merci de joindre votre cotisation : \square membre associé : $5 \in$ ou \square membre adhérent : $26 \in$.				
BULLETIN DE COMMANDE				
☐ Abonnement (deux numéros) : 18 € ☐ Abonnement de soutien : 25 €				
Commande d'anciens numéros de « Connaître » :				
□ \mathbf{N}° nombre d'exemplaires : soit 7 \longleftrightarrow x = \longleftrightarrow				
□ N° 33 nombre d'exemplaires : soit $8,5 \in \times$ = €				
□ N° 34 nombre d'exemplaires : soit $8,5 \in \times$ = €				
□ N° 35 nombre d'exemplaires : soit 9 $\epsilon \times$ = ϵ				
Cotisation (facultative) : \square membre associé : $5 \in$; \square membre adhérent : $26 \in$.				
Date: / / 2011 Somme totale€				
Nom Prénom				

Je joins mon règlement de € (par chèque bancaire ou postal) à l'ordre de " Association Foi et Culture Scientifique "

Adresse:

Code postal: Ville:

Courrier à adresser à : « Connaître » 13, Rue Amodru 91190 GIF/Yvette

Pour recevoir les informations sur la vie de notre association et les comptes rendus des réunions, adresser un courriel à : 91afcs@orange.fr

Site internet: evry.catholique.fr/Foi-et-Culture-Scientifique

Pays:....

CONNAÎTRE

Cahiers de l'Association Foi et Culture Scientifique Réseau Blaise Pascal

SOMMAIRE

N°36-37, décembre 2011

L'humanité de l'homme bouleversée par les neurosciences ?

Colloque du Réseau Blaise Pascal, 2 et 3 avril 2011, Sainte-Foy-lès-Lyon

Éditorial	F	Philippe Deterre	4
Conférenc	es plénières		
La naissance des neurosciences sociale.	s I	Marc Jeannerod	6
Neurosciences, déterminisme et libre ar	bitre	Bernard Feltz	32
Neurosciences et théologie de la foi-con	nfiance Alex	andre Ganoczy	48
Neurosciences + éthique = neuroéthiqu	e Be	rnard Baertschi	62
Table ronde	Philippe Deterre (Modérateur)		
Deux carrefor	urs thématiques		
Neurosciences et âme humaine	Jean-M	lichel Maldamé	110
Compte-rendu du carrefour Brice de Malherbe			123
Neurosciences et psychanalyse : le cas	Thérèse d'Avila	Michel Simon	126
Compte-rendu du carrefour		Michel Dion	141
Abonnements, anciens numéros			143