

expliquer les mécanismes neuronaux associés à la conscience et qu'elles pourraient même détailler les correspondances entre des événements cérébraux et des éléments de conscience. Toutefois, nous ne franchirons pas ce que le philosophe Joseph Levine a nommé le fossé explicatif entre les processus physiques et la conscience tant que nous ne saurons pas pourquoi ces processus secrètent la conscience. Pour faire ce saut, nous devons utiliser un nouveau type de théorie.

Une vraie théorie de tout

De quelle théorie avons-nous besoin ? Observons tout d'abord qu'en physique tous les systèmes ne sont pas décomposables en éléments plus fondamentaux. L'espace-temps, la masse et la charge sont considérés comme des caractéristiques fondamentales du monde : ils ne sont pas réductibles à quelque chose de plus simple. Malgré cette irréductibilité, des théories détaillées et utiles relient ces éléments par des lois fondamentales, expliquant de nombreux phénomènes complexes.

L'idée que la physique contient tous les éléments fondamentaux et toutes

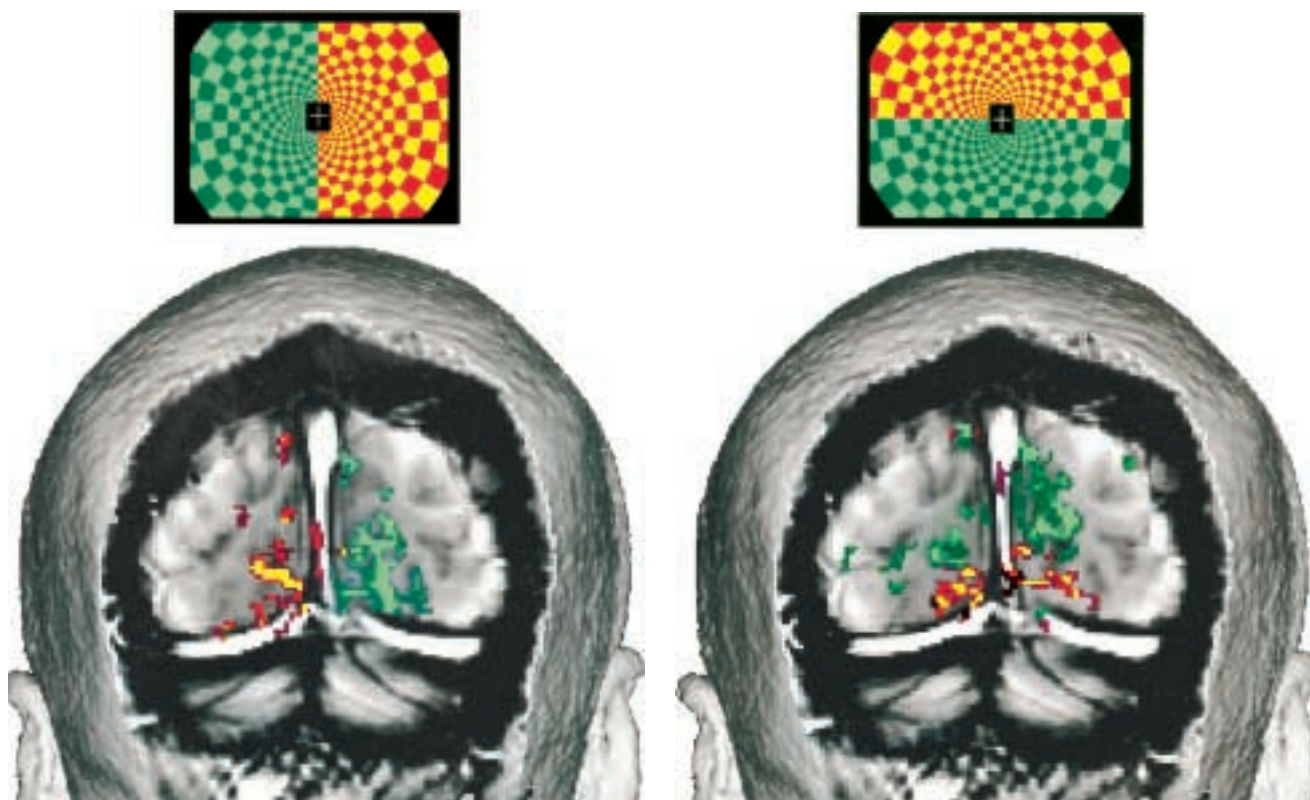
les lois fondamentales de l'Univers est largement répandue. Comme le physicien Steven Weinberg l'a écrit, le rêve du physicien est une «théorie de tout», à partir de laquelle on déduirait tout ce qu'il faut savoir sur l'Univers. S. Weinberg concède toutefois que la conscience pose un problème : malgré la puissance explicative de la théorie physique, on ne peut apparemment pas en déduire l'existence de la conscience. Même si la physique permettait d'expliquer les mécanismes cérébraux objectifs associés à la conscience, elle n'expliquerait pas la conscience elle-même, de sorte qu'aucune théorie physique ne sera une vraie théorie de tout : pour élaborer cette dernière, il faut introduire une composante supplémentaire.

Je propose donc que l'on considère la conscience comme une caractéristique fondamentale, irréductible à quoi que ce soit de plus élémentaire. L'idée peut paraître étrange, mais la cohérence l'exige et l'histoire nous donne quelques précédents : au XIX^e siècle, par exemple, quand les phénomènes électromagnétiques étaient inexplicables à partir des principes connus, les physiciens ont introduit une nouvelle entité fondamentale,

la charge électromagnétique, et ils en ont étudié les lois fondamentales. Ne peut-on reprendre cette méthodologie dans le cas de la conscience ? Si les théories actuelles ne l'expliquent pas, alors il faut introduire de nouveaux éléments.

Un nouvel élément fondamental requiert de nouvelles lois fondamentales : celles-ci relieront la conscience aux éléments de la théorie physique, sans interférer avec les lois physiques, car le monde physique forme un tout fermé. Les nouvelles lois serviront plutôt à franchir le fossé explicatif : elles indiqueront comment la conscience dépend des processus physiques sous-jacents. Une théorie complète aura deux composantes : des lois physiques, qui décriront le comportement des systèmes physiques, de l'infiniment petit à l'infiniment grand, et des lois «psychophysiques», qui diront comment certains systèmes physiques sont associés à une conscience.

Comment découvrir les lois psychophysiques ? L'obstacle majeur, dans cette recherche, est le manque de données. Telle que je l'ai décrite ici, la conscience est subjective, de sorte que nous ne pouvons pas analyser directement celle d'autrui. Cet obstacle



3. LES VARIATIONS DU DÉBIT SANGUIN dans le cortex visuel montrent comment le cerveau d'un sujet répond à une configuration qu'il voit. Les couleurs de cette illustration représentent les zones

cérébrales activées par la vision de chaque moitié du modèle. Cette expérience montre les réactions neuronales associées à la sensation visuelle, qui seraient à la base de la conscience.

n'est toutefois pas insurmontable : chacun d'entre nous a accès à ses propres expériences, une source abondante, utilisable pour formuler des théories. Nous pouvons utiliser des informations indirectes, telles les descriptions par d'autres de leur conscience. Les arguments philosophiques et les expériences de pensée apportent aussi des éléments. Ces méthodes ont leurs limites, mais elles nous fournissent largement de quoi commencer la recherche.

Comme les théories psychophysiques ne pourront pas être testées de manière indiscutable, elles seront plus spéculatives que celles des disciplines scientifiques classiques. Toutefois, nos expériences intimes, ainsi que les données extraites des témoignages d'autres, sont des données assez contraignantes : si nous trouvons une théorie qui en rende mieux compte que d'autres, nous l'accepterons. Comme aucune théorie psychophysique n'existe aujourd'hui, il serait de toute façon prématuré de se préoccuper de la réfutabilité de ces théories.

Conscience et présence à soi

Commençons à chercher des lois générales qui relient les processus physiques à l'expérience consciente. Lorsque nous sommes conscients de quelque chose, nous sommes généralement capables d'agir dessus et d'en parler (grâce à des fonctions objectives, physiques). Réciproquement, une information dont on peut parler et que l'on peut utiliser pour agir est généralement consciente. Par conséquent, la conscience est reliée avec ce que nous pourrions nommer la «présence à soi», le processus par lequel l'information, dans le cerveau, est rendue globalement disponible pour des opérations motrices telles que la parole ou l'action corporelle.

La notion de «présence à soi» semble banale, mais, telle que nous l'avons définie ici, elle est objective et physique, alors que la conscience ne l'est pas. On devra préciser la relation entre la présence à soi et la conscience pour analyser le cas des animaux et des bébés, qui ne parlent pas, mais, dans

les cas simples, nous voyons les grandes lignes d'une loi psychophysique : il y a conscience quand il y a présence à soi, et réciproquement.

Poursuivons cette idée en analysant le phénomène de perception visuelle. Nous ressentons notre champ de vision comme une mosaïque constamment changeante de couleurs, de formes et de configurations qui, en tant que telle, possède une structure géométrique détaillée. Puisque nous pouvons décrire cette structure, nous diriger vers ses composantes et réaliser d'autres actions qui dépendent d'elle, nous sommes incités à admettre que la structure de la conscience correspond directement à la structure de l'information donnée au cerveau par les processus neuronaux de la perception visuelle.

Nos sensations de la couleur ont aussi une dimension tridimensionnelle intrinsèque, traduite par la structure du traitement de l'information dans le cortex visuel. Cette structure s'observe sur les disques de couleurs utilisés par les artistes : les couleurs y sont disposées dans un ordre précis, du rouge

Pourquoi les neurosciences peuvent expliquer la conscience

Nous pensons qu'aujourd'hui la meilleure démarche pour expliquer la conscience est l'étude des mécanismes cérébraux qui en sont le plus directement responsables. En localisant les neurones qui correspondent le plus à la conscience et en comprenant comment ils sont associés à des neurones d'autres parties du cerveau, nous découvrirons de nouveaux éléments de réponse à ce que David Chalmers nomme le «problème difficile» : une explication complète de la manière dont la conscience subjective émerge de ces mécanismes cérébraux.

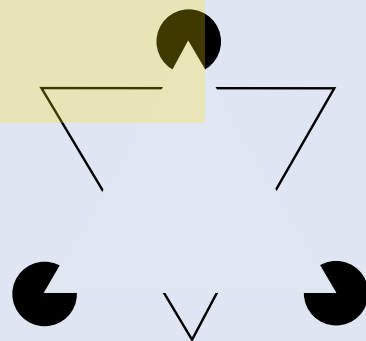
Nous admirons l'audace de D. Chalmers, qui identifie d'emblée le problème difficile et tente de le résoudre, mais nous ne sommes pas convaincus par certaines de ses expériences de pensée. Selon nous, le problème difficile peut être décomposé en plusieurs questions. Pourquoi avons-nous une conscience ? Comment avons-nous des sensations conscientes particulières (telle la sensation du bleu de la couleur bleue) ? Pourquoi certains éléments de l'expérience subjective sont-ils impossibles à communiquer à d'autres personnes (pourquoi sont-ils personnels) ? Nous pensons avoir trouvé une réponse à la dernière question, et nous avons des propositions pour les deux premières, sur la base d'un phénomène nommé la représentation neuronale explicite.

Illustrons le terme «explicite» par un exemple. En réaction à l'image d'un visage, des cellules ganglionnaires sont activées dans toute la rétine, à peu près comme les pixels d'un écran de télévision, et engendrent une représentation implicite du visage. Simultanément, elles réagissent à d'autres caractéristiques de l'image, telles des ombres, des lignes

ou les variations d'éclairage. Des neurones, à des niveaux supérieurs du cortex visuel, réagissent surtout au visage, ou au visage vu sous un certain angle. Cette catégorie de neurones permet au cerveau de représenter le visage de manière explicite. Leur mort, à la suite d'un accident vasculaire cérébral ou d'une lésion, entraîne la prosopagnosie, l'incapacité pour un sujet de reconnaître des visages familiers, y compris le sien, de manière consciente, bien que la personne reste capable d'identifier un visage comme étant un visage. De la même manière, des lésions touchant d'autres parties du cortex visuel peuvent provoquer la perte de la sensation des couleurs, excepté le noir et blanc, et sans qu'il y ait aucun déficit des récepteurs rétiniens de la couleur.

À chaque étape de sa progression dans le cortex, l'information visuelle est codée, généralement de manière semi-hiérarchique : les cellules ganglionnaires de la rétine réagissent à un point lumineux ; les neurones du cortex visuel primaire réagissent préférentiellement à des lignes ou à des arêtes ; des neurones supérieurs détectent les contours en mouvement, etc. Ceux qui réagissent aux visages ou aux objets familiers sont encore plus haut dans la hiérarchie. Au sommet, les neurones sont reliés aux structures prémotrices et motrices du cerveau, où ils activent les neurones qui déclenchent des actions telles que parler ou éviter un véhicule qui passe.

D. Chalmers pense, comme nous, que les aspects subjectifs des perceptions sont étroitement liés à l'activité des neurones qui leur correspondent. Il décrit une expérience de pensée très connue, où une neurobiologiste hypothétique, Marie, spécialisée dans la perception des couleurs, n'a jamais vu une couleur. Nous pensons que Marie ne sait



Le triangle de Kanizsa stimule des neurones qui codent explicitement un contour qui n'existe pas.