

**BULLETIN N° 95**  
**ACADÉMIE EUROPÉENNE INTERDISCIPLINAIRE**  
**DES SCIENCES**



Séance du Mardi 10 mai 2005

**Présentation du Congrès « Physique et Conscience »**  
*et discussion sur le congrès et les suites que nous*  
*pouvons lui donner*

|

Prochaine séance : le Mardi 14 juin 2005 :

*Accueil des Prs. Daniel Andler (ENS),*  
*Michel Bitbol (CREA),*  
*et Jean-Paul Bacqiast (Automates Intelligents).*  
*Discussion sur les thèmes du congrès.*



**ACADEMIE EUROPEENNE INTERDISCIPLINAIRE DES SCIENCES**  
**MAISON DES SCIENCES DE L'HOMME**

**PRESIDENT** : Gilbert BELAUBRE  
**SECRETAIRE GENERAL** : Michel GONDRAN  
**SECRETAIRE GENERAL ADJOINT** : Irène HERPE-LITWIN  
**TRESORIER GENERAL** : Bruno BLONDEL.  
**CONSEILERS SCIENTIFIQUES** :  
**SCIENCES DE LA MATIERE** : Pr. Gilles COHEN-TANNOUDJI.  
**SCIENCES DE LA VIE ET BIOTECHNOLOGIES** : Pr. François BEGON  
**PRESIDENT DE LA SECTION DE NICE** : Doyen RENE DARS.

**PREDIDENT FONDATEUR**  
**DOCTEUR** Lucien LEVY (†)  
**SECRETAIRE GENERAL D'HONNEUR**  
 Pr. P. LIACOPOULOS  
**TRESORIER GENERAL D'HONNEUR**  
 Pr. H. NARGEOLET

mai 2005

**N°95**

**TABLE DES MATIERES**

- P. 5. Compte-rendu de la séance du 10 mai 2005.
- P. 7. Compte-rendu de la section Nice Côte d'Azur du 21 avril 2005
- P.11. Documents

Les documents donnés ci-dessous constituent des approches de thèmes touchés par le congrès « Physique et conscience ». Dans cette livraison, nous avons donné la parole aux thèmes cognitivistes et épistémologiques (qui tendent largement à se confondre).

L'abondance de ces documents nous a privés de vous adresser des commentaires de Gilles Cohen-Tannoudji sur les sujets abordés par M. Crichton (Cf. N° 93 et 94), ainsi qu'un article de notre collègue le Recteur René Blanchet à propos des publications scientifiques.

Ces articles paraîtront dans le prochain bulletin.

Prochaine séance : Mardi 14 Juin 2005  
***Accueil des Prs. Daniel Andler (ENS),  
 Michek Bitbol ( CREA ),  
 et Jean-Paul Bacquiast ( Automates Intelligents ).  
 Discussion sur les thèmes du congrès.***



**ACADEMIE EUROPEENNE INTERDISCIPLINAIRE DES SCIENCES**  
Maison des Sciences de l'Homme, Paris.

*Séance du*  
**Mardi 10 mai 2005**

**Maison des Sciences de l'Homme, salle 215, à 18 h.**

La séance est ouverte à 18 h. 00 sous la Présidence de Gilbert BELAUBRE et en la présence de nos collègues Jacques BARBIER, Michel BERREBY, Noëlle CAGNARD, Alain CARDON, Gilles COHEN-TANNOUDJI, Michel GONDRAN, Irène HERPE-LITWIN, Marie-Louise LABAT, Pierre MARCHAIS, Jean POIRIER, Alain STAHL. Etaient excusés nos collègues Bruno BLONDEL, Françoise DUTHEIL, Henri GULAND , Pierre SIMON.

La séance est exclusivement consacrée à la présentation du Congrès « PHYSIQUE et CONSCIENCE » et discussion sur le congrès et les suites que nous pouvons lui donner

Notre Collègue Gilles COHEN-TANNOUDJI nous donne tout d'abord un certains nombre d'avancées relatives à la préparation du congrès :

1) Financièrement

- Attribution gratuite des amphithéâtres Gay-Lussac le vendredi et Poincaré le samedi de la rue Descartes à l'ancienne Ecole polytechnique comportant chacun 250 places. Cette offre, que nous devons au Ministère, implique la gratuité pour les participants.
- L'obtention promise de subventions de la part de la Mairie de Paris et du CEA (NEUROSPIN) nous permettra d'assurer les dépenses du congrès.
- Contact avec Ministère de la Recherche ,CNRS et la Région pour nous aider

2) Intervenants :

Notre collègue Pierre MARCHAIS , qui interviendra, nous a mis en contact avec les Prs J.-B. GRIZE (Université de Neuchâtel) et C. DEBRUE ( E.N.S. ). Par ailleurs , ont été contactés, et ont donné leur accord de principe, les Prs. Daniel ANDLER, Michel CABANAC, Alain CARDON, Stanislas DEHAENE, Yves FREGNAC, Denis LE BIHAN, Jean PETITOT, Alain PROCHIANTZ et Wolf SINGER..

3) Parrainages

Le congrès « PHYSIQUE et CONSCIENCE » a obtenu le parrainage de l'Académie des Sciences, le Label de l'Année Mondiale de la Physique, le partenariat avec Physique et Interrogations fondamentales (Société Française de Physique) et la BNF.

Un comité de parrainage est en cours de constitution, comprenant la Revue « Science et Avenir »( hors-séries ), les centres Georges CANGUILHEM et Ferdinand GONSETH, le Laboratoire PENSEE DES

SCIENCES de l'ENS, le CEA (NEUROSPIN)...Des contacts ont également été pris avec le CNRS et avec des personnalités comme Dominique LECOURT et Etienne KLEIN.

4) Création du Comité Scientifique : il comprendra les membres de l'Académie impliqués dans le congrès, les intervenants et des personnalités du monde scientifique. Les derniers contacts sont en cours.

5) Autres contacts programmés :

- Gilles BLOCH ancien Directeur Adjoint des Sciences du Vivant(CEA),  
Directeur de l'Agence Nationale de la Recherche.
- Michel BITBOL en tant que philosophe spécialiste de la mécanique quantique  
et membre du CREA.
- Dominique LAMBERT analyste de l'efficacité des mathématiques en physique
- Michel CROZON , Directeur de Recherche au CNRS, responsable de l'Année Mondiale de la  
Physique pour la Région Ile de France,
- Jean Jacques KUPIEC de l'Ecole Normale Supérieure, et son collègue Bertrand Laforêt.
- André SYROTA , Directeur des Sciences du vivant au CEA
- E. EMERY, co-fondateur du Centre Ferdinand Gonseth.

Notre collègue G. COHEN-TANNOUJJI aborde le problème de l'ORGANISATION DES SESSIONS qui devraient comporter selon lui un ou deux exposés relativement longs et quelques communications assez brèves.

Parmi les questions en suspens : il faut avoir confirmation des contacts pris avec « PHYSIQUE ET INTERROGATIONS FONDAMENTALES », et du soutien du CEA (Etienne KLEIN).

Au niveau de la communication il faudrait prévoir une affiche, un numéro hors série de la revue « Science et Avenir », la création d'un site « Web » et dès maintenant se soucier de la publication des actes du congrès. Pierre Gilles de GENNES et Bertrand DUPLANTIER seraient également intéressés par notre projet.

Le dernier point abordé a été celui des visées de notre congrès :

- Mise en place d'un Observatoire sur Physique et Conscience mettant en réseau des laboratoires de biologie, de neurosciences, de physique, de sciences cognitives .
- Développement d'une réflexion épistémologique poussée, de modélisations en intelligence artificielle, en logique , en robotique...

Ceci permettrait progressivement l'utilisation de la méthodologie rigoureuse des sciences physiques dans le domaine des sciences humaines.

Après ce riche exposé, la séance a été levée à 20heures.

Bien amicalement à vous.

Irène HERPE-LITWIN.

# Compte-rendu de la Section Nice-Côte d'Azur :

*Séance du 21 avril 2005*  
(81<sup>ème</sup> séance)

Le savoir est le seul bien qui s'accroisse à le partager. Comprendre est bien sans limite qui apporte une joie parfaite. Baruch SPINOZA (1632-1677)

## Présents :

Jean Aubouin, René Blanchet, Sonia Chakhoff, Pierre Couillet, Patrice Crossa-Raynaud, Guy Darcourt, René Dars, Jean-Pierre Delmont, Jean-Paul Goux, Yves Ignazy, Gérard Iooss, Jacques Wolgensinger.

## Excusés :

Alain Bernard, Emile Girard, Thierry Gontier, Jean Jaubert, Michel Lazdunski, Jean-François Mattéi, Maurice Papo.

## 1- Approbation du compte-rendu de la 80<sup>ème</sup> séance.

**Le compte-rendu est approuvé à l'unanimité des présents.**

## 2- Le mois écoulé.

### Pierre Couillet

- M. Lorenz Kramer collaborateur de Pierre Couillet est décédé brusquement de maladie. Nous regrettons tous sa disparition, notamment parce qu'il était l'animateur du projet Aspen – Mercantour.

- Dans le cadre de l'Année de la physique, il est prévu :

- Deux conférences grand public en mai au CUM sur la physique quantique les 20 et 27 mai 2005 à 18 heures, en liaison avec l'Université et quatre en octobre sur la relativité et l'astrophysique autour du projet ICRANET du professeur Rufini.

- un groupe de chercheurs de Gif sur Yvette décide de se délocaliser pour venir à Nice travailler avec ICRANET, qui pourrait probablement être installé dans la villa Ratti, avenue des Arènes à Cimiez.

René Blanchet accompagnera M. Monteil, Directeur de l'Enseignement supérieur, avec Monsieur le Président Marouani lors de sa visite à Nice. Le programme prévu lui permettra d'avoir un bon contact avec l'Université de Nice qu'il visite pour la première fois.

Il donnera une conférence à Valrose à 16 heures le vendredi 22 sur l'enseignement supérieur et la recherche dans notre pays.

Jacques Wolgensinger : l'*Express* a publié récemment la prise de position de la Présidente de la Société des agrégés de l'Université qui rejoint celle qu'avait prise l'association « Le Nénuphar » dès 1990 contre la réforme de l'orthographe.

Jean Aubouin : outre la défense de l'orthographe, il faudrait veiller à la défense de la prononciation qui est également contaminée par l'anglais.

Les Mexicains qui sont, comme nous, à la frontière du monde anglo-saxon, y veillent plus que nous.

Notons aussi le danger des moteurs de recherche comme Google qui sont uniquement en anglais. C'est ainsi qu'un géophysicien des anomalies magnétiques du plancher océanique nommé en Amérique « Brunies » n'est autre que le Français Brunhes.

Jacques Wolgensinger fait aussi remarquer que le vocabulaire des jeunes se rétrécit de plus en plus ; ils ne comprennent plus des quantités de mots même courants.

Le traité constitutionnel européen n'aborde nulle part le problème des langues et on risque de voir l'anglais devenir la langue européenne de communication.

Actuellement, à Nice, signale Pierre Couillet, les cours de master de géophysique qui accueillent plusieurs étrangers, sont délivrés en anglais.

Yves Ignazi signale la création à Sophia Antipolis d'un studio d'enregistrement et de diffusion doté des équipements les plus modernes.

Au CNAM a été créée une mission chargée de se préoccuper de la conservation de l'immatériel, ce qui est un problème majeur pour les années à venir compte tenu du renouvellement très rapide des instruments de lecture.

3- 6<sup>ème</sup> colloque : les peurs.

Plutôt que risques et précautions, nous tenterons d'analyser les peurs. Y a-t-il des psychoses collectives ?

\*\*\*\*\*

Prochaine réunion  
le jeudi 19 mai 2005 à 17 heures  
au siège  
Palais Marie-Christine

20, rue de France 06000 NICE
---------------------------------



## *Documents*

- P. 13 : Daniel Andler  
Les neurosciences cognitives : une nouvelle « nouvelle science de l'esprit » ?
- P. 29 : Michel Bitbol.  
Des « phénomènes » de Kant à la théorie quantique de l'information.  
Le rôle de la philosophie des sciences
- P. 37 : Jean-Michel MALDAME/  
Mathématiques, science, culture et philosophie
- P. 47 : Shahid Rahman :  
Sujet épistémique et sujet cognitif



## Les neurosciences cognitives: une nouvelle « nouvelle science de l'esprit »?

Daniel Andler

UFR de philosophie et sociologie, Université de Paris-Sorbonne (Paris IV)

Département d'études cognitives, Ecole normale supérieure

### Résumé (français)

Une question controversée est de savoir si, grâce en particulier aux techniques d'imagerie cérébrale fonctionnelle, une nouvelle science du cerveau occupe désormais, sous le nom de « neurosciences cognitives », le cœur des sciences cognitives. Quels sont les fondements de cette nouvelle spécialité? Comment s'articuleront, demain, les neurosciences cognitives et la psychologie? On replacera les techniques d'imagerie dans le contexte théorique des neurosciences fondamentales, et on montrera que les neurosciences cognitives en sont tributaires, comme elles le sont de l'ensemble des sciences cognitives, au sein desquelles elles constituent aujourd'hui un programme de recherche parmi d'autres. S'il venait à constituer, dans un avenir lointain, une science achevée des fonctionnalités du cerveau, il ne remplacerait sans doute pas la psychologie cognitive, appuyée sur les autres sciences de l'esprit, mais entrerait avec elle dans un rapport de complémentarité en un sens fort, très éloigné de toute forme de réduction.

### Abstract (English)

*It is a matter of considerable controversy whether cognitive neuroscience, thanks in large part to functional neuroimaging techniques, is in the process of becoming a new science of the brain and moving into the heart of cognitive science. What are the foundations of this new field? How will neuroscience and cognitive science coexist in the future? The paper will attempt to situate neuroimaging in the theoretical framework of fundamental neuroscience, and will show the extent to which cognitive neuroscience depends on it, as it depends on the rest of cognitive science, within which it stands as one research program among several. Should it lead, in a distant future, to a completed science of the brain's functionalities, such a science would likely not replace cognitive psychology and allied disciplines. Instead, I envisage a form of strong complementarity between the two branches, exclusive of any form of reduction.*

*Mots-clés* : neurosciences cognitives, neuroimagerie, sciences cognitives, philosophie des sciences

8 février 2005, rév. 14 février 2005

## Les neurosciences cognitives: une nouvelle « nouvelle science de l'esprit »?

Daniel Andler

Université de Paris-Sorbonne (Paris IV) & Ecole normale supérieure

Le rôle des neurosciences dans l'étude scientifique de l'esprit s'est prodigieusement accru au cours des dernières années, tout particulièrement depuis que les méthodes d'imagerie cérébrale se sont généralisées. L'étiquette "neurosciences cognitives", apparue récemment, semble témoigner d'une reconfiguration des sciences cognitives sous l'égide des neurosciences. S'agit-il d'une seconde "révolution cognitive"? Les acquis de la première sont-ils mis en question? Peut-on s'en remettre désormais aux neurosciences pour ériger enfin une science progressive de l'esprit, sous la forme d'une science du cerveau? <sup>1</sup>

### 1. Position du problème

**Ces questions sont déterminantes pour l'avenir des sciences de la cognition, et se posent dans un contexte de fortes tensions qui ne sont pas seulement intellectuelles, mais également institutionnelles. Une discussion sereine n'est pas toujours possible. Une autre source de confusion vient de ce que les choses se présentent différemment selon qu'on se place à l'extérieur ou à l'intérieur du milieu professionnel des sciences cognitives, et il est difficile d'en parler d'une manière qui soit recevable par les deux catégories de lecteurs.**

**Le lecteur extérieur cherche à identifier une discipline, un programme de recherche, une thèse, une méthode qui soient caractéristiques des sciences cognitives; il n'a ni les ressources ni le désir de mener sa propre enquête, et veut seulement savoir à quel représentant autorisé il doit s'adresser. C'est auprès de ce représentant qu'il va s'enquérir sur les fins et sur les moyens de ce qu'il considère comme un programme de recherche parmi d'autres, et qu'il va juger dans quelle mesure les sciences cognitives, ainsi représentées, apportent ou ont des chances raisonnables d'apporter des réponses à telle ou telle question qu'il juge essentielle concernant l'esprit.**

**Pour le lecteur de l'intérieur, la seconde interrogation n'est pas d'actualité: revendiquer, ou accepter l'étiquette "sciences cognitives", c'est faire le pari que des questions centrales et urgentes sur l'esprit sont à la portée de ces sciences, et qu'il serait vain, au stade présent, de mettre en question leur pertinence<sup>2</sup>. En revanche, il est partie prenante d'une perpétuelle querelle interne sur la bonne stratégie et sur le rôle des disciplines participantes.**

**Aujourd'hui, un témoin arrivé depuis peu sur la scène, et renseigné par les journaux, les revues savantes, les manifestations scientifiques, n'aurait guère de doute sur la question du représentant: les sciences cognitives ne sont désormais rien d'autre que les sciences du cerveau (les neurosciences), ou peut-être une branche de cette spécialité, celle qui porte le nom de "neurosciences cognitives", et la méthode princeps de cette discipline est l'imagerie cérébrale fonctionnelle.**

**Pour un acteur des sciences cognitives, cette évidence n'en est pas une. Non seulement il ne va pas de soi que les neurosciences cognitives sont le cœur du domaine, mais le sens et la portée de cette présentation font question. A supposer qu'on voudra lui donner son assentiment, ce ne pourra aucunement l'être en vertu d'une définition ou d'un simple constat, mais seulement à titre d'un bilan, terme d'une analyse complexe de l'évolution du domaine. Et l'affichage du bilan sous cette forme**

<sup>1</sup> Je remercie pour leurs remarques critiques sur une première version de cet article Anne Fagot-Largeault, Bernard Pachoud, Arnaud Plagnol, Jérôme Sackur et Marta Spranzi ; et pour leurs réactions les participants aux séminaires de Jacques Glowinski au Collège de France, et de Dag Westerstaal au département de philosophie de l'Université de Göteborg.

<sup>2</sup> On peut être plus radical, et considérer que par définition toute question sur l'esprit susceptible d'une formulation scientifique relève des sciences cognitives. Cela laisse ouverte la possibilité de questions importantes, voire essentielles, concernant l'esprit, auxquelles aucune réponse *scientifique* ne semble pouvoir être apportée dans un avenir proche: Chomsky, l'un des pères de la "révolution cognitive", défend cette opinion, d'autres la jugent incohérente, obscurantiste ou irresponsable.

ramassée ne manquera pas d'avoir des conséquences intellectuelles et institutionnelles, tant au sein du domaine que vis-à-vis de l'extérieur.

Après avoir averti le lecteur de leur importance, on n'insistera pas sur les aspects sociologiques du problème. Et on ne traitera (évidemment) pas la totalité des questions épistémologiques que soulèvent les sciences cognitives. Pour pouvoir formuler celles que l'on abordera, et les réponses qui seront proposées, il faut s'entendre sur les termes.

#### Mise au point terminologique

Les *sciences cognitives* désignent l'ensemble des programmes de recherche consacrés à l'élaboration d'une théorie scientifique de l'esprit humain, et qui souscrivent à une sorte de "cahier des charges" qu'il serait fastidieux de rappeler ici<sup>3</sup>, mais dont il est important de retenir qu'il est essentiellement méthodologique, et que non seulement il autorise, mais qu'il favorise la formulation et le déploiement de doctrines et de théories très différentes. C'est ce pluralisme qui donne leur sens aux désaccords dont il va être question, et qui sont tout à fait comparables aux controverses courantes dans toutes les disciplines scientifiques vivantes.

Les sciences cognitives ont une double structure: d'une part elles développent des programmes de recherche pluri- ou interdisciplinaires sur les différentes fonctions mentales; d'autre part elles s'appuient sur certaines branches de disciplines fondamentales. En particulier, les neurosciences y jouent un rôle important. Par ce terme nous entendrons, conformément à l'usage, l'ensemble des branches de la biologie portant sur le système nerveux, quel que soit le niveau d'analyse et la classe de phénomènes qu'elles privilégient: la neurophysiologie, la neuroanatomie, la neurobiologie cellulaire et moléculaire en sont des parties essentielles, à côté de la neuropsychologie, de la neuropharmacologie et des neurosciences appelées parfois "intégratives"<sup>4</sup>. Il peut alors sembler assez naturel de désigner par "neurosciences cognitives" la partie des neurosciences qui est impliquée dans les programmes de recherche des sciences cognitives. Ce choix est néanmoins très problématique et risquerait d'embrouiller des questions que cet article voudrait contribuer à clarifier. En effet, cette expression est associée aujourd'hui à une certaine pratique scientifique, relevant autant de la psychologie que des neurosciences, et appuyée sur le recours massif à la neuroimagerie fonctionnelle. Il serait maladroit d'introduire dans la terminologie même l'hypothèse que c'est cette pratique particulière qui constitue, pour le dire rapidement, l'intersection des neurosciences et des sciences cognitives, alors que cette hypothèse mérite justement d'être examinée<sup>5</sup>. De plus, pour ce qui est de la place des neurosciences computationnelles, un autre label récent, cette décision créerait une difficulté inutile. Convenons donc de ne pas préjuger de la question de la division des tâches, au sein des neurosciences, entre celles de ses branches qui sont pertinentes pour les sciences cognitives et celles qui ne le sont pas. Et convenons également de considérer les méthodes d'imagerie, d'une part, les techniques de modélisation issues des mathématiques, de la physique, de l'informatique, d'autre part, comme autant d'outils à la libre disposition de toutes les branches des neurosciences et des sciences cognitives.

#### Questions à prendre, questions à laisser

<sup>3</sup> V. par exemple D. Andler, dir., *Introduction aux sciences cognitives*, Paris: Gallimard, coll. Folio, nouvelle éd., 2004; D. Andler et al., *Philosophie des sciences*, Paris: Gallimard, coll. Folio, 2003, vol. 1, chap. 3; O. Houdé et al., *Vocabulaire des sciences cognitives*, Paris: PUF, 1996.

<sup>4</sup> Il n'existe pas aujourd'hui de découpage des neurosciences qui fasse l'unanimité; l'énumération proposée ne l'est qu'à titre indicatif.

<sup>5</sup> On rencontre un problème analogue en linguistique: le label "linguistique cognitive" désigne un programme de recherche particulier, ou plutôt une famille de programmes liés par des affinités théoriques et institutionnelles. Cette orientation, au sein de la linguistique, revendique une pertinence élevée, voire exclusive, pour les sciences cognitives. Une position parfaitement respectable, mais pour la défendre efficacement, comme pour la contester, il est crucial de ne pas l'introduire par stipulation dans la terminologie même, en décrétant que la linguistique cognitive est l'"intersection" de la linguistique et des sciences cognitives. On peut déplorer ce qui apparaît peut-être, vu de l'extérieur, comme des distinguos byzantins, mais ils sont absolument indispensables.

Ce petit « nettoyage de la situation verbale » (Valéry) étant fait, il devient relativement facile de formuler les questions qui se posent dans le présent contexte, et de dire celles qui seront abordées et celles qui seront laissées de côté:

**(Q1) Les sciences cognitives posent-elles des questions importantes relativement à l'esprit (les processus, états, phénomènes mentaux)?**

**(Q2) Les sciences cognitives ont-elles les moyens de leurs ambitions, possèdent-elles des méthodes éprouvées, sont-elles progressives?**

**(Q3) En particulier, la psychologie expérimentale (prise au sens le plus large) s'appuie-t-elle sur une méthodologie éprouvée, antérieure à l'émergence des neurosciences cognitives et de la neuroimagerie?**

**(Q4) Les sciences cognitives ont-elles vocation, à terme, d'absorber la totalité des enquêtes systématiques impliquant l'esprit, la conscience, la pensée?**

Ces quatre questions ne seront pas discutées. Nous postulerons, pour les besoins (urgents) de la discussion, une réponse positive aux trois premières, et écarterons résolument la quatrième. Les questions qui seront examinées, directement ou indirectement, sont les suivantes:

**(Q5) Quelle est l'importance de la neuroimagerie fonctionnelle pour les sciences cognitives aujourd'hui?**

**(Q6) Quelle que puisse être son importance relative, en vertu de quoi modifie-t-elle la situation?**

**(Q7) Les neurosciences occupent-elles désormais le centre des sciences cognitives, et si c'est le cas, est-ce essentiellement du fait de l'imagerie?**

**(Q8) Dans l'hypothèse où la neuroimagerie, associée éventuellement à d'autres progrès techniques et conceptuels des neurosciences, conduirait à une science développée des fonctionnalités du cerveau, quelle forme prendraient les sciences cognitives? Se réduiraient-elles alors aux neurosciences?**

En posant quatre questions distinctes, quoique manifestement liées, là où certains ne voient qu'une unique interrogation: **Oui ou non, les neurosciences cognitives sont-elles le présent, ou l'avenir imminent, des sciences cognitives?**, on se ménage la possibilité d'une réponse nuancée, qui ne soit pas pour autant une réponse de Normand. Ce qu'on entend surtout aujourd'hui, ce sont des réponses en tout ou rien: pour les uns, la neuroimagerie est une découverte fondamentale qui bouleverse dès à présent les sciences cognitives et les rabat en réalité sur les sciences du cerveau; pour les autres, la neuroimagerie est une technique coûteuse, sans portée théorique, dont le principal effet est de retarder ou d'entraver le progrès dans le domaine, en monopolisant les ressources et l'attention, et en en donnant une image ultra-réductionniste peu propice au dialogue avec les autres disciplines et courants de pensée, tant au sein des sciences cognitives qu'à l'extérieur.

Résumée en quelques lignes, la position que je défendrai dans le présent article est la suivante. Non, les neurosciences ne bouleversent pas la situation dans les sciences cognitives. Non, elles n'ont pas les moyens à elles seules d'édifier une science de l'esprit. Elles changent bel et bien la donne, mais ce n'est pas principalement du fait qu'elles apporteraient, enfin, des données empiriques suffisamment nombreuses et détaillées pour nous dispenser de recourir aux spéculations incertaines dont nous devons nous contenter jusqu'à l'arrivée de l'imagerie. Bien entendu, et il ne sera pas nécessaire de le rappeler constamment, les neurosciences apportent des informations importantes, non seulement pour les sciences cognitives théoriques, mais aussi pour la neurologie, peut-être pour la psychiatrie, et la récolte ne fait que commencer. Mais si l'on peut imaginer aujourd'hui que les neurosciences pourraient dans un avenir prévisible modifier la situation en profondeur, c'est davantage en raison des questions nouvelles qu'elles commencent à poser que des réponses qu'elles apportent à certaines questions déjà formulées par les psychologues, les philosophes ou les linguistes. On en verra des exemples, et on s'interrogera pour conclure sur ce que pourrait être la situation d'une science de l'"esprit-cerveau" lorsque l'orientation neurocognitive aura eu le temps de produire son plein effet.

## 2. Un retour en arrière

Les prémices de ce qui ne s'appelait pas encore les sciences cognitives apparurent il y a un demi-siècle. On put bientôt parler d'une "nouvelle science de l'esprit", fût-ce pour s'interroger sur sa réalité, sur sa nouveauté ou sur son unité. Il s'agissait à première vue d'un ensemble de programmes de recherche rangés sous différentes étiquettes disciplinaires. Leur convergence devait résulter de l'adhésion à certaines thèses, de l'adoption d'un certain vocabulaire théorique, d'une méthodologie particulière, d'une pratique interdisciplinaire. Une histoire bien connue, du moins dans ses grandes lignes, sur laquelle on ne revient que pour attirer l'attention sur la place occupée initialement par deux des composantes de ce mouvement. La première était la nouvellement nommée "intelligence artificielle" (IA), la seconde ce qui ne s'appelait pas encore les neurosciences. L'IA occupait le centre de la scène, rassemblant en elle l'essentiel du projet, sa philosophie et sa méthode; c'est du moins ainsi qu'elle-même concevait son rôle (elle était, selon l'un de ses pères fondateurs, l'"interdiscipline"), et, que ce soit par conviction ou par calcul, on ne songeait généralement pas, dans les milieux de la "nouvelle science", à lui contester cette prééminence. Les arts et sciences du système nerveux occupaient à l'inverse une position marginale. Cela ne tenait pas à leur moindre mérite, mesuré à l'aune de l'ancienneté ou des résultats, théoriques et pratiques (à ce compte, l'IA ne pesait pas bien lourd). Ce qui limitait le rôle de ces précurseurs des neurosciences, c'était leur faible pertinence dans le projet tel que l'IA et les disciplines alliées l'avaient formulé: il s'agissait de caractériser, d'expliquer, de simuler ou de modéliser les processus de la pensée "intelligente". Or cette pensée était conçue comme étant essentiellement constituée de processus conceptuels, délibérés, verbalisables; elle était une capacité des représentants adultes typiques de l'espèce humaine; enfin, et surtout, la caractérisation recherchée se situait à un niveau d'abstraction nettement distingué, et en un sens indépendant, de celui des processus concrets, de l'enchaînement des causes. Ce que l'on voulait comprendre, c'était la *structure logique* de certains processus de "traitement de l'information". La façon dont ces processus étaient effectués, physiquement, dans le tissu nerveux n'avait qu'une importance relative.

Cette conception de l'esprit (ou de ses productions) alliant un monisme ontologique (tout processus mental est un processus matériel) à un dualisme méthodologique (les catégories explicatives au niveau mental ne sont pas identiques aux catégories explicatives au niveau matériel) porte le nom de "fonctionnalisme" et demeure la doctrine catholique des sciences cognitives, même si elle traverse une crise profonde. Ce qui importe ici, c'est que le fonctionnalisme assignait des limites de principe à toute contribution potentielle des neurosciences, sans toutefois la ramener à zéro, puisqu'ils fournissaient au moins aux architectes de l'esprit des contraintes d'ingénierie. Il fallait bien que les processus informationnels postulés par le théoricien de l'IA ou de la psychologie cognitive fussent matériellement exécutables par le cerveau tel qu'on le connaissait (composé, par exemple, de neurones lents et peu fiables). En pratique, ce genre de contrainte n'était guère exploitable par les théoriciens de l'IA. A cette restriction de principe imposée par le cadre théorique au ressources explicatives des neurosciences s'ajoutait le fait qu'à l'époque elles avaient surtout prise d'une part sur la perception (essentiellement sur la vision), sur la motricité, sur les émotions ou dispositions caractérielles, d'autre part sur les pathologies massives ou les différences interindividuelles, tous domaines apparemment bien éloignés de la pensée conceptuelle caractéristique de l'adulte humain moyen.

### **Un spectaculaire renversement de la situation**

Aujourd'hui, l'IA est reléguée aux marges (alors même qu'elle s'est enrichie de résultats et de savoir-faire, et guérie des excès de sa jeunesse), et ce sont les neurosciences qui revendiquent un rôle central dans les sciences cognitives, rôle qui lui est souvent accordé –conviction ou calcul?– par les autres branches du domaine. On pourrait penser que ce renversement résulte d'une concurrence théorique directe, l'IA ayant joué "STI" (systèmes de traitement de l'information) et finalement perdu contre les neurosciences qui misaient sur "SNC" (système nerveux central). L'histoire aurait ainsi montré que la voie de l'abstraction informationnelle se perdait dans les sables, tandis que l'approche concrète de l'organe siège de la cognition, une fois munie des instruments nécessaires, mènerait au but. Ou, autre lecture, l'IA aurait misé sur le modèle de l'ordinateur, et perdu tout simplement parce que ce n'est pas un bon modèle pour le cerveau et que seule une modélisation de l'organe, à partir des données empiriques disponibles, est susceptible de conduire à de bons modèles de la

fonction. Or ces deux lectures, malgré les bribes de vérité qu'elles contiennent l'une et l'autre, sont factuellement fausses. Et ce n'est pas là seulement un point d'histoire, car l'échec, avéré, de l'IA est porteur de leçons applicables à la situation présente.

Ce qui a perdu l'IA (celle de la première période, parfois dite “prométhéenne”) n'est ni le choix du niveau informationnel ni celui de l'ordinateur comme système de traitement, mais une erreur stratégique consistant à ériger la simulation en méthode universelle de découverte et de justification. L'idée était de prendre un à un les processus de l'intelligence (grosso modo, les tâches intellectuelles ou mentales, au sens le plus large, que l'être humain peut accomplir), de les représenter comme résolutions de problèmes, et d'en proposer des simulations informatiques: pour un processus donné, un programme était proposé, ses performances étaient comparées au processus étudié, les écarts constatés suggérant des modifications, et le programme obtenu au terme d'une suite d'approximations successives constituait à la fois la *théorie* du processus et une *preuve* de la théorie (la théorie étant prouvée si, moyennant idéalisation, les résultats obtenus par le programme sont ceux qu'obtient l'agent humain).

Or cette méthode présupposait une identification suffisamment claire et précise du processus à simuler (ce qu'on appelait en IA le *task domain*): jouer aux échecs, se remémorer une liste d'objets présentés oralement, empiler des blocs de bois en structures stables, traduire un texte d'italien en russe, démontrer une proposition logique ou géométrique, identifier une scène visuelle, rédiger un chèque bancaire, répondre à une demande de renseignement fiscal, calculer les doses d'insuline nécessaires au traitement d'un diabète... Comment catégoriser et caractériser cette immense variété de tâches? Sans doute n'était-il pas nécessaire que l'identification initiale soit parfaite, car la recherche du programme adéquat était censée fournir les corrections utiles, mais il fallait que le “prédécoupage” respecte certaines frontières naturelles, faute de quoi l'itération divergerait, ou bien conduirait à une “machine” hybride, arbitraire et fragile, manière de robot ménager répondant au téléphone et faisant monter les œufs en neige dans les cas prévus par le fabricant.

Quelle devait donc être la source de ces identifications préalables? L'IA avait là-dessus deux idées complémentaires: le sens commun (des spécialistes de l'IA) et, dans certains cas, le psychologue ou le linguiste, voire le logicien, le philosophe ou l'anthropologue. L'ironie est que cette réponse était bonne, mais que l'IA n'a pas su en conjuguer les deux volets: elle a dramatiquement surestimé les ressources du sens commun<sup>6</sup>, et corrélativement sous-estimé la difficulté de la tâche du psychologue, du linguiste etc., commettant du coup la troisième et fatale erreur de se croire capable d'orchestrer les recherches relevant des disciplines fondamentales (en quoi consistait son rôle d“interdiscipline”). Bref, l'espoir de substituer à l'étude empirique des capacités mentales, par les méthodes scientifiques traditionnelles, la mise au point de simulations, dans un cadre théorique réduit à l'hypothèse computationnelle, cet espoir était vain (a posteriori, son audace surprend). Et ce sont l'expérience de l'enlisement et la critique par les disciplines directement impliquées (dont les neurosciences n'étaient pas) qui ont mis un terme à l'IA prométhéenne.

L'IA se heurtait à d'autres obstacles, qui l'auraient peut-être arrêtée ultérieurement si, en concluant une alliance plus humble avec la psychologie cognitive et la linguistique théorique, elle avait évité de s'étouffer prématurément faute d'alimentation scientifique. Ces obstacles (intellectualisme, formalisme, fonctionnalisme, problème du cadre, problème du sens commun,...) sont aujourd'hui l'affaire des sciences cognitives dans leur ensemble, qui les négocient ou les contournent dans leur cadre théorique propre. Les neurosciences n'ont pas les ressources conceptuelles nécessaires pour éliminer ces perplexités ou pour développer un point de vue qui leur soit propre.

### **D'où vient le succès actuel des neurosciences?**

Puisque ce n'est pas d'avoir eu raison d'un adversaire naturel (au sens où l'on peut dire par exemple, en première approximation, que le cognitivisme naissant a eu raison du béhaviorisme) que les neurosciences tirent leur dynamisme actuel, quelles en sont les causes? La première est que les progrès des sciences cognitives (y compris, naturellement, de leur composante neurobiologique) ont préparé le terrain: le “prédécoupage” du mental est bien plus avancé qu'il n'était il y a même vingt ans, a fortiori cinquante, et la pratique de l'interdisciplinarité, dans le domaine de la cognition, a atteint une certaine maturité.

---

<sup>6</sup> ... tout en les sous-estimant: elle a surestimé le sens commun comme heuristique scientifique (nos intuitions sur nos processus mentaux sont très limitées), et l'a inversement sous-estimé comme composante de la cognition naturelle.

La seconde cause est que les neurosciences, en partie grâce à l'imagerie, trouvent énormément de choses à dire sur le cerveau et ses fonctions. Et c'est là que la situation évoque, *mutatis mutandis*, celle de l'IA à l'heure de ses premiers triomphes. Comme l'IA naissante, les neurosciences cognitives semblent parfois revendiquer la responsabilité de refonder l'étude de l'esprit sur des bases réellement scientifiques, comme si la nouvelle science de l'esprit que l'IA entendait incarner était en instance de remplacement. Ouvertes aux autres disciplines, elles sont tentées de se considérer comme l'interdiscipline en charge de rassembler les connaissances nécessaires pour l'enquête scientifique sur l'esprit-cerveau, dont elles sont le maître d'œuvre naturel en vertu du fait qu'elles détiennent la compétence s'agissant du cerveau. Car, dans le complexe esprit-cerveau, le cerveau jouit du privilège d'être un système biologique concret, soumis à une série de contraintes repérables par les méthodes éprouvées de la biologie, et plus récemment rendu accessible à l'observation *in vivo*. Symétriquement, l'IA se réclamait naguère des méthodes éprouvées de la logique et de l'analyse conceptuelle, qu'elle entendait appliquer directement à l'objet de l'enquête, à savoir non le cerveau mais ses facultés. Aujourd'hui les neurosciences, comme hier l'IA, fixent donc l'objet privilégié, mais aussi la forme des réponses: aux programmes de l'IA répondent les régions ou systèmes impliqués, les codages neuraux. Comme hier l'IA, les neurosciences estiment avoir par définition un rôle à jouer dans toute enquête sur les facultés mentales, et nombreux sont les psychologues, les linguistes, les philosophes, les anthropologues qui ne concevraient plus de publier un article dans leur domaine sans référence à des données neuroscientifiques, acquises ou attendues, si ténue que puisse être leur connexion réelle à la question traitée (hier c'était à l'IA que s'adressait ce genre d'hommage). Comme hier l'IA, les neurosciences sont certaines que les découvertes d'aujourd'hui sont les premiers pas vers une explication asymptotiquement achevée des fonctions mentales.

Enfin, les neurosciences, tout particulièrement lorsqu'elles sont engagées dans le programme de recherche des neurosciences cognitives, sont promptes à s'offusquer de toute réflexion critique sur leur rôle: elles y voient l'expression d'une attitude anti-scientifique ne pouvant avoir d'autre effet, quelles que soient les intentions proclamées de leurs auteurs, qu'un retour à une psychologie coupée de tout fondement naturaliste. Elles l'estiment également inopportune sur le plan politique, alors qu'elles doivent se défendre, au sein de leur discipline-mère, contre des mastodontes tels que la biologie moléculaire ou la génomique. A l'échelle nationale, elles attendent un soutien loyal, pour affronter dans les meilleures conditions possibles la concurrence mondiale. L'IA, autrefois, faisait valoir les mêmes arguments.

Ce parallèle avec l'IA n'a pas pour but de suggérer que les neurosciences sont aujourd'hui dans une situation à tous égards semblable à celle de l'IA hier, et en particulier qu'elles connaîtront demain le sort qui est celui de l'IA aujourd'hui. Il sert seulement de mise en garde contre un certain type d'exagération, et peut nous aider à préciser la nature des apports spécifiques des neurosciences aux sciences cognitives, qui leur donnent un avantage incontestable sur l'IA première manière.

## 2. L'imagerie fonctionnelle cérébrale

Pendant des décennies (et même, en un certain sens, depuis l'Antiquité), philosophes et psychologues se sont demandé si les images mentales existent réellement, c'est-à-dire si certains processus cognitifs (par exemple, s'imaginer aller d'un endroit à un autre, ou déterminer par la seule réflexion si telle forme géométrique est superposable à telle autre) mettent en jeu des représentations mentales irréductiblement iconiques, ou si au contraire tous les processus cognitifs consistent en manipulations de "phrases" d'une manière de langage interne. Ce débat, écrivait Stanislas Dehaene il y a déjà sept ans, "est aujourd'hui tranché. [...] Quelques expériences de tomographie par émission de positons et d'imagerie fonctionnelle par résonance magnétique auront suffi"<sup>7</sup> à établir définitivement la réalité de l'imagerie mentale Jugement qui figure aussi dans le titre même du livre de Stephen Kosslyn, *Image and Brain: The resolution of the imagery debate*<sup>8</sup>. En réalité, le débat n'est pas tranché, comme le montre en particulier Zenon Pylyshyn<sup>9</sup>, même s'il se pourrait qu'a posteriori le jugement de la communauté s'aligne sur celui de Dehaene et Kosslyn. Pour notre propos, plus intéressant que cette double

<sup>7</sup> Stanislas Dehaene, Introduction, *Le cerveau en action*, collectif, Paris: PUF, 1997, p. 2.

<sup>8</sup> Cambridge, MA : MIT Press, 1994.

<sup>9</sup> V. par ex. Le débat sur l'imagerie est-il terminé? Si oui, de quoi s'agissait-il ?, in E. Dupoux, dir., *Les langages du cerveau*, Paris: Odile Jacob, 2002, 65-88; et aussi Nigel Thomas, article "Mental imagery", *Encyclopaedia of Cognitive Science*, Lynn Nadel, ed., Londres: Nature Publ. & Macmillan, 2003.

controverse (sur l'existence de l'imagerie mentale<sup>10</sup> et sur la question de savoir si l'affaire est aujourd'hui entendue) est l'argument principal avancé par Dehaene: le cortex visuel primaire V1 est impliqué dans les tâches d'imagerie. Ainsi, les bases neurales du phénomène psychologique de l'imagerie mentale comprennent le cortex visuel primaire, lequel fait partie des bases neurales de la perception visuelle, donc l'imagerie met en jeu de véritables images, c'est-à-dire des représentations iconiques. Nous nous référons plus loin à cet exemple.

Parmi de très nombreux autres exemples disponibles, un second cas, plus subtil et plus récent, est fourni par la découverte que le contrôle contextuel de l'action (en gros, la préparation prémotrice d'une action planifiée tenant compte des stimuli perceptifs) dépend des aires de Brodmann 44 et 45, dont la partie incluse dans l'hémisphère gauche n'est autre que l'aire de Broca dont on connaît le rôle dans la production du langage<sup>11</sup>. Contrôle de l'action et langage se trouvent ainsi rapprochés, d'une manière que ne livre nullement l'évidence d'une simple analyse conceptuelle.

### La méthode caractéristique des neurosciences cognitives

**De manière générale, le domaine des neurosciences cognitives produit une quantité considérable de résultats mettant en relation des événements mentaux (psychologiques) ou comportementaux A, B,... et des activations différentielles dans des zones X, Y... du cerveau. Ces résultats prennent l'une des formes caractéristiques suivantes :**

- (i) L'événement mental A est corrélé avec une activité accrue des zones X, Y,... du cerveau ("implication de X, Y,... dans A").
- (ii) La zone X du cerveau est activée aussi bien lorsque l'événement mental A est observé que lorsque l'événement mental B est observé ("implication de X dans A et dans B").
- (iii) L'événement mental A est corrélé avec une activité accrue de la zone X dans le contexte psychologique C et de la zone Y dans le contexte psychologique D ("implication de X dans A quand C et de Y quand D").

Ces résultats sont souvent présentés comme procédant d'une inspiration unique combinant neurobiologie et psychologie, caractéristique de la nouvelle spécialité dont ils relèvent. Ils ne se prêtent pas moins à deux lectures. Selon la première, les processus ou événements psychologiques sont supposés déterminés et connus, et les activations mises au jour en X, Y,... nous renseignent sur tel ou tel aspect du fonctionnement cérébral.. Selon la seconde, les événements cérébraux sont tenus pour compris, la signification qu'ils prennent dans les meilleures théories neuroscientifiques disponibles étant admise, et c'est notre compréhension de certains processus psychologiques qui pourrait s'en trouver améliorée.

### **Le soupçon de vacuité**

Examinons maintenant certaines critiques adressées à ce type de travaux. Laissons de côté tout ce qui touche à la méthodologie et qui relève d'une compétence strictement neurobiologique: caractère indirect des mesures, artefacts dus aux algorithmes de traitement, faiblesse du signal, résolution insuffisante, variabilité individuelle et valeur douteuse des moyennages, non-standardisation des résultats, précarité des constats de non-activation (faux négatifs), etc. La présente discussion n'a d'intérêt que sous l'hypothèse que ces difficultés sont en voie d'être réduites. La question de savoir si une neuroimagerie infirme ferait progresser la connaissance de l'esprit ne se pose pas vraiment, puisqu'elle ne serait pas même en mesure de contribuer significativement à une science du cerveau. La question intéressante est de savoir si une neuroimagerie mûre et forte, en démultipliant l'efficacité des sciences du cerveau, peut révolutionner les sciences cognitives.

<sup>10</sup> A ne pas confondre avec l'imagerie cérébrale: celle-ci désigne les techniques de visualisation des aires actives du cortex au cours de l'exécution de tâches ou de processus mentaux particuliers; celle-là est le phénomène psychologique particulier dont il est question dans la controverse.

<sup>11</sup> E. Koechlin, C. Ody, F. Kouneiher, The architecture of cognitive control in the human prefrontal cortex, *Science* **302** (2003), 1181-5 ; E. Koechlin, T. Jubault, From action to human language : Broca's area and the hierarchical organization of behavior, *soumis*.

Les sceptiques pensent que non: selon eux la neuroimagerie ne fournit que des réponses sans questions, elle n'est qu'un ensemble de techniques « photographiques », elle accumule des faits sans théorie. Une partie de ces reproches est sans intérêt ici, pour la même raison que précédemment: qu'il y ait dans ce domaine en émergence des expériences mal conduites, des résultats sans grande portée, n'a rien d'exceptionnel. Ce qui nous intéresse est la portée de la *meilleure* neuroimagerie possible. C'est à son propos que le soupçon de vacuité doit être examiné.

Les cas les plus favorables sont ceux dans lesquels l'un des membres de l'«équivalence»  $X, Y, \dots \sim A, B, \dots$  est effectivement identifié au-delà de tout doute raisonnable; ainsi des expériences portant sur des actions motrices ou perceptives élémentaires, dont la dimension comportementale observable détermine sans réelle ambiguïté la nature mentale ou cognitive (le  $A, B, \dots$  de l'équivalence); d'apprendre que  $X, Y, \dots$  sont impliqués dans le décours de  $A, B, \dots$  constitue alors un accroissement sensible d'intelligibilité quant aux propriétés de  $X, Y, \dots$ , car  $A, B, \dots$  sont de «bons objets», des catégories naturelles dont les événements singuliers intervenant dans l'expérience sont des représentants. Plus généralement, les processus cognitifs rapides et essentiellement automatiques, associés à des comportements globalement peu modulables, sont un terrain favorable, comme l'attestent du reste les succès de cet ancêtre des neurosciences cognitives qu'est la neuropsychologie traditionnelle. Restent deux familles de cas: (1) ceux dans lesquels  $A, B, \dots$  sont censés être compris mais présentent une complexité ou une variabilité qui font douter qu'ils constituent vraiment des catégories naturelles dont la nature soit conceptuellement claire; (2) ceux dans lesquels ce sont les variables neurales  $X, Y, \dots$  qui sont tenues pour fixes et les variables mentales  $A, B, \dots$  que l'on cherche à mieux cerner. Pour les cas de ce genre, qui n'ont rien de rare, l'interprétation des résultats est hasardeuse. Et c'est ici que le bon sens du neurobiologiste, si cultivé et génial qu'il puisse être, peut se trouver en défaut, et que l'interdisciplinarité trouve tout son sens.

### **Pourquoi, et à quelles conditions, ces résultats peuvent présenter de l'intérêt**

Revenons à la situation simple, du type (i) ci-dessus: «L'événement mental A est corrélé avec l'activation de l'aire X». Ce résultat, disent les critiques, est trivial, car personne (dans le monde scientifique) ne doute que tout événement mental ait un corrélat neural, et un corrélat neural doit bien se situer quelque part dans le cerveau. Il n'y a donc pas de question scientifiquement plausible à laquelle un tel résultat pourrait répondre. Arrêtons-nous un instant sur cette objection.

Prenons d'abord le cas dans lequel la seule chose que l'on sache de X, c'est sa position topographique dans le cerveau. Il n'est pas trivial que A soit toujours corrélé avec une activité de X: cela tend à montrer que le type mental A possède un caractère naturel sur le plan neurobiologique. Si les progrès que l'on nous promet en matière de finesse de résolution se réalisent, et si l'aire X était identifiée sans doute possible comme étant homologue chez tous les sujets, on pourrait en arriver à une réfutation<sup>12</sup> du fonctionnalisme. De même, elle pourrait conduire à rejeter l'idée d'une absence de corrélat neurologique des troubles psychiatriques, notamment de pensées délirantes, qui définit encore aujourd'hui, malgré les effets connus, et partiellement compris, des psychotropes, une manière de ligne de démarcation du territoire du psychiatre,

Cependant, pour en revenir à la pratique actuelle, les résultats du type (i) le plus simple sont rarement très intéressants, et la plupart des expériences portent sur des situations dans lesquelles à X sont attachées des propriétés qui vont au-delà de la seule localisation spatiale. Les neurobiologistes ont accumulé des connaissances considérables sur les fonctionnalités des différentes régions du cerveau, et même si elles sont entachées d'incertitudes, et sont constamment révisées, parfois de manière radicale, elles n'en ont pas moins suffisamment de consistance pour asseoir des résultats de la forme (ii) ci-dessus: l'aire X est impliquée aussi bien dans des tâches de type A que dans des tâches de type B. La leçon que l'on est tenté de tirer est qu'il y a une parenté entre les tâches de type A et les tâches de type B, et une telle parenté peut être un «scoop». Par exemple, l'hippocampe, avec ses structures associées, semble jouer un rôle double: pour la navigation et la mémoire spatiale (répérage égocentrique sans doute dans le lobe pariétal, et allocentrique dans l'hippocampe

<sup>12</sup> Réfutation partielle: d'une part la même fonction mentale pourrait être réalisée matériellement d'une façon différente chez d'autres espèces, ou chez des sujets anormaux, ou dans des systèmes artificiels; d'autre part, du seul fait qu'en réalité cette fonction n'est réalisée que d'une seule manière chez les sujets normaux on ne pourrait déduire qu'elle n'a pas une nature irréductivement fonctionnelle, c'est-à-dire qu'elle a d'autres propriétés que le rôle qu'elle joue dans l'économie mentale et comportementale.

droit), d'une part; pour la mémoire épisodique et autobiographique (dans l'hippocampe gauche)<sup>13</sup>, d'autre part. Que s'orienter dans l'espace et construire une image ou un concept de soi en tant que centre biographique soient liés (via leur soubassement neural respectif) est loin de résulter d'une simple analyse conceptuelle ou de l'introspection. Cependant, pour passer d'une hypothèse de ce genre, finalement assez floue, à une découverte psychologique précise, certaines conditions doivent être remplies: (1) L'activation de l'aire X doit être un phénomène unitaire, et non une conjonction fortuite de processus distincts. (Imaginons un détective qui aurait remarqué que chaque fois qu'un vol a lieu, les volets du 5<sup>e</sup> étage d'un immeuble donné restent closes dans les heures précédant le crime, et que chaque fois que *Don Giovanni* est joué à l'Opéra, les volets du 6<sup>e</sup> étage du même immeuble restent closes dans les heures précédant la représentation. Imaginons alors que le détective, par faiblesse de la vue ou de l'intelligence, classe sous la même rubrique X la fermeture des volets du 5<sup>e</sup> et du 6<sup>e</sup> étage de l'immeuble en question. Il serait tenté de conclure que les vols (A) et les représentations de *Don Giovanni* (B) sont "liés", alors qu'on pourrait conjecturer que le voleur habite le 5<sup>e</sup> étage, le chanteur le 6<sup>e</sup>, et que cette proximité est fortuite et dépourvue de signification.) (2) L'activation de X doit être, sinon l'unique, du moins le principal corrélat neural de A comme de B, à défaut de quoi le lien entre A et B pourrait être accidentel. (Imaginons que X soit activé uniquement lorsque de petits objets noirs sur fond blanc apparaissent dans le champ visuel, que A soit la lecture, et B l'exécution d'un morceau de violon. A et B ont en commun, dans nos cultures et dans les circonstances habituelles, de mettre en jeu, au stade de l'apprentissage ou de l'exécution, la perception de lettres ou de notes imprimées en noir sur fond blanc. Pourtant, A et B ne sont pas liés (du moins en première analyse). Pour les non-voyants, comme dans une culture où les partitions seraient imprimées en rouge sur fond noir, ou encore dans une culture sans notation musicale, X, on peut le présumer, ne serait pas activé par A et par B. Un neuroimageur qui ne disposerait pas d'un accès indépendant à ces distinctions psychologiques, anthropologiques ou comportementales, et qui s'en tiendrait à ses images, serait incapable de tirer une conclusion valide, et n'aurait de choix qu'entre le silence et l'erreur.

Il ne s'agit pas ici d'affirmer que les conditions qui viennent d'être énoncées ne sont jamais réalisées, ni que les spécialistes ne sont pas conscients de ce genre de réquisits, mais seulement d'insister sur le fait qu'à eux seuls, les résultats d'imagerie de type (ii) ne suffisent pas à conclure quoi que ce soit concernant le lien entre A et B, *sauf* parfois, mais alors ce n'est pas rien, sur le plan clinique: on peut penser que des patients privés, par exemple du fait d'une lésion cérébrale, de la fonctionnalité A, le seront également de la fonctionnalité B, ce qui peut être d'une grande importance pour le diagnostic, le pronostic et la thérapeutique.

Les résultats concernant l'imagerie mentale mentionnés au début de la présente section illustrent bien la situation. Le problème de départ est de savoir si l'imagerie mentale A met en œuvre des processus ou représentations iconiques. L'imagerie fonctionnelle montre que A est corrélé à une activation de X, et l'on sait que X est impliqué dans la vision (B). Si l'on admet (ou que l'on stipule) que la vision fait intervenir à son tour des processus iconiques, on est tenté de conclure que A est de nature iconique et que "le débat est tranché". Mais il faut aussi s'assurer que le cortex visuel primaire est uniquement activé lorsque des représentations iconiques sont manipulées, et que l'imagerie mentale ne déclenche pas V1 pour une raison indirecte (par exemple, si l'imagerie était accompagnée d'un écho, non essentiel pour l'accomplissement de la tâche, en sorte qu'elle serait réalisable malgré une inhibition temporaire de V1 par SMT<sup>14</sup>)<sup>15</sup>. A ces difficultés générales s'ajoute dans ce cas la question du bon « découpage » : où faire passer la frontière entre représentations iconiques et représentations symboliques, sachant que l'on se place au niveau des processus eux-mêmes, et non sur le plan émergent des représentations conscientes ou de la reconstruction conceptuelle : de cette difficulté-là il serait téméraire d'affirmer qu'elle est aujourd'hui surmontée.

### **Descendre au niveau du neurone individuel: le cas des neurones-miroirs**

<sup>13</sup> N. Burgess, E.A. Maguire & J. O'Keefe, The human hippocampus and spatial and episodic memory, *Neuron* 35 (2002), 625-641.

<sup>14</sup> Stimulation magnétique transcrânienne. V. p. ex. *PSNI*, 1 (2003), 40-43 ou E.M. Robertson, H. Théoret & A. Pascual-Leone, Studies in cognition : The problems solved and created by transcranial magnetic stimulation, *J. of Cognitive Neuroscience* 15 (2003), 948-960.

<sup>15</sup> Il ne s'agit évidemment pas ici d'une discussion scientifiquement réaliste de cette question complexe et controversée, dans laquelle interviennent quantité d'autres considérations et que l'auteur n'a pas compétence à mener. Mais précisément l'un des arguments développés dans cet article est qu'il n'y a en général pas de résolution simple par la seule neuroimagerie de questions psychologiques intéressantes.

Pour écarter la menace d'une conjonction fortuite, au sein d'une même aire, de deux processus essentiellement indépendants, on peut procéder (mais seulement chez l'animal) à des enregistrements de neurones individuels (*single neuron recordings*). Il semble en effet plausible que si deux processus cognitifs A et B s'accompagnent de la décharge X d'un même neurone, ils sont liés. Ici encore, il faut supposer que la décharge du neurone joue un rôle essentiel dans les processus en question, sans toutefois qu'elle puisse à elle seule caractériser à la fois A et B, auquel cas ou bien il s'agirait en réalité d'un seul événement cognitif, ou bien A et B ne se distingueraient *que* par une propriété non cérébrale (dans le jargon des philosophes, le mental ne "surviendrait" pas sur le cérébral) –deux éventualités qu'on peut écarter. Un exemple qui fait couler beaucoup d'encre depuis quelques années est celui des "neurones-miroirs"<sup>16</sup>. Certains neurones prémoteurs de l'aire F5 du cortex des macaques ont la propriété de décharger aussi bien lorsque l'animal tend la main pour prendre des cacahuètes disposés devant lui que lorsqu'il observe un expérimentateur (ou un autre macaque) en faire autant. Des systèmes miroirs analogues semblent bien exister chez l'homme (dans l'aire de Broca qui est homologue de l'aire F5 du macaque). C'est une découverte<sup>17</sup> que beaucoup considèrent comme capitale, car elle fournirait selon eux une explication naturaliste directe de la compréhension de l'action intentionnelle d'autrui: en observant le mouvement de l'autre, je comprends ce qu'il veut faire, en vertu du fait que ce geste provoque chez moi une disposition à agir dont la nature intentionnelle m'est révélée par l'anticipation que crée cette disposition. C'est l'idée de base de la *théorie motrice* de la cognition sociale; pour le dire rapidement, je te perçois comme un *agent* semblable à moi, capable d'entretenir des intentions analogues, parce que mon cerveau, d'un côté, identifie ton geste et le mien, et de l'autre me fait connaître le sens de mon propre geste en me permettant d'anticiper ses conséquences. La discussion ne peut être même résumée ici, mais on peut en retenir quelques leçons:

(1) L'existence des neurones-miroirs est significative, et ne pouvait être prédite à partir de considérations seulement théoriques; elle n'est pas empiriquement triviale.

(2) Sa portée dépend crucialement de l'hypothèse qu'il s'agit de neurones prémoteurs, donc remplissant une fonction que l'on a préalablement identifiée avec certitude (à savoir la préparation du mouvement), et cette hypothèse à son tour s'enracine dans des strates épaisses de théorisation tant en psychologie qu'en neurobiologie qui ne doivent rien à l'imagerie.

(3) Sa signification exacte est encore loin d'être claire, et dans les débats en cours les imageurs sont amenés à se placer sur le terrain psychologique, philosophique, voire linguistique, où ils rencontrent des spécialistes de ces disciplines<sup>18</sup>; ils n'ont pas le monopole de la compétence.

(4) Dans la mesure où une théorie substantielle stable des neurones-miroirs émergera, elle constituera une contribution non triviale des neurosciences dans leur ensemble aux sciences cognitives, en sorte que...

(5) ...neurosciences et psychologie auront à parts à peu près égales apporté ainsi à la connaissance de l'esprit humain une contribution dont le caractère non trivial est attesté par le fait qu'elle fournirait une explication de la capacité à reconnaître *directement et non intentionnellement* l'intention d'autrui à partir de ses seuls mouvements, invalidant ainsi une conception solidement ancrée dans le sens commun comme dans certaines traditions de psychologie savante (le caractère inférentiel et indirect de cette reconnaissance).

### **Vers une carte ou un organigramme cérébral?**

Plus s'accumulent les résultats de corrélation que nous venons d'examiner, moins on peut se satisfaire d'une simple juxtaposition. La question se pose de savoir comment les articuler. A quel genre de connaissance ou de représentation du cerveau concourent-ils? S'ils avaient tous la forme élémentaire "A ~ X", A étant une fonction cognitive ou comportementale donnée d'entrée de jeu, et X l'activation d'une aire

<sup>16</sup> Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. Premotor cortex and the recognition of motor actions. *Cog. Brain Res.*, 3 (1996): 131-141; Gallese, V., Fadiga, L., Fogassi, L. and Rizzolatti, G. Action recognition in the premotor cortex. *Brain* 119 (1996): 593-609

<sup>17</sup> V.S. Ramachandran par exemple y voit la plus importante découverte de la décennie en neurosciences: Edge (2000) - [www.edge.org/3rd\\_culture/ramachandran](http://www.edge.org/3rd_culture/ramachandran).

<sup>18</sup> V. par exemple le colloque en ligne sur [www.interdisciplines.org](http://www.interdisciplines.org).

cérébrale, et s'ils étaient simplement juxtaposés, leur assemblage ne constituerait qu'une nouvelle "carte" phrénologique, dont l'intérêt se limiterait à certaines situations cliniques, et dont la portée théorique serait faible.

L'erreur de certains sceptiques, tels Fodor<sup>19</sup>, est de croire que c'est l'unique objectif que puissent viser les neurosciences cognitives adossées à l'imagerie. En réalité, elles disposent de résultats des trois types ci-dessus, qui se prêtent à une combinatoire et à un enchevêtrement de fonctions: à la pure mise en correspondance de type (i) simple—une fonction, une aire— s'ajoutent avec le type (i) complexe —une fonction, plusieurs aires, voire plusieurs processus neuraux concomitants— des décompositions en éléments plus simples, avec le type (ii) des partages de ressources (ii), et, avec le type (iii) des différenciations,. On n'a donc plus affaire à une simple carte, mais à un système, à un "organigramme" cérébral. Mais évidemment la constitution de l'organigramme à partir des résultats partiels fait appel à des connaissances et à un art de la conjecture qui ne doivent rien à la neuroimagerie, et tout aux neurosciences fondamentales et à d'autres branches des sciences cognitives.

Reste qu'il semble difficile de contester l'intérêt d'une théorie "organigrammatique" du cerveau, quels qu'en soient les auteurs. Le sceptique ne doit-il pas se rendre? Il devrait au moins être amené à réfléchir. Mais il lui reste des portes de sorties. Il peut d'abord rappeler qu'on est encore bien loin du moment où une telle théorie sera disponible, et soupçonner qu'elle ne le sera peut-être jamais: l'IA n'avait-elle pas été victime, en son temps, de l'illusion du premier pas (penser que monter à un arbre nous rapproche inexorablement du moment où nous pourrions monter à la lune)? Il peut aussi réaffirmer sa conviction fonctionnaliste et maintenir, pour des raisons théoriques, le principe d'une séparation de deux ordres d'intelligibilité: aucun organigramme du cerveau, qui par définition ne met en jeu que des causes, ne peut nous éclairer sur la nature et sur la dynamique de l'esprit, siège de significations et de raisons. Ou enfin, si le fonctionnalisme lui semble reposer lui-même sur des confusions, il peut contester que la neuroimagerie fournisse des explications, car *ce qui est à expliquer* (le domaine que vise la psychologie) demeure entouré d'un épais brouillard<sup>20</sup>.

### 3. Les neurosciences sont une discipline théorique

Si la discussion précédente a permis, on peut du moins l'espérer, de clarifier les enjeux, elle laisse néanmoins planer le doute. D'un côté, on croit comprendre que la neuroimagerie contribue incontestablement à la connaissance du cerveau, et qu'ainsi elle apporte aux sciences cognitives une moisson de faits qu'elles ne peuvent ignorer. D'un autre côté, on nous invite à relativiser son apport, soit en soulignant sa dépendance vis-à-vis d'autres disciplines, soit en niant qu'elle puisse amener les sciences cognitives à clarifier, réviser ou enrichir notablement leurs conceptions de l'esprit.

Désaccords théoriques sur fond de rivalités institutionnelles? Sans doute, mais il y a également une difficulté épistémologique. Le développement rapide de la neuroimagerie semble encourager l'idée que les faits, qui sont concrets et solides, s'opposent aux théories, abstraites et fragiles, et que les neurosciences, grâce à l'imagerie, fournissent les faits permettant à la science de l'esprit/cerveau d'échapper à l'incertitude des théories, produit d'une psychologie encore trop asservie au modèle spéculatif de la philosophie et d'un certain mode idéologique de pratiquer les sciences de l'homme. Ce "neurofactualisme" (ou "neuropositivisme"<sup>21</sup>) doit être clairement récusé: les neurosciences sont une science théorique, ce qui signifie notamment que les connaissances qu'elles produisent sont, au sens technique, des *théories*. Les théories vont au-delà des faits; ceux-ci les sous-déterminent; les faits eux-mêmes sont imprégnés de théorie; les hypothèses d'une théorie

<sup>19</sup> V. p.ex. J. Fodor, *Let your brain alone*, *London Review of Books*, 21, 19, 30 sept. 1999. La critique de Fodor s'applique sans doute à certains programmes de recherche, qui semblent en effet se contenter de nourrir une néophrénologie. Nous nous intéressons ici, nous l'avons dit, à la meilleure contribution possible de l'imagerie à la connaissance de l'esprit.

<sup>20</sup> Un sceptique de ce genre s'inspirerait probablement de Wittgenstein, ou peut-être de la phénoménologie existentielle issue de Heidegger, ou encore de philosophes pragmatiques contemporains.

<sup>21</sup> Un terme qu'il vaut mieux éviter, car il prête à confusion: il peut également servir à une certaine psychologie "humaniste" pour discréditer l'entreprise des neurosciences elles-mêmes, en tant que science de l'homme. En rejetant le "neurofactualisme", je n'entends nullement me ranger dans le camp des "humanistes" (le véritable humanisme se place à un autre niveau): je m'en prends seulement à une conception naïve de la nature et du rôle des faits dans les sciences, abandonnée depuis belle lurette s'agissant des sciences mûres, mais retrouvant du service dans une discipline immature telle que les sciences cognitives.

confrontent collectivement le tribunal de l'expérience; enfin les théories, ainsi que les explications qu'elles suggèrent, font intervenir des entités “cachées”, c'est-à-dire non directement accessibles à l'expérience.

Affirmer que les neurosciences sont une discipline théorique, c'est donc dire qu'elles possèdent ces traits. Pour être plus précis, elles sont une branche d'une science théorique, à savoir la biologie, et elles héritent de leur caractère théorique. Quant à ce qu'on appelle (depuis peu) “neurosciences cognitives”, ce n'est pas, contrairement aux apparences lexicales, une simple spécialité au sein des neurosciences, mais un programme de recherche, ou mieux une toute jeune “matrice disciplinaire”, pour emprunter à Thomas Kuhn un terme qu'il a finalement préféré au trop versatile “paradigme”<sup>22</sup>. On peut maintenant préciser les raisons de ce qui a pu apparaître au début de cet article comme une simple précaution verbale, voire une chinoiserie. Les neurosciences cognitives relèvent non pas d'une seule discipline fondamentale, mais de plusieurs, dont l'articulation appelle des décisions qui ne s'imposent pas également à tous. Elles ont de fait une manière déterminée d'aborder leur objet, fondée notamment sur l'option méthodologique de l'imagerie et sur le privilège donné à un certain type d'explication (celle dont on distinguait les types (i), (ii), (iii) au début de la section 2). Ce sont précisément ces décisions théoriques qui leur fournissent les moyens de constituer une représentation de leur objet qui pourrait (une hypothèse que l'on précisera dans un moment) modifier profondément la donne dans les sciences cognitives. Enfin, les neurosciences cognitives ne constituent pas une *théorie* unique, au sens strict, car elles produisent et accueillent une pluralité de théories particulières, qui diffèrent entre elles soit par les phénomènes qu'elles visent, soit par les hypothèses qu'elles défendent dans un même domaine, soit des deux façons.

### **La juste place de la neuroimagerie et neurosciences cognitives**

Sous cet angle, certaines de nos perplexités se dissipent. En replaçant l'imagerie au sein des neurosciences, et les neurosciences au sein de la biologie, on voit que les sciences cognitives ne sont pas davantage fondées à récuser la contribution de la neuroimagerie que les astronomes du début du XVIIe siècle ne le sont à refuser de regarder dans la lunette de Galilée: s'ils veulent élaborer une théorie de la cognition, ils ne peuvent se dispenser de chercher à comprendre en vertu desquelles de ses innombrables propriétés le seul système dont nous sachions à coup sûr qu'il est cognitif l'est effectivement, ni rejeter a priori l'avis globalement favorable des spécialistes du cerveau sur la fiabilité et l'utilité des instruments de neuroimagerie. Mais on voit aussi que *cet avis, scientifiquement motivé, est une garantie essentielle*: les contemporains de Galilée avaient raison de s'enquérir sur la véridicité et sur les conditions d'emploi de la lunette; ils avaient également raison d'exiger des arguments en faveur des thèses copernico-galiléennes qui soient indépendants des images procurées par la lunette, et jamais Galilée n'a pensé fonder la défense de ses théories sur ses seules observations. Enfin, il est clair que les faits que fournit la neuroimagerie sont imprégnés de théorie, d'une part en raison du caractère hautement indirect de la méthodologie, d'autre part pour la raison générale illustrée par l'exemple précédent: jamais la lunette à elle seule n'aurait permis au premier astronome médiéval venu de conclure à la vérité du copernicanisme. Ce sont là des banalités, c'est bien pourquoi il ne faut pas les perdre de vue.

Une question reste en suspens, celle qui figure dans le titre du présent article: les neurosciences cognitives ont-elles un potentiel révolutionnaire au sein des sciences cognitives? Avant d'y répondre, remarquons que seule une matrice disciplinaire peut révolutionner une science: c'est pourquoi il est préférable de conserver l'étiquette “neurosciences cognitives” pour désigner une certaine matrice disciplinaire, et non comme le proposent certains, l'ensemble des recherches en sciences cognitives revendiquant un lien privilégié avec les neurosciences, voire les sciences cognitives dans leur totalité au prétexte qu'elles *devraient* faire allégeance aux neurosciences. (Il est possible que la distinction devienne oiseuse dans un avenir proche ou lointain; ce serait justement le signe d'un ralliement général de la discipline à la matrice en question.)<sup>23</sup>

### **Un potentiel révolutionnaire?**

<sup>22</sup> T. Kuhn, *The Essential Tension*, University of Chicago Press, 1977 ; trad. fr. *La tension essentielle*, Paris : Gallimard, 1990, chap. 12 (article original paru en 1974).

<sup>23</sup> On peut aussi comprendre que les administrations de la recherche préfèrent dès à présent la solution extensionnelle.

Se demander si les neurosciences cognitives révolutionnent les sciences cognitives, ou peuvent peut-être le faire, c'est d'abord, pour le redire, leur assigner une place séparée (sans nier leur imbrication, puisque les neurosciences ont toujours été considérées comme un membre actif de la fédération). C'est ensuite exprimer, au risque de fâcher certains éminents collègues, l'opinion que ce n'est pas encore le cas. Mais c'est aussi, à l'inverse, aller au-delà de ce que leurs partisans modérés se contentent de revendiquer pour elles: un rôle dans une construction commune déjà bien engagée – apporter à point nommé des réponses à des questions formulées par les autres disciplines. Révolutionner une science, c'est *aussi* répondre à des questions qu'elle a été incapable de poser.

Parmi ces questions, certaines sont exprimables mais ne se posent pas, car dans le cadre conceptuel considéré on ne peut imaginer une expérience qui imposerait, ou suggérerait seulement une réponse. Un exemple est celui-ci: existe-t-il des représentations mentales? Dans le cadre classique des sciences cognitives, qui définissent la cognition comme une dynamique de représentations internes, on voit mal quelle épreuve empirique conforterait une réponse quelle qu'elle soit. Daniel Amit, il y a dix ans, a remarqué qu'une série d'expériences d'enregistrement de neurones dans le cortex temporal du singe avait permis à Miyashita de mettre en évidence des états neuraux (plus précisément des assemblées de neurones) remplissant les fonctions essentielles attribuées a priori aux représentations internes<sup>24</sup>. Pour le dire autrement, les représentations internes comme nécessité conceptuelle (dans un cadre donné) ou comme métaphore sont remplacées par des phénomènes empiriquement repérables remplissant les mêmes fonctions.

D'autres questions ne sont pas même formulables, car elles nécessitent un vocabulaire dont la science constituée ne dispose pas. Les neurosciences cognitives nous présentent à certains moments une image étrange, qui n'est ni l'image familière du cerveau, ni l'image familière de l'esprit. L'étrangeté ne tient pas à des différences repérables par rapport aux conceptions admises. Elle est due à ce qu'on commence à appeler les "actions à distance": la structure corticale serait radicalement différente de tout ce que la philosophie et la psychologie peuvent imaginer comme "découpage" concevable de l'esprit, en sorte que ce qui semble, dans la conception classique, constituer une unité fonctionnelle, sinon anatomique, est en réalité un simple *pattern* résultant de l'action conjuguée de systèmes cérébraux éloignés; inversement, le même système pourrait contribuer à des processus extrêmement différents, conceptuellement très distants. C'est ce que désigne parfois l'expression ambiguë de *complexité fonctionnelle*.

Non seulement donc, comme les connexionnistes l'avaient envisagé, les *représentations* cérébrales seraient distribuées<sup>25</sup>, mais il en serait de même des *fonctions* cérébrales. Ainsi, les schèmes caractéristiques de l'imagerie formulés au début de la section 2 ne seraient que des échafaudages provisoires, car les fonctions cognitivo-comportementales A, B,... n'entreraient pas dans des relations systématiques, si complexes soient-elles, avec des processus neuraux X, Y,..., ce dont on s'apercevrait en échouant dans les tentatives de combiner ces résultats en une "neurocarte" légendée dans le vocabulaire psychologique.

### **Un exercice de prospective scientifique: neurophilosophie ou émergence d'un niveau macroneurobiologique autonome?**

Des prémices de cette situation sont déjà apparus, et elles suffisent à montrer que les neurosciences cognitives bousculent dès à présent l'assise des sciences cognitives. Mais rien pour le moment n'indique que les réajustements nécessaires soient hors de portée. La question que je voudrais poser pour conclure, c'est de savoir si les neurosciences cognitives pourraient déboucher sur une révision si radicale de l'ontologie du mental qu'elle provoquerait une déconnexion de leur domaine avec la psychologie telle que nous la connaissons, déclenchant ainsi une révolution dans les sciences cognitives.

<sup>24</sup> D.J. Amit, The Hebbian paradigm reintegrated: Local reverberations as internal representations, *Behavioral and Brain Sciences* **18**, 4 (1995), 617-626; Y. Miyashita & H.S. Chang, Neural correlate of pictorial short-term memory in the primate temporal cortex, *Nature* **331** (1988) 68

<sup>25</sup> V. p. ex. G.E. Hinton, J.L. McClelland & D.E. Rumelhart, Distributed representations, in D.E. Rumelhart, J.L. McClelland & the PDP Research Group, *Parallel Distributed Processing*, vo. 1, Cambridge, MA : MIT Press, 1986, 77-109. Voir aussi, sur ce point mais aussi pour la perspective générale qu'il propose, P. Smolensky, On the proper treatment of connectionism, *Behavioral and Brain Sciences* **11** (1988), 3-71 ; trad. in G. Fisette & P. Poirier, dir., *Philosophie de l'esprit*, vol. 2, Paris : Vrin, 2002.

Cette situation n'a-t-elle pas été depuis longtemps envisagée, voire prédite avec assurance par Paul et Patricia Churchland sous le nom de “neurophilosophie”<sup>26</sup>? Il y a deux grandes différences entre leur vision et celle qui vient d’être esquissée. Leur thèse était que les neurosciences fourniraient une image scientifique unifiée de l’“esprit/cerveau”, qu’elles constitueraient à la fois une science du cerveau (comme leur nom l’indique) et une science de l’esprit, autrement dit une psychologie scientifique. Au contraire, ce que j’envisagé ici, à titre de pure hypothèse, comme horizon des neurosciences cognitives, est la constitution de *deux* représentations ou niveaux descriptifs au sein d’une branche de la biologie, une neuroscience fondamentale et une “neurocognition” ou “macroneurobiologie”. Certes, l’idée de deux niveaux de description neurobiologique n’a rien d’original, elle est présente dans l’organisation même de la recherche et dans les dénominations « neurosciences intégratives » ou « neurobiologie fonctionnelle » opposées à « neurobiologie moléculaire et cellulaire ». Dans ce contexte, le niveau supérieur ou intégratif a vocation à entrer en correspondance ou à s’articuler systématiquement avec la description psychologique.

Ce que proposaient et proposent toujours les Churchland, c’est d’éliminer purement et simplement la description psychologique, plus précisément de considérer que la description macroneurobiologique se substituera graduellement à l’actuelle psychologie fonctionnaliste, qui tendrait vers l’extinction. Le scénario proposé ici présente une première différence avec le leur : d’une part, il donne au niveau supérieur ou macro une certaine autonomie ontologique relativement au niveau micro ; mais d’autre part, il ne postule nullement le dépérissement de la psychologie scientifique comme science du mental. Au contraire, et à la différence également avec la conception « correspondantiste-articulatoire » (le schéma unificateur mais non réductionniste selon lequel les descriptions articulées aux niveaux fonctionnel et cérébral présentent des correspondances systématiques dont la mise au jour étanche la soif d’explication, la possibilité se dessine qu’entre macroneurobiologie et psychologie cognitive existerait une complémentarité, au sens où Bohr l’a proposé pour la mécanique quantique, interdisant l’accès simultané aux descriptions macroneurobiologique et psychologique d’un même processus.<sup>27</sup>

La seconde différence avec la « neurophilosophie » concerne la « psychologie naïve » ou de sens commun (*folk psychology*), que les Churchland vouent également à l’élimination<sup>28</sup>. Au contraire, dans le scénario proposé, comme chez un philosophe fonctionnaliste partisan du « réalisme intentionnel » comme J. Fodor, la psychologie scientifique pourrait l’abriter, ou du moins en abriter une partie, convenablement amendée ; mais il se pourrait aussi qu’elle constitue une troisième sphère d’intelligibilité. Elle serait ce que le philosophe Wilfrid Sellars appelait l’« image manifeste »<sup>29</sup> de l’esprit, l’« image scientifique » étant alors composée de deux représentations complémentaires (au sens fort), macroneurobiologique et (proprement) psychologique d’autre part<sup>30</sup>.

Contrairement aux Churchland, je ne prétends pas que le scénario que je viens d’esquisser soit nécessaire ou seulement probable. Il se pourrait que les neurosciences cognitives ne détiennent aucun potentiel révolutionnaire, et qu’elles se contentent d’un rôle analogue à celui qu’elles remplissent aujourd’hui au sein des sciences cognitives. Pour revenir à la comparaison initiale, comme l’IA première manière, elles auront alors été une matrice disciplinaire qui doit renoncer à s’emparer d’un domaine de la nature pour endosser le statut plus modeste de branche spécialisée. Il subsisterait néanmoins une différence de principe entre les deux programmes. Les neurosciences cognitives sont une science empirique, l’IA mûre est une discipline formelle ou structurale. Les liens de l’IA avec les sciences cognitives sont contingents, mais inversement son domaine d’application peut à tout moment s’étendre à d’autres systèmes que le cerveau, dès lors qu’ils abritent une forme même dérivée ou rudimentaire d’intentionnalité. Les neurosciences cognitives, en tant que spécialité, sont au

<sup>26</sup> Paul M. Churchland, *Eliminative materialism and the propositional attitudes*, *J. of Philosophy* 78 (1981), 67-90; – *A Neurocomputational Perspective*, Cambridge, MA : MIT Press, 1989. Patricia Smith Churchland, *Neurophilosophy. Towards a unified science of the mind/brain*, Cambridge, MA : MIT Press, 1986; trad.fr. *Neurophilosophie. L’esprit-cerveau*, Paris: PUF, 1999.

<sup>27</sup> Le transfert du concept de complémentarité de Bohr vers d’autres domaines a été récemment proposé par Rom Harré, *Resolving the reduction/emergence debate*, in M. Kistler, dir., *Emergence and Reduction*, numéro spécial de la revue *Synthèse*, à paraître. (V. aussi mon commentaire.)

<sup>28</sup> Pour les spécialistes : la position des Churchland est présentée ici à l’envers. C’est d’abord la psychologie naïve qu’ils veulent éliminer au profit des neurosciences, et la psychologie scientifique cognitiviste ou fonctionnaliste ne mérite l’élimination que parce qu’elle est, selon eux, indissociable de la psychologie naïve.

<sup>29</sup> W. Sellars, *Science, Perception and Reality*, Londres : Routledge & Kegan Paul, 1963 ; trad. fr. in G. Fiset & P. Poirier, dir., *Philosophie de l’esprit*, vol. 1, Paris : Vrin, 2002.

<sup>30</sup> Une situation qu’on retrouve justement dans la connaissance des phénomènes physiques fondamentaux.

contraire riviées au système nerveux central, mais elles possèdent une dimension supplémentaire qui leur est propre, la compétence clinique. Si Diderot<sup>31</sup> a raison de penser que seuls les médecins peuvent faire de la bonne métaphysique, les neurosciences cognitives sont peut-être, tous compte faits, la reine des sciences cognitives. Et s'il a tort, il leur restera sans doute l'éminente dignité de servir, à leur rang, à la fois la médecine et les sciences pures de l'esprit.

---

<sup>31</sup> *Encyclopédie*, article « Locke ».

# Michel Bitbol

## Des "phénomènes" de Kant à la théorie quantique de l'information

### Le rôle de la philosophie des sciences

*Michel Bitbol est directeur de recherche au CNRS et chargé de cours à l'université Paris 1. Ses domaines sont la Philosophie de la physique moderne, l'Histoire de la physique du vingtième siècle, la Philosophie des sciences*  
*Il est "membre plein" du CREA: Centre de Recherches en Epistémologie Appliquée - (CREA/Ecole Polytechnique), 1, rue Descartes, 75005 Paris*

*Il a publié de nombreux articles et ouvrages. Parmi ceux-ci, le lecteur s'intéressera particulièrement à :*

- ▶ *Mécanique quantique, une introduction philosophique, champs-Flammarion, 1997*
- ▶ *L'aveuglante proximité du réel, champs-Flammarion, 1998*
- ▶ *Physique et Philosophie de l'Esprit, Flammarion 2000*

#### *Interview par Jean-Paul Bacquiast, d' « Automates intelligents »*

*Automates Intelligents (AI) : Nous avons pensé qu'il était indispensable de vous faire mieux connaître de nos lecteurs, notamment de ceux qui n'ont pas eu la possibilité d'étudier la philosophie des sciences. Vous représentez, si vous me permettez de le dire, une sorte d'idéal en matière de connaissance, dans la mesure où vos recherches réalisent la synthèse entre la philosophie et un domaine scientifique de plus en plus actuel, la physique quantique. Les philosophes des sciences ne manquent sans doute pas en France, mais peu d'entre eux apportent le regard nécessairement renouvelé qu'impose la mécanique quantique. Depuis les origines de celle-ci, certes, les commentaires philosophiques n'ont pas manqué, parfois à la limite du surréalisme. Mais pour être pertinent aujourd'hui, il faut avoir soi-même travaillé ce domaine, ce que vous avez fait et ce qui n'est pas le cas de tous.*

*Faut-il cependant que des gens qui travaillent principalement dans le domaine des technologies de l'information, comme c'est le cas de la plupart de nos lecteurs, s'intéressent à la mécanique quantique ? La réponse est évidemment affirmative, dans la mesure où, aujourd'hui, les composants électroniques et nanotechnologiques atteignent la taille de l'atome individuel, échelles où les phénomènes quantiques ne peuvent plus être ignorés. Par ailleurs, de plus en plus de dispositifs, soit dans l'instrumentation, soit dans le calcul, manipulent ou vont manipuler des bits quantiques. Le concept de théorie quantique de l'information prend progressivement de la consistance(1). Il serait loisible de ne pas chercher à savoir le type de monde éventuellement sous-jacent à ces phénomènes, en se satisfaisant des résultats pratiques des manipulations. Mais la curiosité humaine est telle que l'on s'interrogera de plus en plus sur les soubassements physiques possibles, invisibles et inévitablement mystérieux, de ces phénomènes. Il faut alors que des philosophes de la connaissance, tels que vous, interviennent pour apprendre aux citoyens à prendre le recul nécessaire face aux images utilisées pour essayer de traduire le formalisme quantique d'une façon intelligible. C'est tout le problème de la critique du réalisme naïf ou métaphysique, sur laquelle nous allons revenir, qui se trouve posé.*

*Voulez-vous cependant, pour en venir à vous, nous rappeler comment vous vous êtes investi dans la philosophie des sciences ?*

**Michel Bitbol (MB)** : J'ai d'abord reçu une formation médicale, pendant laquelle j'ai également étudié la physique dans le but de poursuivre des recherches en biophysique. Durant les dix années que j'ai consacrées à la biophysique, j'ai utilisé diverses techniques de résonance paramagnétique à travers lesquelles j'ai acquis une certaine pratique du formalisme quantique. J'ai ensuite rencontré des physiciens préoccupés par la philosophie des sciences. Ce fut le cas notamment de Bernard d'Espagnat, qui m'a beaucoup appris par sa critique systématique des multiples versions possibles du réalisme scientifique direct en physique microscopique. A son contact, puis seul, j'ai rassemblé la matière et les réflexions qui font l'objet des trois principaux ouvrages que vous avez cités dans votre introduction, ainsi que de nombreux articles.

**AI** : *Aujourd'hui, vous juxtaposez la recherche et l'enseignement...*

**MB** : Oui, au sein du CREA et à l'Université Paris 1. Pour le CREA, Centre de recherches en épistémologie appliquée, le mot épistémologie est très important. Il s'agit vraiment du cœur de son domaine. On parle aussi de "sciences cognitives", terme de plus en plus d'actualité dans les thèmes de recherche. Nous sommes un centre interdisciplinaire de recherche sur la cognition. Vous trouvez au CREA à la fois des mathématiciens qui font des modèles de cognition, des biologistes qui proposent des expériences à partir de ces modèles et des philosophes de la connaissance qui s'inscrivent dans les travaux de leurs collègues sous l'angle de la réflexivité, c'est-à-dire de l'analyse des présupposés de la recherche.

J'apporte moi-même à ces travaux une thématique de réflexion sur la connaissance qui vise notamment à ramener ce que l'on considère souvent comme des traits du monde à des traits spécifiques de notre interaction avec le monde ou à ce que nous présupposons quand nous essayons de le connaître.

Il s'agit d'un exercice fructueux et salutaire, que j'ai expérimenté voici maintenant plus de dix ans. A chaque fois que l'on essaye, comme le disait Einstein, non pas de croire ce que font, on révèle une quantité de traits caractéristiques du contenu même de leurs connaissances. Aussi curieux que cela puisse sembler, le contenu de leurs connaissances est déterminé par la manière ou les méthodes qu'utilisent les scientifiques pour connaître... Comme le disait Wittgenstein *"Dis-moi comment tu cherches, et je te dirai ce que tu cherches"*.

**AI** : *C'est-à-dire en fait aussi que ce contenu de connaissances est fonction de l'histoire et même du tempérament du sujet qui élabore les connaissances...*

**MB** : Là, vous allez un peu trop loin. Ce qui informe le contenu des connaissances n'est ni d'ordre individuel, ni historiquement contingent. C'est un ensemble de règles collectives d'investigation, épuré par les contraintes techniques et les nécessités de la communication. L'histoire est seulement ce qui permet d'atteindre ce niveau d'épuration.

C'est pourquoi, au CREA, nous sommes revenus avec profit sur les débats philosophiques des siècles précédents, par exemple entre Hume et Kant. Nous sommes loin de considérer que ceux-ci, comme on le dit parfois, ont perdu toute actualité ? aujourd'hui. La philosophie de Kant était certes adaptée à la physique newtonienne. Il n'en reste pas moins qu'un certain nombre de grandes orientations de sa philosophie sont extrêmement porteuses et fécondes. Il faut seulement faire le tri entre ce qui est attaché à une époque et ce qui est de l'ordre plus général de la découverte d'une attitude philosophique nouvelle.

**AI** : *Pouvez-vous nous préciser, puisque vous faites allusion à Kant, comment cette philosophie peut nous aider aujourd'hui à expliquer ce paradoxe du réalisme. Je veux dire expliquer la juxtaposition entre la remise en cause du réalisme, qui est de plus en plus aujourd'hui partagée par les philosophes de la connaissance, et le fait que chacun, qu'il s'agisse du grand public ou des scientifiques eux-mêmes, continuent à parler comme s'il existait bien un réel indépendant de l'homme et qu'il faut s'attacher à découvrir. En s'inspirant de Kant, pourrait-on introduire, comme le propose Mme Miora Mugur-*

**Schächter(2), une méthodologie qui rappelle en permanence que celui qui parle ne décrit pas le réel mais une interaction entre lui et un environnement qu'il ne peut pas qualifier en termes transcendants ?**

**MB** : C'est effectivement la question. Ce paradoxe dont vous parlez est dû au langage lui-même, qui est un langage référentiel : il tend à faire (ou apparaît faire) référence à quelque chose qui est censé lui être complètement extérieur. En fait, la difficulté naît non pas quand on utilise ce langage, mais quand on le prend trop au sérieux, c'est-à-dire lorsqu'on transforme ce que certains appellent le réalisme naturel ou spontané, propre au langage et au mode d'expression courant des scientifiques, en un réalisme métaphysique. Pour le réalisme métaphysique, les termes qu'emploient les scientifiques, les théories scientifiques, sont des réalités complètement indépendantes des manières dont on les a acquises.

Un premier point important est à noter. Il se trouve que déjà à l'époque de Kant, on se rendait parfaitement compte au vu de la physique newtonienne que cette physique ne pouvait pas être descriptive de quelque chose qui aurait été complètement indépendant de tout moyen d'y avoir accès. Cette nouvelle vision de la physique s'opposait à celle, d'inspiration cartésienne, que défendaient encore les philosophes français de l'époque. La physique newtonienne comme la concevait Newton lui-même, ne faisait pas d'hypothèses visant à expliquer l'ordre des phénomènes par des réalités sous-jacentes conçues sur le modèle de "figures et mouvements" ("Hypothèses non fingo", écrivait Newton). Cette physique se contentait de construire des modèles mathématiques des phénomènes. Par exemple, au lieu de chercher à expliquer l'interaction gravitationnelle par des tourbillons, Newton et les newtoniens s'en tenaient à l'énoncé de la loi en inverse du carré de la distance, et à la démonstration de sa capacité à servir de règle à tous les phénomènes astronomiques et terrestres connus. Kant l'a remarqué, Hume d'ailleurs avant lui. Ces deux philosophes en ont conçu des systèmes de pensée extrêmement critiques à l'égard du réalisme en général.

Mais l'intérêt supplémentaire de leurs philosophies est qu'elles montraient l'une et l'autre que dans certaines conditions l'on risquait d'oublier que la physique classique décrivait uniquement un ordre mathématique des phénomènes, et non pas un ordre qui irait au-delà des phénomènes et qui expliquerait ceux-ci. Par exemple, Hume remarquait que lorsque le scientifique se livre à ce qu'il appelait "*carelessness and inattention*" (le manque de soin et l'inattention), il peut très bien perdre de vue qu'il n'a affaire qu'à des phénomènes, toujours transitoires, devant simplement être liés par des lois. Il peut penser ou laisser penser qu'il décrit des objets intrinsèquement existants, transcendants - au sens de ce qui dépasse l'expérience.

De même Kant a construit tout un système pour montrer que l'ordre des phénomènes en question se comporte à tel point "*comme si*" il relevait d'une manifestation d'objets dépassant les phénomènes que la réserve du "*comme si*" risque à tout moment d'être oubliée. A partir de là, on est tenté de croire intégralement au réalisme métaphysique et de s'exprimer en conséquence. Et c'est ce qui s'est passé. Après Kant, au fur et à mesure que le 19e siècle a progressé, et malgré quelques philosophes néo-kantiens dispersés, les physiciens ont complètement oublié les mises en garde de Kant et ont considéré que la physique classique décrivait la réalité telle quelle était "en soi".

Mais au XXe siècle est survenue la nouveauté de la physique quantique. Celle-ci représente peut-être par elle-même cette méthode dont vous parliez qui permettrait de ne pas oublier que ce à quoi l'on a affaire n'est pas transcendant, mais constitue le produit d'une interaction avec notre appareil cognitif. En fait, la physique quantique a rendu totalement impossible le parachèvement du "*comme si*" kantien. A chaque fois que l'on essaye de reconstituer des objets spatio-temporels (ou "corps matériels") à partir de l'ordre des phénomènes, on échoue. Ou on n'y arrive que par parties, par fragments, par fragments complémentaires comme le disait si bien Bohr. Certains fragments manifestés lors de l'utilisation de certains types d'appareillage se comportent comme une onde classique, d'autres comme une particule classique. Mais ce ne sont jamais que des parties de phénomènes relatives à des appareillages mutuellement exclusifs. Elles ne permettent pas une recombinaison, une synthèse qui désignerait de façon non ambivalente une réalité extérieure complètement détachée des moyens de l'appréhension des phénomènes.

*AI : On peut estimer que les physiciens l'ont compris aujourd'hui... Sans cela, comment pourraient-ils faire avancer leurs travaux. Mais pratiquement, aucun autre scientifique, dans les disciplines de la science macroscopique, n'a transposé ces considérations aux descriptions du monde auxquelles ils se livrent dans leur discipline. C'est notamment le cas en physique des particules ou en cosmologie, qui sont pourtant très imbibées de physique quantique. On parle du boson de Higgs ou du multivers comme s'il s'agissait de choses existant en dehors de l'homme et de ses systèmes de connaissance.*

*Enfin, et c'est plus grave, la société laisse toujours la parole aux autoritarismes de toutes sortes, qui prétendent s'appuyer sur leur connaissance prétendue de ce qu'est le monde en soi pour nous dicter nos croyances et nos comportements - je ne parle pas seulement des tenants des religions dites révélées, mais des experts de toutes origines, se disant scientifiques, qui s'expriment constamment dans les médias, et influencent les gouvernements.*

*Il y a là une sorte de défaillance de la raison collective qui conduit à penser que les travaux de philosophes comme vous ne sont pas assez connus, notamment au niveau de la théorie politique dont s'inspirent nos sociétés. Il y a encore un très grand pas à faire, mais comment le faire ?*

**MB :** Oui, effectivement. Vous soulevez un problème important. La vraie raison pour laquelle la base philosophique de la mécanique quantique, bien connue pourtant depuis l'invention de celle-ci, ne se propage pas davantage, c'est qu'il y a une sorte d'espoir, toujours vivace, que l'enseignement de la physique quantique n'a pas à être pris au sérieux. La physique quantique ne serait qu'une théorie provisoire. Un jour, on découvrira des "variables cachées", selon le terme consacré, qui devraient permettre de penser autrement, c'est-à-dire en revenant au réalisme.

C'est absurde. Il est totalement inapproprié de considérer comme transitoire la leçon d'une théorie aussi durable, aussi efficace, aussi féconde que l'est la physique quantique sous prétexte que, peut-être, un jour, celle-ci changera. Elle changera certes, comme toute théorie, mais la probabilité la plus grande à mon sens est que l'étrangeté de la future théorie qui englobera et dépassera la physique quantique sera non seulement aussi grande mais bien supérieure. Au lieu de tenter sans cesse de reculer, il vaudrait mieux aller jusqu'au bout, dire "Arrêtons de regarder en arrière, prenons la physique quantique à la lettre et regardons ce qu'elle nous enseigne sur la nature de la connaissance scientifique en général, et plus particulièrement de la connaissance en physique". A partir de là, on peut progresser.

Un autre facteur qui explique la persistance sociologique du réalisme est que la familiarité avec la physique quantique (superficiellement comprise) devient telle que l'on finit par ne plus se poser la question de la nature et de l'usage de ses symboles. En prenant trop au sérieux le langage que l'on surimpose à ces symboles, on finit par se dire que le langage dit vrai. Et on omet d'étudier soigneusement le fonctionnement des symboles. Ici encore, on tend malencontreusement à attacher plus d'importance à ce que disent les scientifiques qu'à ce qu'ils font dans leurs manipulations symboliques et expérimentales. Pire: les scientifiques eux-mêmes finissent par se laisser fasciner par leur propre rhétorique au lieu de rester axés sur leur pratique. Par exemple, on oublie que le vecteur d'état(3) n'est utilisé que comme moyen de calcul de probabilités d'événements expérimentaux à travers l'algorithme de Born. On dit (en utilisant un vocabulaire critiquable dû à Dirac) : "ceci est l'état d'une particule". Mais on perd ainsi de vue que le concept de particule est complètement éclaté à cause de l'absence de possibilités d'attribuer des propriétés à ces objets comme aussi à cause de l'impossibilité de leur attribuer une identité permanente à travers le temps. Et on omet aussi de soumettre à l'examen le concept d'état qui n'a presque plus rien à voir(4) avec celui d'état classique. On se laisse donc abuser par le langage que l'on surimpose au formalisme. Ceci pose effectivement la question de savoir quel genre d'événement sociologique pourrait faire que les scientifiques changent d'avis à ce sujet et se laissent complètement imprégner par les méthodologies de la physique quantique.

*AI : Qu'en pensez-vous ?*

**MB** : Je pense qu'en fait c'est en train d'arriver. Là je crois pouvoir vous apporter des nouvelles fraîches. Depuis deux ou trois ans beaucoup de physiciens viennent me voir en me disant que les nouveaux procédés de redérivation de la physique quantique à partir de postulats relevant de la théorie de l'information sont exactement en phase avec le type de philosophie que je défend depuis longtemps.

**AI** : *Pouvez vous préciser ce que vous entendez là ?*

**MB** : La particularité de la théorie de l'information est que les bits d'information ne sont ni d'ordre objectif ni d'ordre subjectif. Ce sont des entités qui relèvent déjà d'une interaction accomplie entre le sujet connaissant et le milieu qu'il explore. Tous les axiomes de la physique quantique, qui étaient posés habituellement d'une façon un peu arbitraire et dont la seule justification était a posteriori, c'est-à-dire consécutive à leur corroboration expérimentale, peuvent maintenant être redéduits à partir d'axiomes sur les limites de l'accès à un certain nombre de bits d'information propres à chaque système. L'analyse des limites de l'accès à l'information nous permet de redériver des choses compliquées qui relèvent par exemple de l'équation de Schrödinger, des espaces de Hilbert, de tout ce qui fait la particularité de la physique quantique.

Une fois que l'on a compris que des axiomes portant sur des entités "interfaciales" (plutôt que sur des entités qui soient ou bien objectives ou bien subjectives) permettent de faire tant de choses, on se convainc plus facilement que la physique quantique est peut-être non une théorie du monde que l'on explore, mais une théorie du domaine d'interaction, de la surface qui sépare sujet et objet. L'interface serait alors première par rapport au couple sujet-objet, si bien qu'à partir d'elle on pourrait dans certaines circonstances (et on ne pourrait pas dans d'autres circonstances) distinguer entre le quelque chose qui revient au sujet et le quelque chose qui revient à l'objet.

**AI** : *Est-ce que vous ne dites pas ici, autrement que nous le faisons dans les articles de notre revue, que les progrès qui sont faits dans la réalisation de systèmes tels que les robots autonomes et leurs constructions émergentes vont nous conduire à rapprocher de celles-ci les émergences scientifiques humaines. Les roboticiens montrent que des robots autonomes en interaction darwinienne en arrivent à élaborer des conventions de type langage puis ensuite, des contenus de connaissance langagiers ou comportementaux. On pourrait imaginer qu'un jour, vraiment laissés à eux-mêmes, ces robots finiraient par élaborer des descriptions scientifiques du monde que nous pourrions ne pas reconnaître comme telles, de même que nous ne reconnâtrions pas les manifestations éventuelles d'extra-terrestres qui seraient parvenues jusqu'à nous - mais qui seraient néanmoins aussi pertinentes que les nôtres.*

**MB**. Il y a un lien incontestable entre la robotique, les sciences cognitives et ce que j'ai dit de la physique quantique. Ce lien est apparu à travers les déboires du cognitivisme classique selon lequel les machines agissent de façon programmée et entièrement dépendante de ce que les concepteurs ont voulu y mettre. Confronté à ces déboires, on s'est mis de plus en plus à envisager la cognition artificielle sur le modèle de la cognition naturelle, c'est-à-dire sur le modèle non pas d'une reconnaissance d'un monde déjà pré-engrammé, mais de la construction d'un monde à partir de l'interaction d'une entité connaissante et de son environnement. Cette interaction produit des connaissances qui n'étaient pas préexistantes et qui finissent progressivement par s'autosélectionner. Cette auto-sélection dépend elle-même de l'organisation rétroactive de l'environnement par l'activité de l'entité qui l'explore. Il ne s'agit pas de considérer que l'environnement est déjà tout constitué et qu'il suffit de l'identifier, il s'agit de montrer que l'entité cognitive parvient à façonner l'environnement où elle agit et survit.

**AI** : *Vous retrouvez ainsi la construction de niche, concept que l'on utilise en biologie évolutionnaire(5).*

**MB** : Oui, tout à fait. Si on procède de cette façon pour la robotique, le modèle dominant de la cognition doit abandonner le modèle classique en termes d'input-output (la cognition provenant de la réception d'informations venant d'un monde déjà constitué et provoquant une action à partir de cette réception). Du fait qu'il n'existe plus de schèmes pré-engrammés, on doit adopter une conception entièrement interactive de la cognition, qui est parfaitement isomorphe à celle à laquelle on arrive tout naturellement quand on réfléchit sans préjugés aux processus de la physique quantique. On retrouve ainsi des convergences absolument indubitables entre les deux

directions de recherche relatives à la philosophie de la connaissance : la direction de recherche par analyse de modèles artificiels de cognition et la direction de recherche par analyse de sciences déjà constituées et révélant par leurs propres structures ce qu'elles ne sont pas ou ne sont plus.

**AI : Tenez-vous compte des travaux de la cognition artificielle dans votre propre recherche ?**

**MB :** J'ai travaillé avec Francisco Varela. Nous avons beaucoup discuté de thèmes voisins : philosophie de la cognition artificielle, auto-poïèse... Mais je continue pour ma part dans le domaine de la physique quantique, d'autant plus que j'y vois sans cesse des nouveautés. Par exemple ce dont je vous parlais tout à l'heure en évoquant la théorie quantique de l'information. Celle-ci se développe rapidement à cause de la contrainte technique que vous avez citée, la miniaturisation des ordinateurs qui nous conduit à la barrière de fait et conceptuelle entre le domaine classique et le domaine quantique. Ceux qui ont élaboré la théorie quantique de l'information se sont rendu compte que toute la théorie quantique était dérivable de là et que cela avait des implications philosophiques. Voici l'exemple d'un progrès récent et très important.

Je voudrais aussi évoquer dans mes prochains travaux la gravitation quantique(6) et toutes les théories sophistiquées qui s'y rattachent : super-cordes, géométries non commutatives et autres variantes intéressantes et moins connues. L'une des grandes nouveautés que j'ai vu apparaître récemment est le principe holographique, qui est vous le savez un principe général dont toutes les théories capables de prévoir les effets de gravitation quantique sont dérivables. Le principe énonce (en gros!) que toute l'information sur le volume des entités est contenue dans leur surface, cette surface d'interaction dont nous parlions tout-à-l'heure(7).

**AI : Vous évoquez ici ces questions, notamment la gravitation quantique, comme s'il ne s'agissait pas seulement de constructions théoriques encore inaccessibles aux vérifications expérimentales. Pourtant, on ne peut pas les comparer aux modèles construits à partir d'observations (ou interactions) instrumentales. N'est-on pas là plus proches de la science-fiction que de la science ? L'hologramme, je vois ce que c'est. Mais appliquer l'hologramme à l'univers entier, c'est une autre affaire. De même, on voit ce qu'est la réalité virtuelle, mais peut-on faire l'hypothèse, comme certains auteurs, que nous serions des avatars dans un monde virtuel construit par des intelligences supérieures. Sommes-nous encore là dans le champ constructiviste de la science ?**

**MB :** Effectivement, l'aspect spéculatif de tout cela est considérable. Les meilleurs théoriciens le disent : la base empirique est étroite par rapport à l'immense étendue de leurs sauts formellement spéculatifs. Mais un philosophe des sciences qui a l'expérience de ces questions ne peut pas s'empêcher de penser que les théories considérées comme standard sont déjà elles-mêmes très sous-déterminées par l'expérience. Voilà maintenant de nouvelles théories qui surgissent, dont la gamme de possible est très grande, et qui reposent sur des tests étroits, peu variés. Mais pourquoi ne pourraient-elles acquérir le même statut que les autres, c'est-à-dire se voir reconnues de plus en plus largement, en dépit de leur sous-détermination, et au nom de leur capacité à servir de guide à de nouvelles explorations ?

**AI : Voulez-vous dire qu'au fond elles contribueront, comme les autres, à construire un cosmos résultant de l'interaction de nos systèmes biologiques, cognitifs et computationnels avec un univers indéterminé riche de tous les possibles mais dont rien ne pourrait sortir s'il n'y avait pas cette interaction. Ce serait une vue résolument optimiste de notre devenir, un peu analogue à celle qu'avait proposée il y a quelques années David Deutsch ?(8)**

**MB :** Pourquoi ne pas être en effet un peu optimiste et ambitieux dans ce domaine du renouvellement de la pensée?

**AI : J'aimerais ajouter deux questions aux précédentes, si vous voulez bien. La première question concerne vos projets. Envisagez-vous une suite à vos trois précédents ouvrages, compte-tenu notamment des évolutions qui se produisent dans le domaine de la physique ? Cette suite nous intéresserait certainement beaucoup.**

**MB** : J'envisage effectivement un nouvel ouvrage, mais j'ai encore beaucoup de travail à y consacrer. Il ne paraîtra donc pas tout de suite. Je compte y développer, notamment, les idées d'un article récent que j'ai publié dans un ouvrage collectif : M. Bitbol, " Relations et corrélations en physique quantique ", in: M. Crozon & Y. Sacquin (eds.), Un siècle de quanta, EDP sciences, 2003.

**AI** : *Et enfin, ma toute dernière question : vous vous intéressez au bouddhisme comme théorie ou pratique de connaissance. Pourquoi ? Je suppose que vous n'y voyez pas un substitut aux méthodes de la recherche expérimentale occidentale ?*

**MB** : Non, pas du tout. Mais je suis fasciné par la façon dont le bouddhisme est philosophiquement et "existentiellement" sceptique vis-à-vis de la réification des croyances qui nous servent à nous orienter dans le monde de l'expérience. J'y vois une médiation intéressante vers la conception réflexive et critique du monde qui s'avère si féconde en physique quantique. Le réalisme métaphysique dont nous avons parlé tout au long de ces entretiens, et qui pèse encore lourdement sur la science et la pensée occidentale à la suite de l'héritage de Parménide et de Platon, s'atténue. En nous ouvrant à d'autres cultures, nous devenons insensiblement plus réceptifs à des conceptions du cosmos et de l'homme proches de celles qui sont dorénavant nécessaires pour que se développent les sciences de demain. Y a-t-il seulement là une coïncidence?

### Notes

(1) Sur la Théorie quantique de l'information, on peut regarder

<http://www.ulb.ac.be/rech/inventaire/projets/9/PR2159.html>

(2) Voir [notre article sur Mioara Mugur-Schächter](#)

(3) " Vecteur d'état : entité mathématique par laquelle on représente l'état physique d'un système quantique. En vertu du principe de superposition, les vecteurs d'état ont la propriété de pouvoir s'ajouter entre eux : la somme de deux vecteurs d'état possibles d'un système est aussi un vecteur d'état possible du système ". Etienne Klein, *Petit voyage dans le monde des quanta*, Champs-Flammarion 2004

(4)...sauf passage à la limite quand la constante de Planck tend vers zéro (note de MB)

(5) Sur la construction de niches, voir <http://www.automatesintelligents.com/biblionet/2003/nov/niche.html>

(6) Sur la gravitation quantique, voir <http://www.automatesintelligents.com/biblionet/2002/oct/smolin.html>

(7) On trouvera sur le web une présentation rapide mais suffisante de l'hologramme et du principe hologrammatique à <http://www.cvconseils.com/laser.html> (où la science va-t-elle se nicher?)

(8) Sur David Deutsch, voir <http://www.automatesintelligents.com/biblionet/2004/jan/deutsch.html>

# *Mathématiques, science, culture et philosophie*

PAR Jean6Michel MALDAME , O P

- [I. Le Nombre et le sacré](#)
- [II. Construire le réel, la contribution des mathématiques aux techniques et aux sciences naturelles](#)
- [III. Logique des mathématiques](#)
- [Bibliographie](#)

Mathématique, science de la nature et philosophie se partagent le champ du savoir et structurent de fait l'esprit humain : toutes trois ont une prétention l'universel et reconnaissent cependant que leurs méthodes ne leur permettent pas d'être exhaustives. Toutefois, c'est conjointement dans la reconnaissance de leur visée d'universalité et de leurs limites inhérentes que peuvent se comprendre leurs relations. On peut en déceler la trace dans l'antique désir de sagesse, " savoir de toute chose dans la simplicité ".. Le " sage " en effet sait voir ce qui advient dans le réseau d'une explication qui perçoit l'ordre dans le désordre, l'unité dans la multiplicité, la certitude dans les variations et les confusions. Dans cette visée de la "sagesse", les philosophes et les mathématiciens se rencontrent. Connaître les choses mêmes implique que l'on reconnaisse les mathématiques comme un instrument privilégié pour l'intelligibilité des structures des existants concrets. L'histoire de la pensée montre une relation étroite entre la connaissance du monde et le savoir mathématique et que cette relation ne se laisse pas réduire à un partage des domaines dans lequel les mathématiques s'occuperaient du mesurable et la philosophie du subjectif ou du sentiment. Il y a entre l'une et l'autre une étroite corrélation.

## **I. Le Nombre et le sacré**

Lorsque des mathématiciens aussi différents que Jacques Hadamard, Henri Poincaré, Claude Bruter ou Jacques Roubaud rendent raison du chemin de leur invention et de leur passion pour le travail mathématique, ils se réfèrent à une intuition fondatrice qu'ils expriment en termes d'esthétique.

Dans un monde plus religieux, la référence à la créativité de l'esprit était faite à la catégorie du sacré, que nous entendons ici comme catégorie anthropologique fondamentale, dégagée par les analyses phénoménologiques de Mircea Eliade, James Frazer, Gérard Van der Leeuw, Michel Meslin, G.E.R. Lloyd, qui montrent que pour un esprit religieux l'espace n'est ni homogène, ni isotrope. Il est orienté selon une direction privilégiée (axis mundi) ou estimé à partir d'un centre. Le lieu sacré est toujours lié à une forme pure rattachée à un archétype divin. Dans la simplicité de la forme (cercle, carré, coupole, clocher ou minaret verticaux,...) ou dans la complexité (labyrinthe), le lieu sacré est censé reproduire sur la terre un modèle céleste. La perfection de la forme sacralisée se prolonge en perfection du chiffre qui scande la proportion, le rythme, et préside à l'ordre qui se déploie dans le bâtiment.

Pour cette raison l'homme religieux a toujours été attiré vers la connaissance de la forme et du nombre. Aussi, il n'est pas surprenant que les commencements du savoir mathématique soient mêlés aux systèmes religieux. Le prêtre égyptien, l'astronome chaldéen ou, avant eux, le scribe sumérien justifiaient la légitimité de leur usage des chiffres et des figures en termes de révélation. Les chiffres et les figures alimentent toujours quelques mythes qui réapparaissent en cette époque dite de postmodernité, par exemple chez les apologistes de la connaissance symbolique marquée par la philosophie de Jung.

Cette situation ne saurait mener à conclure à une identité, car les mathématiques n'usent pas d'un langage religieux. En effet, la pensée religieuse est une pensée de la participation, tandis que la pensée mathématique se veut rigoureuse par la distance qu'exige l'objectivité du savoir. Le terme de symbole est équivoque quand il qualifie le comportement religieux ou le langage mathématique. Le premier vit en osmose et en fusion avec les forces portées par le corps et l'esprit ; le second est fondé sur un travail d'abstraction qui forme un objet qui se tient à distance, non seulement de la perception sensible ou de l'affectivité, mais de toute volonté fusionnelle. Aussi le développement des mathématiques est-il lié à une mise à distance de toute sacralité.

Pythagore est paradigmatique de ce moment fondateur. D'une part, ils présente les chiffres et les proportions avec l'aura du divin. Mais, d'autre part, il y a un souci d'universalité et de généralité qui arrache l'objet à sa matérialité. On retrouve un cheminement analogue au début de la science classique, lorsque Kepler cherche à décrire l'harmonie universelle dans *Le Secret du monde* ; de même lorsque Bolzano étudie l'infini, en référence avec le concept théologique d'infini parfait, hérité de la philosophie chrétienne, il témoigne d'un rapport entre les mathématiques et la vision religieuse du monde. En effet, le terme d'infini (le non-fini), qui était lié à l'imperfection du contingent et du devenir, est devenu un attribut divin, désignant une perfection absolue - ce qui a orienté l'esprit vers l'infini actuel. Aussi, de Nicolas de Cues et Giordano Bruno à Leibniz et Cantor, cette notion n'a cessé d'être à l'articulation des questions posées par les rapports du local et du global, du physique et du métaphysique ; tant pour l'analyse que pour la théorie des nombres. De même la sacralité du chiffre a été investie dans la virtuosité d'esprit de la Kabbale et des traditions ésotériques ou romantiques, pour le meilleur et pour le pire.

Or le progrès des mathématiques n'a cessé d'inviter à récuser les équivoques entre des ordres de savoir différents, l'objet mathématique excluant toute référence à une transcendance absolue. Aussi, à la fin du XIXe siècle, l'apologétique de Cauchy parut bien naïve, puisque les concepts mathématiques d'infini et d'unité au principe du nombre ne sont en rien les attributs de perfection du Dieu de la Bible. La psychologie génétique, développée par Jean Piaget et ses disciples, montre comment le progrès de l'esprit consiste à passer des perceptions globales et fusionnelles de la causalité préformelle au stade où l'esprit est clair, conscient de ses pouvoirs et de ses limites. Le passage d'un stade à un autre n'est pas de simple substitution, car l'esprit se structure de manière globale et la réorganisation de la pensée n'abolit jamais les stades antérieurs. Aussi, une dimension symbolique reste matricielle pour la pensée mathématique. Elle apparaît dans les moments de création et d'invention. Même si elle est cachée par le labeur de l'écriture rigoureuse, elle demeure comme un fondement non aboli. L'intuition ne cesse pas d'être présente aux principes et aux fondements, qui sont tout autre chose qu'un préalable.

Dans une autre perspective, Gaston Bachelard a montré la permanence d'archétypes qui soutiennent les constructions rationnelles de la science et donc habitent l'élaboration des formes pures, objet des mathématiques.

Les travaux épistémologiques ayant montré que la référence au sacré est inutile, pour thématiser et justifier la créativité de l'esprit, l'attention s'est orientée vers l'esthétique. La beauté de l'objet mathématique et la fascination qu'exerce le travail du mathématicien introduit un élargissement de la raison. Les termes de simplicité, pureté, harmonie, fécondité, finesse, pénétration, montrent comment les mathématiques sont un des lieux majeurs du développement de l'esthétique, au sens kantien du terme. En même temps qu'une réflexion systématique sur l'esthétique ne saurait ignorer les mathématiques, l'épistémologie ne saurait mésestimer le rôle de la beauté tant pour la construction que pour la présentation des travaux mathématiques. Il ne manque pas aujourd'hui d'essais pour produire des oeuvres d'art en suivant des algorithmes inspirés des fractals qui retrouvent la beauté des formes de la nature. Si ceux-ci ont été connus par un vaste public, c'est à cause de la "beauté des fractals" formalisés par Benoit Mandelbrot en référence aux grandes réalités de la nature - des galaxies à la forme des végétaux ou des minéraux. Le travail du mathématicien rejoint celui du poète, comme le fait Roger Caillois écrivant la beauté des pierres et des formes de la vie en même temps qu'il réfléchit sur la symétrie et les formes abstraites présentées par la géométrie élémentaire. Il en va de même, pour les attracteurs étranges, ainsi nommés parce que par itération d'une opération simple, quelle que soit la position de départ, le comportement à long terme mène à un point fixe et, si l'opération est plus complexe, on construit une forme géométrique. De telles opérations permettent de construire des arborescences et des figures qui correspondent aux créations des grands artistes, comme l'a montré, entre autres, Roland Fivaz. Aussi l'esthétique est-elle une zone frontière où mathématiques et philosophies se rencontrent. Une telle rencontre invite à poser la question des rapports entre mathématiques et connaissance de la nature, puisque la simplicité ou l'exubérance des formes de la nature s'accordent à ce qui peut être formalisé et construit par les mathématiques.

L'aspect esthétique ne concerne pas seulement les objets mathématiques, elle porte également sur le raisonnement mathématique. Pour dire la qualité d'une démonstration, on parle de son élégance ou de la

séduction qu'elle opère sur l'esprit. L'enchaînement des concepts et des propositions est d'autant plus justifié qu'il procède avec simplicité et sobriété.

## **II. Construire le réel, la contribution des mathématiques aux techniques et aux sciences naturelles**

Si les mathématiques se construisent en lien avec le mode symbolique de penser, elles entretiennent un rapport plus riche encore avec les sciences de la nature.

Depuis les origines de la culture, le travail du mathématicien a été en lien avec la maîtrise de l'espace par le cadastre ou la géographie et avec la maîtrise du temps par détermination du calendrier ou le comput astronomique. Les mathématiques ont pour point de départ les demandes de l'artisan, de l'architecte ou de l'urbaniste : mesure des distances, des aires, des angles et des circonférences. Les problèmes de construction ou de calcul appellent, aujourd'hui comme hier, la compétence mathématique. Que seraient la physique théorique sans langage mathématique ? Que serait aujourd'hui la biologie sans modèles mathématiques ? Que seraient les sciences informatiques sans outils mathématiques ? Corrélativement, le vocabulaire des mathématiques garde la marque de ce lien étroit, puisqu'on parle de construction, de fondement et de développement et que les concepts de la topologie empruntent encore le langage des artisans, comme le montrent les travaux de René Thom. Cette intrication est telle qu'il est hélas commun qu'on réduise les mathématiques à être un instrument d'exploration de la nature, dont elles seraient le décalque théorique et idéal.

C'est à cause de cette intrication que les renouvellements dus à l'approfondissement et à l'extension du domaine mathématique ont suscité de grands débats. A la fin du XIX<sup>e</sup> siècle, l'élaboration des géométries non euclidiennes a amené à remettre en cause l'axiomatique présidant à la construction des objets mathématiques habituels, référés aux travaux d'Euclide, pour représenter la réalité ; ce qui a déchiré une fois encore le sens commun, structuré par l'apprentissage scolaire de la géométrie euclidienne ! Il est apparu clairement que les mathématiques ne sont pas une partie des sciences de la nature ; elles se développent pour elles-mêmes, indépendamment des questions posées par celles-ci. Les objets mathématiques ont un statut spécifique ; ils sont irréductibles à un rôle ancillaire d'instrument d'exploration et de maîtrise de la nature. Les arpenteurs savaient mesurer pratiquement la diagonale du carré et si les mathématiciens s'étaient contentés de cet usage, ils n'auraient jamais été plus avant. Les physiciens et les ingénieurs savaient résoudre bien des problèmes, et si le seul motif de la recherche était l'économie de moyen, les mathématiciens n'auraient pas eu besoin de construire rigoureusement une théorie des fonctions. De même aujourd'hui, le développement des mathématiques est indépendant de l'utilisation de leur rôle par les biologistes, les physiciens ou les cybernéticiens.

A cause de cette indépendance vis-à-vis des sciences de la nature, le savoir mathématique est autre chose qu'une collection de disciplines -arithmétique, géométrie, analyse,...- utiles pour la connaissance du monde. Il constitue un savoir unifié qui demande à être nommé au singulier : la mathématique ! Le singulier signifie ici une cohérence propre et une démarche spécifique qui se développe dans son ordre pour elle-même. L'emploi du singulier ouvre sur des questions ontologiques.

Pour Pythagore, les nombres sont l'essence même de la réalité physique qui, pour cette raison, constitue un cosmos. Les mathématiques disent la nature des choses, dès que l'on dépasse les apparences confuses. L'irrationalité des nombres que l'on peut construire empêche que l'on tienne ce réalisme pour explication ultime de la réalité. Instruit par l'échec de cette cosmologie, Platon a renoncé à l'adéquation du chiffre et du réel, pour penser le monde en terme de participation. L'ordre des choses sensibles, qui sont données à l'expérience, n'est intelligible que par sa référence à l'idéalité qui est réalisée purement et simplement dans le seul ordre de l'intelligible. L'idéalité mathématique est plus réelle que le monde mouvant et changeant, indéterminé et confus.

Contre cette idéalisation, Aristote a vu dans les mathématiques le résultat d'une abstraction, qui de ce fait s'éloigne de la réalité et donc ne saurait être comprise comme source première d'intelligibilité, car le Stagiritte privilégie une philosophie de la nature dont les concepts proviennent de l'étude du devenir par une élaboration qui ne se limite pas au seul quantitatif, mais accorde une prééminence au qualitatif. L'opposition entre platonisme et aristotélisme s'enracine dans l'estime ou le refus du savoir mathématique. Une fois encore, les

mathématiques influent sur les philosophies qui, en retour, par la place qu'elles leur accordent, influent sur leur développement et leur usage tant pour la formation de l'esprit que pour la maîtrise technique du donné matériel ou social.

Lorsque la Renaissance s'est libérée de l'influence de l'aristotélisme scolaire, elle a demandé à des thèmes platoniciens de lui apprendre à regarder le monde d'un regard libre et dominateur. Une nouvelle philosophie a commencé, fondée sur la conviction que les mathématiques étaient l'instrument privilégié pour connaître le réel. Non seulement, des physiciens comme Galilée ou Newton, mais surtout les philosophes prenant pour modèle du savoir, l'*organon* mathématique au développement duquel ils participaient. Descartes joue un rôle exemplaire dans ce domaine. La notion de *mathesis universalis* exprime l'horizon de cet idéal du savoir dont le raisonnement et la méthode mathématiques sont le paradigme. La mathématique est alors la clef de l'intelligibilité de toute chose et le modèle de toute quête du vrai. Les grandes constructions de l'âge classique, celles de Spinoza ou de Leibniz, utilisent largement cette perspective. La *mathesis universalis* définit un horizon du savoir où le réel peut devenir transparent grâce à la méthode mathématique.

Le progrès des mathématiques au seuil du XXe siècle et la crise des fondements qui en est résultée ont mené les mathématiciens à tenir à distance les confusions possibles entre la métaphysique et les mathématiques, en se gardant de donner une dimension ontologique à leurs concepts. Les termes qui semblaient décrire la réalité - point, ligne, cercle, droite, surface, volume,...- ont été perçus dans leur idéalité. Les termes qui avaient un aspect métaphysique -infini, un, continu,...- ont été réduits à leur dimension opératoire dans le champ du savoir mathématique. Les concepts mathématiques sont le fruit d'une activité de l'esprit, par une démarche d'abstraction et de construction. L'essentiel est l'activité de l'esprit ; comme celle-ci n'est pas arbitraire, elle peut rendre raison de la réalité.

A l'âge classique, la mathématique était essentiellement algèbre et géométrie. A l'âge moderne, la théorie des fonctions et du calcul différentiel et intégral permettaient de décrire les phénomènes de la nature. Aujourd'hui, elle a construit de nouveaux instruments dont la fécondité étonne, car ils permettent de formaliser des transformations qui échappaient à la mécanique rationnelle. La biologie use désormais de modèles qui proviennent des mathématiques, en particulier de la théorie des catastrophes de René Thom qui, venue de la topologie la plus abstraite, peut s'appliquer à l'étude des organismes vivants. L'étude des populations et de leur évolution -et donc corrélativement celle de l'origine des espèces et de la vie même- puise dans les outils mathématiques qui proviennent des probabilités et des lois stochastiques. La géométrie fractale donne de la réalité une description plus fine que celle qui était obtenue à la règle et au compas. La théorie du chaos, développant certaines intuitions de Henri Poincaré, l'étude des systèmes non linéaires par Mitchell Feigenbaum, ont permis de donner une représentation des phénomènes les plus complexes de la nature. De même, les attracteurs étranges permettent d'expliquer des stabilités et des morphogénèses.

Ainsi la mathématique s'affirme-t-elle dans une indépendance souveraine ; indépendance, en effet, parce qu'elle ne se développe pas dans le souci de répondre à des questions qui lui seraient étrangères ; souveraineté, parce qu'elle ne cesse de donner des instruments d'analyse de la réalité qui dépassent même les espérances des disciplines qui usent de ses résultats.

Cette possibilité de rendre raison avec simplicité du réel est due à la nature de l'objet mathématique, fruit d'une activité de l'esprit. Le premier rapport au monde est dans le sensible où l'esprit est investi. La première appréhension est riche, mais encombrée et paralysée. Pour progresser, l'esprit doit procéder par séparation. Il ne retient que certains aspects et par là construit un objet spécifique. L'objet est donc le fruit d'une activité de la raison et leur donne un statut particulier. Ce n'est pas une perception passive, mais un jugement. La mathématique est le fruit d'une décision de l'esprit qui construit un objet qui n'a plus besoin d'être immédiatement référé à un donné extra-mental (comme dans les sciences de la nature). Il se développe pour lui-même par fécondité naturelle mue par un désir d'universalité et d'unité, en répondant à des problèmes nouveaux dans des domaines nécessairement spécialisés. Le tableau donné par Jean Dieudonné dans son *Panorama des mathématiques pures* expose les thèmes de recherche des mathématiciens et montre de façon synthétique à la fois la très grande abstraction de ces recherches, leur indépendance vis-à-vis de toute application extérieure et

leur ouverture sur des utilisations par d'autres savoirs. Si les sciences de la nature peuvent user du résultat de ce travail de la raison, elles construisent un objet propre qui ne procède pas de la même méthode d'abstraction.

### III. Logique des mathématiques

Puisque les mathématiques ne sont pas, comme les sciences de la nature, seulement une science du concret, et que leurs objets sont le fruit de l'activité de l'esprit, ne faut-il pas rapprocher mathématiques et logique ?

La démonstration est un aspect de la démarche fondamentale de la raison qui est soumis à la logique. D'où la question : les mathématiques sont-elles une partie de la logique ?

La réponse affirmative à cette question est fondée sur le fait que les mathématiques ont toujours progressé, grâce à une mise en forme rigoureuse. Il ne suffit pas, en effet, au mathématicien de produire des théorèmes et des algorithmes pour résoudre les problèmes qu'il se pose, il lui faut un *organon*. Les *Éléments* d'Euclide ont joué ce rôle ; ils héritaient d'un savoir mathématique diversifié ; ils unifiaient en un tout cohérent un savoir dispersé et pragmatique. Ils mettaient en lumière des liens entre des propositions ayant un statut divers. Par là, ils créaient d'autres questions et relançaient l'esprit à la recherche d'une meilleure intelligibilité de ses objets. Ce travail est une exigence propre aux mathématiques ; il n'a jamais cessé de l'animer. Les travaux de Bourbaki en ont donné une réalisation exemplaire au cours du XXe siècle. Le groupe de mathématiciens rassemblés dans cette école a mis en ordre le savoir mathématique de manière méthodique et cohérente à partir d'une formalisation très générale de la théorie ensembliste de manière à présenter de manière rigoureuse l'enchaînement des concepts et objets mathématiques. La mise en ordre permet de présenter de manière organique les différents champs de la recherche - topologie algébrique et différentielle, variétés différentielles, analyse harmonique commutative ou non commutative, géométrie analytique, géométrie algébrique, théorie des nombres, algèbre homologique, logique mathématique, catégories et faisceaux,.... L'unité et la cohérence d'un tel corpus montre pourquoi les mathématiques réalisent de manière exemplaire un savoir méthodique, fondé sur des principes clairs et distincts, qui fascine l'esprit en quête de vérité. Aussi la mathématique sert de modèle à la pensée rigoureuse, généralisant l'idéal de la *mathesis universalis*. L'existence d'un tel corpus ne saurait clore le progrès des mathématiques ; aussi l'unité réalisée par l'école Bourbaki éclate-t-elle aujourd'hui, comme le montre le chapitre 10 du premier volume de l'Encyclopédie philosophique universelle consacré aux "démarches mathématiques".

Un tel développement et la créativité dont il témoigne, montre que la démonstration et l'unification en un corpus méthodique, ne dit pas tout le travail du mathématicien. Il y a un rôle décisif pour l'intuition et l'imagination et la mise en ordre déductive n'est qu'un moment du développement des mathématiques. A quelque moment que l'on se place dans l'histoire des mathématiques on constate que l'esprit ne saurait rester enfermé à quelque stade du savoir. L'idéal de la *mathesis universalis* a pour corrolaire une incessante ouverture de l'esprit tout à la fois désireux d'unité et d'universalité.

Conformément à cet idéal, il était inévitable que l'on fasse de la logique une branche des mathématiques, en donnant aux règles de la logique une formalisation qui permettait le calcul. Poursuivant les premiers travaux de Leibniz et de Hobbes, la formalisation rigoureuse faite par Boole, puis par Morgan, Pierce et Frege, a donné les éléments fondamentaux pour développer le calcul propositionnel.

A cause de la fécondité de ce programme, des mathématiciens ont cherché les fondements de leur savoir dans la logique. Whitehead et Russell ont voulu fonder les propositions fondamentales des mathématiques non par un rapport de l'esprit à un réel extra-mental, mais aux activités spécifiques de l'esprit. Cherchant les notions les plus générales -en particulier celle de classe- ils ont construit une immense synthèse qui visaient à enserrer les recherches axiomatiques, qui se situaient déjà à l'intérieur des mathématiques (celles de Peano en tout premier lieu), dans une reconstruction fondée sur les notions d'ordre et de classe, répondant aux opérations fondamentale du raisonnement. Une telle mise en oeuvre donnait à la logique le statut de norme universelle du vrai et ouvrait le débat sur les fondements des mathématiques et les diverses écoles qui prirent naissance à partir

des recherches de Bolzano, de Morgan, Boole, Frege, Cantor, Peirce, Dedekind, Russel, Hilbert, Borel, Zermelo, Brouwer et Russell.

L'exigence d'une instance de savoir apte à vérifier la cohérence et la non-contradiction des propositions a donné naissance à ce que l'on appelle la méta-mathématique, savoir qui participe de la logique et des mathématiques et à l'intérieur duquel les travaux de Gödel ont définitivement tranché.

Les mathématiques ouvrent ainsi sur des questions spécifiquement philosophiques. La logique serait-elle donc plutôt une partie des mathématiques ? La réponse aux questions du fondement peut-elle être donnée sur le seul plan de la logique ? La diversité des options entre écoles mathématiques peut-elle être réduite par l'élaboration d'une théorie mathématique plus large que la théorie ensembliste, ou par le jugement extérieur du philosophe ?

L'itinéraire philosophique de Husserl comme celui de Wittgenstein sont alors exemplaires puisque c'est à partir de la construction des mathématiques que l'un et l'autre se sont interrogés sur la valeur de la représentation et du langage humain.

Le philosophe n'a pas de peine à reconnaître que les mathématiques ne peuvent se donner à elles-mêmes leur fondement. L'origine reste inaccessible, puisque nul ne peut être avant d'avoir été... Les mathématiques naissent d'une décision de l'esprit qui n'est pas enfermée dans le formalisme mathématique. Pour autant le formalisme mathématique n'est pas subjectif. Il doit obéir aux règles strictes de la démonstration et de l'exposé logiquement construit qui reflètent une réalité plus profonde, l'ordre des raisons.

Les progrès des mathématiques posent ainsi la question de la vérité : s'il n'y a de vérité que dans la non-contradiction, le vrai est-il réductible ? La norme ultime du vrai est-elle dans la construction d'un système de signes ou se trouve-t-elle dans un donné premier - plus universel et transcendant ?

Le rapport entre la logique et les mathématiques n'est pas seulement délimité par la question des fondements des mathématiques. Il est posé aujourd'hui par l'essor considérable des sciences cognitives. Celles-ci ont été possibles grâce à la formalisation de la logique en calcul propositionnel, mais aussi grâce aux progrès techniques des machines aptes à traiter de l'information et à une meilleure connaissance de l'activité cérébrale humaine. Si personne ne pense que l'expression d'intelligence artificielle ne doit être prise à la lettre, on ne peut pas ne pas prendre acte du fait que les ordinateurs sont capables de mener à leur terme des processus de raisonnement et de calcul : la construction des machines s'inspire du fonctionnement des neurones et, en retour, l'élucidation des problèmes de construction et d'optimisation du fonctionnement des machines permet d'explorer les processus cérébraux. Le développement des sciences cognitives pose donc la question du rapport entre penser et calculer. La question philosophique revient au premier plan, portée par les réalisations où la logique mathématique joue un rôle structurel. On voit donc s'opposer des familles d'esprit dans des philosophies diverses, en fonction de leur jugement sur les réalisations de la raison pratique. Les machines dites "intelligentes" ne sont pas seulement des réalisations isolées, elles s'intègrent dans une vision d'ensemble, la théorie de l'information, due à Léon Brillouin, qui pose la question du langage et donc de ce qui est spécifique à l'homme et de la communauté de pensée qui préside à tout acte d'intelligence.

Les mathématiques et les philosophies entretiennent donc des rapports étroits. En premier lieu, les philosophes ont souvent porté une grande attention aux objets mathématiques (nombres et figures, grandeurs et relation...) et au mode de production de leurs objets (règles de démonstrations, axiomes et théorèmes...). De plus, le renouvellement des mathématiques, qui ne s'est pas fait de manière continue mais, comme tout autre savoir, en passant par des révolutions conceptuelles, suppose une clarification interne dans une conception claire de la spécificité du travail mathématique, constituant pensée interne présidant au développement des connaissances. Il y a enfin le regard externe des philosophes sur les travaux mathématiques qui sont, à leur yeux, une réalisation exemplaire de la vie de l'esprit humain et de la puissance de l'entendement. La prise en compte de l'existence mathématique ouvre sur l'éloge de la pensée humaine, en tous ses éléments, c'est-à-dire sa créativité, l'imagination, l'intuition et la force du raisonnement. La philosophie est ainsi invitée à mieux s'interroger sur les

actes fondamentaux de la connaissance ou à fonder la philosophie de l'action en dialogue avec la rationalité des méthodes mathématiques.

## **Bibliographie**

BACHELARD Gaston, *Essai sur la connaissance approchée*, Paris, Vrin, 1928 ; *La Formation de l'esprit scientifique. Contribution à une psychanalyse de la connaissance objective*, Paris, Vrin, 1969.

BARREAU Hervé, *L'Epistémologie*, Paris, PUF, 1990.

BELAVAL, *Leibniz critique de Descartes*, Paris, Gallimard, 1960.

BOREL Emile, *Valeur pratique et philosophie des probabilités*, Paris, Gauthier-Villars, 1952.

BOUTOT Alain, *L'Invention des formes*, Paris, Odile Jacob, 1993.

BRUNSCHWIGC Léon, *Les Etapes de la philosophie mathématique*, Paris, Felix Alcan, 1912, Albert Blanchard, 1993.

BRUTER Claude Paul, *Sur la nature des mathématiques*, Paris, Gauthier-Villars, 1973.

CAVAILLES Jean, *Philosophie mathématique*, Paris, Hermann, 1962.

DIEUDONNÉ Jean, *Panorama des mathématiques pures. Le choix Bourbachique*, Paris, Gauthier-Villars, 1977; *Pour l'honneur de l'esprit humain. Les mathématiques aujourd'hui*, Paris, Hachette, 1987 ; *Penser les mathématiques*, Paris, le Seuil, 1984.

FIELD Michael et GOLUBITSKY Martin, *La Symétrie du chaos. A la recherche des liens entre mathématiques, art et nature*, Paris, InterEditions, 1993.

FIVAZ Roland, *L'Ordre et la volupté. Essai sur la dynamique esthétique dans les arts et dans les sciences*, Lausanne, Presses polytechniques romandes, 1989.

GONSETH Ferdinand, *Les Mathématiques et la réalité. Essai sur la méthode axiomatique*, Paris, Albert Blanchard, 1974.

PIAGET Jean, *Logique et connaissance scientifique*, *Encycopédie de la Pléiade*, Paris, Gallimard, 1967.

THOM René, *Modèles mathématiques de la morphogénèse*, Paris, Christian Bourgois, 1980 ; *Paraboles et catastrophes. Entretiens sur les mathématiques, la science et la philosophie*, Paris, Flammarion, 1980.

## Sujet épistémique et sujet cognitif

### dans la théorie des modèles mentaux de l'inférence

Shahid Rahman

Les relations entre logique et psychologie ont une vieille histoire touchant à la nature profonde de ces deux disciplines, à la fois quant à leur dépendance mutuelle et quant à leur autonomie.

**Mon exposé évoquera deux types de difficultés méthodologiques soulevées par les relations entre logique et psychologie, dans le contexte de l'approche cognitive du raisonnement développée par Johnson-Laird et ses collaborateurs. Rappelons que cette approche — la théorie des modèles mentaux — est l'une des principales théories cognitives du raisonnement logique, voire la théorie dominante.**

Je reconnais d'emblée que ma position prêtera à contestation car je discuterai la théorie des modèles mentaux en me plaçant uniquement du point de vue de la philosophie de la logique.

Je laisserai de côté une revue historique des passionnantes relations entre psychologie et logique. Je commencerai plutôt par quelques brèves remarques sur l'œuvre de Frege et sur ses postulats platoniciens. D'autres auteurs comme Husserl ou Bolzano auraient pu offrir un point de départ analogue, *mutadis mutandi*, mais je pense que l'œuvre de Frege exerce une influence plus directe sur la théorie des modèles mentaux.

Le point de départ de la philosophie analytique : la logique sans la psychologie

Résumons en termes très simples un débat profond et complexe : Frege a cherché à éliminer la psychologie de la logique parce que, selon lui, la psychologie ferait de la logique une science purement subjective, ce qui priverait la logique de son but le plus cher, à savoir la détermination de lois logiques universelles. Plus précisément, selon Frege, la logique est essentiellement l'étude de la relation de conséquence, et cette relation de conséquence est une relation entre propositions. Et le fait qu'une proposition est une conséquence logique d'un ensemble de prémisses vraies est indépendant d'un sujet quelconque [qu'il soit épistémique, cognitif ou affectif].

Cependant, afin d'expliquer la connaissance des propositions logiques, Frege a besoin de quelque chose comme un "moment" épistémique. Frege utilise alors le concept traditionnel d'acte de jugement pour rendre compte de ce moment où un sujet connaissant entre en relation avec une proposition. Ce moment épistémique est conçu comme un type d'acte par lequel les êtres humains peuvent reconnaître des propositions préexistantes vraies et leurs relations. Frege pense donc que le jugement, au sens propre, a affaire avec l'épistémologie et la psychologie, mais non avec la logique comme telle.

**Autrement dit, sa théorie du jugement n'a pas d'influence sur la logique et sur les propriétés de la relation de conséquence : la logique est "bien comme elle est". Pour Frege, le jugement ne joue un rôle central qu'en épistémologie, dans l'abord des contextes de découverte scientifique, ou bien en psychologie, lorsque nous étudions la performance logique humaine dans ses réussites ou ses échecs.**

Appelons cette théorie frégenne la *logique sans psychologie*. Dans ce qui suit, je vais discuter deux approches allant à l'encontre de cette logique sans psychologie. La première approche est plus

conservative quant au point de vue de Frege, car nous verrons qu'elle accepte l'idée que la logique est "bien comme elle est", et accepte aussi la distinction stricte entre logique d'un côté, psychologie et/ou épistémologie de l'autre. La deuxième approche, qui conteste de façon beaucoup plus radicale le point de vue de Frege, soutient que la théorie du jugement devrait guider l'étude des propriétés logiques des propositions.

J'appellerai la première approche *l'approche gricéenne*, même si je ne m'occuperai pas ici de discuter si cette approche était effectivement celle que Grice avait en tête. La seconde approche sera dite *antiréaliste*. Je vais soutenir que ces deux approches posent deux défis méthodologiques sérieux à la théorie des modèles mentaux.

*Psychologie sans logique ? L'approche gricéenne et le sujet cognitif de la théorie des modèles mentaux.*

Dans l'approche gricéenne, la logique est donc pensée comme "bien comme elle est". Cependant les "usages" de la logique dans le raisonnement quotidien sont déviants relativement aux normes logiques frégréennes, même s'ils ne sont pas purement subjectifs. Les déviations sont expliquées par le recours à une théorie de la communication. Prenons un exemple. Rappelons-nous la table de vérité de la disjonction : A ou B est vraie si et seulement si au moins l'une des deux propositions est vraie. Dans une situation de communication supposée "normale", si je sais que A est vraie, je ne communiquerai pas "A ou B", [mais "A"] même si A ou B est vraie. Je ne le ferai pas car si je veux vraiment communiquer, je ne veux normalement pas diminuer l'information que je transmets. De cette façon, tout contraste entre l'usage de la logique par un sujet qui raisonne et les lois logiques prescrites par Frege est supposé résulter de quelque maxime sous-tendant l'usage dans la communication du schéma logique en jeu.

La théorie des modèles mentaux de Johnson-Laird semble construite dans le même esprit, les lois de la communication étant remplacées par des contraintes cognitives qui sont surtout fonction de la mémoire et du temps. Ici aussi, les anomalies ne sont pas purement subjectives, cependant elles suivent les lois contraignant le raisonnement d'un sujet cognitif. L'inférence, ici comprise comme le raisonnement de ce sujet cognitif, est le résultat de l'application des lois logiques lorsque cette application est restreinte en raison des limitations des ressources psychiques du sujet cognitif. En fait, les maximes de la communication de Grice sont réduites à ces contraintes temps-mémoire.

Les modèles mentaux sont conceptualisés comme des sortes de graphes construits par l'esprit, et, pour la logique propositionnelle, ces modèles mentaux sont très semblables aux représentations spatiales des fameuses tables de vérité.

Prenons l'exemple d'un raisonnement spatial, là où la théorie semble la plus plausible. Si Jacques Dubucs est derrière moi et Daniel Andler derrière lui, cela semble raisonnable de conclure que Daniel Andler est derrière moi.

Selon la théorie des modèles mentaux, en effectuant ce raisonnement, le sujet construit un modèle spatial abstrait ou un graphe plutôt qu'une construction linguistique ou une déduction dans le style d'un système de déduction naturelle.

La principale thèse de Johnson-Laird est que des phénomènes similaires se produisent dans tout type de raisonnement. Dans le cas d'un raisonnement relié à ce que nous appelons la logique propositionnelle, la construction mentale ressemble à une représentation spatiale des tables de vérité des constantes logiques.

Cependant, la construction de ces graphes et le raisonnement qui les utilise demande non seulement beaucoup d'énergie mais aussi du temps et de la mémoire à court terme. En raison de ces contraintes, nous faisons des erreurs en appliquant ces "tables éternelles" de vérité, qui, comme dans le poème d'Horace, paraissent "éternelles autant que l'air et les pyramides d'Égypte". Nous commettons des erreurs parce que la construction mentale d'une table de vérité peut être horriblement longue.

Les résultats empiriques apportés par Ruth Byrne semblent confirmer cette théorie. Cependant, pour la plupart de ces résultats, il est possible de donner une explication différente des performances déviantes, en supposant simplement que les raisonneurs n'admettent pas nécessairement les tables de vérité classiques ou même de quelconques tables de vérité.

Peut-être pensez-vous que je suis injuste. Pourquoi ne pas simplement étendre les constructions mentales à d'autres logiques possibles avec l'aide de quelque méthode de décision appropriée qui serait elle-même conçue comme un graphe mental ? De fait, dans un bref article critique sur les modèles mentaux, Hintikka a suggéré une telle stratégie.

Mais je soutiens justement que ce n'est pas une option disponible pour la théorie des modèles mentaux ou au moins que la théorie nécessiterait d'importants développements si cette option devait être prise en considération. Ce qui importe ici dans la théorie des modèles mentaux, c'est que les ressources cognitives n'ont pas d'influence sur les lois de la logique, comme chez Frege. Cette prémisse explique les performances déviantes. Si nous acceptons plusieurs logiques, alors nous aurions à expliquer la différence entre une performance erronée et un changement de logique. Et pour faire cela, nous aurions besoin de trancher le sujet cognitif en deux pour obtenir un sujet épistémique, responsable des changements de logique, et un sujet psychologique, responsable des performances déviantes.

De plus, et cela est le point central, nous aurions besoin d'expliquer pourquoi nous devrions admettre des logiques différentes si les ressources cognitives n'influencent pas la logique. Même si quelque solution pouvait être découverte, cette solution putative impliquerait une révision radicale des principes fondamentaux de la théorie. Car, philosophiquement parlant, un des défis majeurs de cette révision serait de reprendre à zéro la tâche de différencier la pragmatique de la sémantique. Pour remplir ce défi, il faudrait concevoir une relation beaucoup plus dynamique entre les aspects pragmatiques et sémantiques de la signification (*meaning*). Or, la tradition frégréenne, qui a influencé l'approche gricéenne, et, via cette approche, la théorie des modèles mentaux, fonde son analyse sur une stricte séparation entre les niveaux sémantiques et pragmatiques en logique. Ce que j'entends par "stricte séparation" est le postulat selon lequel l'usage des connecteurs logiques n'a pas d'influence sur leur sémantique. Soulignons que ce postulat a aussi pour conséquence non négligeable la stricte séparation entre la logique d'une part, la psychologie et/ou l'épistémologie d'autre part.

Je ne peux m'appesantir ici sur ces questions, mais, pour suivre l'argument jusqu'au bout, supposons à nouveau que le problème a été résolu d'une façon ou d'une autre et que nous sommes préparés à modifier notre concept du sujet cognitif et à abandonner cette séparation stricte entre pragmatique et sémantique, pour essayer la voie opposée. Ceci revient à admettre que les ressources cognitives guident la construction de la logique en question. Ceci nous conduit en fait au type d'approche que j'ai appelé l'"antiréalisme radical", et qui est défendu par Jacques Dubucs.

### **Les inférences à ressources limitées et le défi de l'antiréalisme radical**

**Dans cette approche, l'idée principale est que si l'objet préférentiel de la logique est l'inférence, c'est-à-dire la relation entre un sujet épistémique et des propositions, alors la logique elle-même devrait prendre plus sérieusement en compte la notion de sujet épistémique et les limitations cognitives de ce sujet. Le sujet épistémique [responsable de la logique] devient donc ici plus proche du sujet cognitif aux ressources limitées [responsable de la performance]. Dans cette approche, le point crucial est que les ressources cognitives du sujet épistémique devraient déterminer d'une façon ou d'une autre la logique résultante [: les ressources cognitives n'altèrent pas seulement les performances du sujet mais la logique qu'il met en œuvre.]**

Du point de vue de la *cognitive science*, l'avantage principal de cette approche, qui est explicitement procédurale, est qu'elle prend en compte les limitations de ressource en tant que complexité, ce qui est immédiatement pertinent pour toute théorie psychologique relative à la construction de modèles mentaux pour le raisonnement. En effet, dans le travail original de Johnson-Laird, les modèles mentaux associés à des schémas de raisonnement même très simples, sont terriblement compliqués et computationnellement complexes, aussi une théorie plus consciente de ces limitations pourrait conduire à des progrès significatifs.

Du point de vue de la philosophie de la logique, l'approche procédurale semble offrir un concept plus plausible de la signification (*meaning*) associée aux constantes logiques, dès lors que des considérations contextuelles ou dynamiques jouent un rôle central. En effet, sur le plan épistémologique, la stratégie suggérée par l'antiréalisme radical semble en général cohérente et suffisante — à un point près sur lequel je reviendrai dans un instant. Cependant, du point de vue [psychologiste] de la théorie des modèles mentaux, des questions difficiles paraissent surgir en nombre.

Commençons par le point épistémologique que je mentionnais : dans l'antiréalisme radical, la méthode de décision est étroitement reliée à la logique concernée, c'est-à-dire qu'il existe une méthode effective pour décider si une expression du langage logique appartient à l'ensemble des théorèmes ou non. On peut prendre comme exemple ce qu'on appelle la "logique classique" ou n'importe quelle autre logique. Faisons l'hypothèse que deux méthodes de décision différentes existent pour cette logique, l'une très complexe computationnellement, l'autre non. Face à cette situation, l'approche antiréaliste conduit directement à boire le calice jusqu'à la lie et à établir que l'on parle en réalité de deux logiques différentes, ce qui revient à reformuler l'hypothèse de départ. Dov Gabbay a soutenu cette position en défendant une notion purement procédurale de la logique.

Mais cette lie est encore plus difficile à boire pour le psychologue [qui construit des modèles mentaux]: si la méthode de construction du modèle mental définit la logique en question, en raison de certaines limitations de ressources comment séparer ces limitations [d'ordre épistémique] des limitations individuelles du sujet? Il faudrait alors dire que chaque méthode de décision différente suppose un processus de construction mental différent. L'unité de la procédure de construction des modèles mentaux semble alors difficile à défendre.

La tâche, au-delà des questions théoriques mentionnées plus tôt, semble être de distinguer les limitations de ressource (ou compétence) des limitations subjectives (ou performance), et de faire cela sans tomber dans les postulats platoniciens de Frege. Remarquons que dans l'approche que je qualifiais de *gricéenne*, cette difficulté ne se posait pas car la logique était strictement séparée de la psychologie, et la signification logique des connecteurs logiques était donc d'emblée préservée de l'usage ou du mésusage par les sujets.

Malheureusement, je n'ai pas de solution développée à vous proposer, bien que je sois tenté de suivre la voie suggérée par Hintikka et de mettre en avant ma propre conception de la logique, à savoir la dialogique, qui est proche de l'antiréalisme radical. Mais, pour être honnête, je ne sais pas encore comment remplir ce programme. Peut-être serait-il fécond de considérer les dialogues comme des jeux de langage, afin d'étudier et de normer les performances logiques. Cela sonne beaucoup plus comme un programme épistémologique qu'un programme cognitif. Mais, en fait, je me demande, où réside la différence ?