

Simon Blackburn

Penser.

Une irrésistible introduction à la philosophie

Chapitre 2

Esprit

Supposons que nous laissons de côté le problème général de l'harmonie entre notre perception du monde et ce qu'il est. Nous croiserons les doigts en supposant que nous savons vraiment ce que nous croyons naturellement savoir. Mais comment nos vues s'agencent-elles ? Avec Descartes, notre moi et notre esprit sont devenus les objets intimes et particuliers d'un savoir immédiat. Car même si je puis m'extraire d'un océan de doute pour grimper sur le rocher du *Cogito*, je ne puis me hisser sur la nature de *votre* esprit. Comment puis-je donc savoir quoi que ce soit de votre vie mentale ? Comment puis-je savoir, par exemple, que vous voyez la couleur bleue comme moi ? Se pourrait-il que certains d'entre nous ressentent la douleur avec plus d'acuité, mais en parlent moins, quand d'autres la ressentent moins mais en font tout un fromage ? Comment nous mettre à réfléchir à l'esprit et au corps, au cerveau et au comportement ?

LE FANTÔME DANS LA MACHINE

Nous avons vu comment la stratégie de Descartes l'a conduit à juger la connaissance de nos esprits plus sûre et certaine que la connaissance du reste du monde. Mais Descartes était aussi un homme de science. On lui doit des découvertes fondamentales en optique. Il pratiqua la dissection et en savait long sur la transmission des impulsions au cerveau à travers les nerfs¹⁸. Il savait qu'elle passait par un moyen de transmission physique, une « traction » ou un « mouvement violent » des nerfs, ou comme nous dirions aujourd'hui, une impulsion électrochimique transmise par le système nerveux. Les sens ordinaires de la vue, du toucher, du goût, de l'odorat et de l'ouïe activent le système nerveux, qui transmet des messages au cerveau. Bien entendu, celui-ci n'est pas un bloc indifférencié. Des secteurs du cerveau transmettent des signaux à d'autres secteurs, qui les répercutent au corps : des systèmes entiers d'activation se mettent en place. Tout cela relève de la neurophysiologie. En principe, tous ces phénomènes sont visibles en public : avec les bons instruments, on peut faire une démonstration des systèmes d'activation en salle de cours.

Et après ?

Eh bien, vient ensuite le moment magique. L'« esprit » (la chose pensante, la *res cogitans*) en est également affecté, et c'est tout l'univers de l'expérience qui s'ouvre. Le sujet voit des couleurs, entend des sons, sent des textures ou des températures, a des sensations gustatives ou olfactives. Ce monde de l'expérience se compose d'événements

mentaux ou d'événements au sein de la conscience subjective, qui, eux, ne sauraient être vus en public. Ce sont des événements privés. Les élèves d'une classe ne peuvent voir la mise à feu de certains neurones : seul le peut celui qui ressent la douleur. Descartes localisa effectivement l'endroit où se produit l'événement magique. Pour de fort raisonnables raisons neurophysiologiques, il crut que la glande pinéale, structure nichée au centre du cerveau, devait être le lieu où les messages passaient du champ du physique à celui du mental.

Pour Descartes, les événements mentaux ne sont pas seulement distincts des événements physiques. Ils relèvent aussi d'un autre type de substance – d'une substance immatérielle – un genre de matériau spectral ou d'ectoplasme. À rigoureusement parler, si je dis « J'ai pensé à la reine et j'ai salué », il y a une ambiguïté : le « je » qui est le sujet de la pensée n'est pas le « je », le corps, qui salue. Les pensées et les expériences sont des modifications d'un genre de matériau ; mouvement et position appartiennent à un autre. Cette partie de sa doctrine fait de Descartes un « dualiste de la substance ». Il y a non seulement deux types de propriétés (mentales et physiques) que peuvent avoir les personnes, mais aussi deux types de porteurs des dites propriétés. Bien entendu, c'est théologiquement commode : cela ouvre la voie à l'immortalité de l'âme, puisqu'il n'y a aucune raison pour que le matériau de l'âme ait la même durée de vie que tout ce qui ressemble au corps physique. Mais le dualisme de la substance n'a rien d'obligatoire. On pourrait soutenir que les propriétés mentales et physiques sont très différentes, mais que l'unique

corps organisé possède les deux : après tout, masse et vitesse sont deux types de propriété très différents, mais les projectiles les possèdent toutes deux. On appelle dualistes des propriétés les gens qui pensent qu'il existe deux sortes de propriétés (mentales et physiques), mais qu'elles appartiennent à une seule espèce de matériau (de quoi que puissent être faits les gros animaux).

Descartes nous conduit au point de vue qu'a bien résumé Gilbert Ryle (1900-1976) en parlant de l'être humain comme d'un « spectre dans une machine¹⁹ ». Les événements qui se produisent dans la machine, dans le corps physique, sont comme d'autres événements du monde physique. Ils consistent en interactions de types de matériaux familiers : molécules et atomes, champs et forces électriques. Les événements de la partie spectrale, l'esprit, sont complètement différents. Peut-être sont-ils des événements dans une sorte de matière spectrale – d'ectoplasme ou de cette substance non physique dont sont faits les esprits et les anges. Dans l'esprit populaire, en effet, ceux-ci se passent de toute incarnation. Chez l'être humain normal, en revanche, existe une étroite corrélation entre les événements d'une espèce et ceux de l'autre : piquer quelqu'un avec une aiguille induit des changements physiques, mais crée aussi un événement mental sous forme de douleur. Inversement, l'événement mental qu'est le souvenir d'une bévue peut susciter gémissement et rougissement, c'est-à-dire des événements physiques. Les événements d'un domaine peuvent donc affecter ceux d'un autre. En principe, cependant, les deux champs sont entièrement distincts.

ZOMBIES ET MUTANTS

Bien entendu, ce point de vue n'est pas propre à Descartes. Nombreuses sont les grandes religions du monde qui le présupposent : il est présent dans toute doctrine soutenant que nous pouvons survivre à la mort physique, ou que notre âme peut suivre un chemin, et notre corps un autre. Mais c'est un point de vue qui se heurte à des problèmes énormes, sinon insurmontables.

La première famille de problèmes est épistémologique. Chez l'être humain normal, je viens de le dire, existe une corrélation étroite entre les deux sortes d'événements. Mais qu'est-ce qui nous autorise à le croire ? Voici l'une des possibilités :

Le zombie. Les zombies sont comme vous et moi et se conduisent comme vous et moi. Leurs natures physiques sont identiques. Ouvrez le cerveau d'un zombie, vous découvrirez qu'il marche exactement comme le vôtre ou le mien. Piquez un zombie, il fera « aïe », toujours comme vous et moi. Mais les zombies ne sont pas conscients. Il n'y a aucun spectre en eux.

Parce que les zombies sont pareils à nous et se conduisent comme nous, il n'y a pas moyen de dire qui est un zombie et qui est conscient comme vous et moi. En tout cas, comme moi. Car maintenant que j'ai soulevé la possibilité des zombies, je vois bien que je ne puis vraiment être sûr de vous ni de personne d'autre. Peut-être la conscience est-elle le corrélat extrêmement rare d'un système complexe corps-cerveau. Peut-être en suis-je le seul exemple : peut-être êtes-vous tous des zombies, vous autres.

Mais voici une autre possibilité :

Le mutant. Les mutants sont comme vous et moi et se conduisent comme vous et moi. Leurs natures physiques sont identiques. Ouvrez le cerveau d'un mutant, vous découvrirez qu'il marche exactement comme le vôtre ou le mien. Piquez un mutant, il fera « aïe », toujours comme vous et moi.

Contrairement aux zombies, les mutants sont conscients. Il y a un spectre en eux. Mais les événements à l'intérieur du mutant ne sont pas ce qu'on attendrait. Un mutant que l'on pique, par exemple, peut expérimenter un événement mental pareil à l'audition d'un do sur une clarinette. Il n'en fait pas moins « aïe » parce que son cerveau fonctionne comme le nôtre et qu'il se conduit comme nous : le fait d'être piqué avec une aiguille amorce des processus produisant des changements qui l'amènent à faire « aïe », exactement comme nous. Peut-être ressent-il une affreuse douleur lorsqu'il entend un do à la clarinette, mais cela lui arrache alors un sourire béat. Un mutant voyant des boîtes aux lettres les voit sans doute roses ; voit-il des lauriers roses, ils seront jaunes. Les événements qui se produisent dans la conscience du mutant sont sans rapport avec les événements qui se produisent dans votre esprit ou le mien. En tout cas, sans lien aucun avec les événements de mon esprit. Car maintenant que j'ai évoqué la possibilité du mutant je vois bien que je ne puis vraiment être sûr de vous ni de personne d'autre. Peut-être êtes-vous tous des mutants, vous autres, par rapport à moi.

Ce qui distingue ces possibilités, c'est qu'elles paraissent grandes ouvertes en vertu de l'explica-

tion cartésienne dualiste du corps et de l'esprit. Ce sont des possibilités déroutantes, de celles qu'on ne considère pas d'ordinaire, même si je soupçonne qu'elles nous traversent l'esprit *plus* souvent que les extravagantes possibilités du chapitre premier.

Une des réactions consiste à mordre la balle et à dire : très bien, supposons que ce sont des possibilités grandes ouvertes. Peut-être ne puis-je pas vraiment savoir ce qu'il y a dans la tête d'une autre personne comme moi, les événements mentaux qui se produisent en elle, ou même s'il s'y déroule la moindre vie mentale. Mais ne puis-je tout de même *supposer* que la vie mentale des autres ressemble beaucoup à la mienne ? Ne puis-je me prendre pour modèle de tous les autres ? Il s'agirait moins de *connaissance* que d'hypothèse ou de *conjecture*, mais peut-être est-ce une conjecture *raisonnable*. C'est ce qu'on appelle l'argument par analogie avec l'existence d'autres esprits.

Le problème de cet argument est qu'il paraît incroyablement faible. « Et comment puis-je généraliser ce cas *unique* de manière aussi irréfléchie²⁰ ? » demandait avec humeur le grand philosophe autrichien, Ludwig Wittgenstein (1889-1951). Le simple fait que dans *un* cas – le mien – comme par hasard il y ait une vie mentale d'un type particulier, bien défini, associée à un cerveau et à un corps, paraît être une raison très fragile de supposer la même association dans toutes les autres cas. Si j'ai une boîte avec un coléoptère à l'intérieur, j'ai bien peu de raisons de supposer que tous ceux qui ont une boîte ont aussi un scarabée à l'intérieur.

Pis encore, cela me donne de très piètres raisons

de nier qu'il y ait des coléoptères ailleurs que dans des boîtes. Dès lors, peut-être des choses qui sont très *différentes* de vous et de moi physiquement sont-elles conscientes exactement comme je le suis : les cailloux et les fleurs, par exemple.

Peut-être serez-vous tenté d'écarter d'un haussement d'épaules l'hypothèse du zombie ou du mutant, de n'y voir que pure fantaisie philosophique, irréelle et en tout état de cause invérifiable. Les neurophysiologistes, par exemple, ne trouvent pas d'expérience consciente comme ils trouvent des neurones, des synapses et des structures d'activité cérébrale : ou comme on dit, ils ne peuvent la projeter à leurs étudiants sur l'écran de leur amphithéâtre. Mais alors, suivant le dualisme cartésien, les possibilités auxquelles nous croyons tous naturellement, à savoir que les autres ne sont *pas* des zombies *ni* des mutants sont elles-mêmes invérifiables ! Ce sont de purs articles de foi. Quelqu'un qui croit à la possibilité des zombies n'est pas logé à plus mauvaise enseigne que nous de ce point de vue.

En fait, si notre conception de l'esprit admet les possibilités du zombie et du mutant, nous pourrions même les supposer fort probables ou, tout au moins, aussi probables que toute autre possibilité. Car s'il n'est pas faux *a priori* que les autres soient des zombies, pourquoi serait-il *a priori* moins probable qu'ils sont conscients comme moi ?

Pourquoi les philosophes parlent-ils tant de possibilités bizarres que les autres se font un bonheur d'ignorer (l'une de ces choses qui donnent à la discipline des airs intimidants et assurent sa mauvaise réputation) ? La raison en est que les possibilités

en question servent à tester une conception des choses. Ici, elles servent à tester la conception de l'esprit et de la matière qui leur a donné naissance. Le raisonnement est que, *si* l'esprit et la matière sont envisagés de manière cartésienne, il y aurait des possibilités grandes ouvertes de nature bizarre, sur lesquelles nous ne pourrions rien savoir. *Donc*, puisque c'est intolérable, nous devons repenser la conception des choses (c'est ce qu'on appelle la métaphysique). Une meilleure conception de l'esprit et de sa place dans la nature devrait exclure ces possibilités. Nous dirons : selon le dualisme cartésien, la possibilité du zombie et celle du mutant demeurent toutes deux grandes ouvertes, mais cela montre simplement que quelque chose ne va pas dans le dualisme cartésien. Le mental et le physique ne sont pas aussi distincts qu'il l'assure. Parce qu'il n'est tout simplement *pas* possible que quelqu'un qui vient de se cogner le pied et hurle de douleur le fasse *parce qu'il* est dans le même état mental que moi quand une clarinette fait entendre un do. *Cet* état mental-là ne saurait s'exprimer par un hurlement ou un grognement. Le lien entre la nature intrinsèque de l'état mental – ce qu'il ressent – et son expression est plus proche que cela. Nous *savons* que quelqu'un qui vient de se cogner le pied *ne* hurle *pas* parce que son expérience est la même que la mienne quand j'entends un do sur une clarinette. On sait qu'il exprime une chose très proche de la mienne quand je me cogne le pied.

L'argument de l'analogie avec les autres esprits était la cible de Wittgenstein. Sa principale objection ne tient pas seulement à sa faiblesse. Il essaie

de montrer que, si tout notre savoir des états mentaux nous venait de notre propre cas, il serait impossible ne serait-ce que de penser en termes de conscience des autres. Comme si, laissant tomber une brique sur votre pied, il n'y avait aucune douleur, puis que je n'en éprouvais aucune, et que l'affaire était ainsi entendue. Mais puisque nous pensons bel et bien en termes d'autres esprits et de leurs expériences, il nous faut les conceptualiser d'une autre façon.

De ce point de vue, l'issue consiste à rejeter l'image de l'esprit et du corps que nous donne le dualisme cartésien. Et des pressions aussi bien *métaphysiques* qu'*épistémologiques* doivent nous encourager à rejeter le dualisme cartésien. Peut-on réellement obtenir une image possible du monde à partir du dualisme cartésien, peu importe que nous sachions ou non comment il est ? Prenons de nouveau le cas du zombie. Son fonctionnement physique est identique au nôtre. Il répond au monde de la même façon ; sa santé dépend des mêmes variables que la nôtre. Il peut rire aux bons moments et pleurer aux tragédies appropriées. Il peut être de bonne compagnie. Que *fait* donc l'absence de conscience ? Ou, pour dire les choses autrement, qu'est-ce que la conscience est censée *faire* pour nous ? Devons-nous en conclure que chez nous, qui ne sommes pas des zombies, les événements mentaux existent, mais ne *font* rien ? La conscience est-elle comme le sifflet du moteur, qu'elle n'est pas une partie de la machine qui fait arriver les choses ? (C'est la doctrine connue sous le nom d'« épiphénoménalisme ».) Mais si les

esprits ne font rien, pourquoi ont-ils évolué ? Pourquoi la nature les a-t-elle sélectionnés ? Et si les états mentaux ne font vraiment rien, comment, par exemple, entrent-ils dans la mémoire ?

C'est tout le problème de l'interaction cerveau-esprit tel qu'il se présente pour le dualisme cartésien.

LOCKE, LEIBNIZ ET LE BON PLAISIR DE DIEU

Le problème est bien résumé dans le débat qui opposa John Locke et son contemporain, le grand mathématicien et philosophe Gottfried Wilhelm Leibniz (1646-1716). Locke est un autre penseur du XVII^e siècle à s'être inquiété des implications de la vision scientifique moderne du monde. En particulier, il s'est préoccupé du point de causation, auquel les mouvements des particules dans le cerveau donnent naissance aux idées, comme celles de couleur, dans l'esprit. Dans le passage suivant, il parle de la façon dont les bombardements de particules atomiques donnent naissance à des choses telles que des odeurs, des goûts, des sons et des couleurs :

Cela étant, nous sommes en droit de supposer que ces sortes de particules, différentes en mouvement, en figure, en grosseur et en nombre, venant à frapper les différents organes de nos sens, produisent en nous ces différentes sensations qui nous causent les couleurs et les odeurs des corps ; qu'une violette, par exemple, produit en nous les idées de la couleur bleuâtre, et de la

douce odeur de cette fleur, par l'impulsion de ces sortes de particules insensibles, d'une figure et d'une sorte de grosseur particulière, qui diversement agitées viennent à frapper les organes de la vue et de l'odorat. Car il n'est pas plus difficile de concevoir que Dieu peut attacher de telles idées à des mouvements avec lesquels elles n'ont aucune ressemblance, qu'il est difficile de concevoir qu'il a attaché l'idée de la douleur au mouvement d'un morceau de fer qui divise notre chair, auquel mouvement la douleur ne ressemble en aucune manière²¹.

Locke partageait l'idée que nous avons déjà rencontrée chez Newton et Descartes, à savoir que certains processus causaux étaient relativement intelligibles, notamment ceux dans lesquels une qualité, comme le mouvement, se transmet par impact d'une particule à l'autre. Mais le moment de la causation du corps vers l'esprit, dans lequel les mouvements du cerveau produisent quelque chose d'entièrement différent, les sensations olfactives, les couleurs, ou la douleur, demeurerait totalement obscur. Il est tout à fait stupéfiant que les événements mentaux se produisent quand ils le font. C'est dû à ce que Locke appelle ailleurs la « volonté arbitraire et le bon plaisir » de Dieu, « le sage architecte » qui « annexe » des modifications particulières de la conscience à des événements physiques particuliers. Dans la terminologie de Descartes, Locke pense que nous n'avons pas d'idée « claire et distincte » des types de système que Dieu a pu choisir comme lieux adéquats pour surajouter la conscience. Que l'univers soit orga-

nisé en sorte que certains systèmes possèdent une conscience et d'autres pas ne serait qu'un fait brut. De même qu'est simplement un fait brut que leur conscience change et acquière des propriétés définies au moment où leurs moi physiques changent et acquièrent des propriétés particulières. Le contraste est entre une connexion rationnelle et intelligible, telle que nous la trouvons dans la discipline *a priori* des mathématiques, et le fait que certains « mouvements » produisent en nous les sensations qu'ils produisent. Tel est le fait brut, conséquence du bon plaisir de Dieu.

En fait, Locke, ici, n'est pas loin de la doctrine connue sous le nom d'*occasionalisme*, que devait embrasser un autre contemporain, Nicolas Malebranche (1638-1715). D'après celle-ci, les événements physiques ne causent ni ne produisent, à rigoureusement parler, des événements mentaux, mais ménagent des occasions pour Dieu d'insérer dans nos biographies des événements mentaux de nature appropriée. À rigoureusement parler, nos corps n'affectent pas nos esprits, mais ménagent simplement des occasions où Dieu le fait. Locke lui-même ne le dit pas, mais nous pourrions songer que la différence est mince mais précieuse entre, d'un côté, un Dieu qui intervient selon son bon plaisir pour que le découpage de la chair par le fer procure une sensation de douleur et, de l'autre, un Dieu qui injecte une sensation de douleur dans l'âme chaque fois que le fer incise la chair.

La doctrine de Locke troubla profondément Leibniz. Dans ce passage de ses *Nouveaux Essais*, qui sont un commentaire serré de Locke, Philalèthe est le porte-parole de Locke, et Théophile celui de

Leibniz. Remarquez la citation directe du passage de Locke reproduit plus haut :

PHILALÈTHE. Maintenant certaines particules, frappant nos organes d'une certaine façon, causent en nous certains sentiments de couleurs ou de saveurs ou d'autres qualités secondes, qui ont la puissance de produire ces sentiments. Et il n'est pas plus difficile de concevoir que Dieu peut attacher telles idées *comme celle de chaleur* à des mouvements, avec lesquelles elles n'ont aucune ressemblance, qu'il est difficile de concevoir qu'il a attaché l'idée de la douleur au mouvement d'un morceau de fer qui divise notre chair, auquel mouvement la douleur ne ressemble en aucune manière.

THÉOPHILE. Il ne faut point s'imaginer que ces idées comme de la couleur ou de la douleur soient arbitraires et sans rapport ou connexion naturelle avec leurs causes : ce n'est pas l'usage de Dieu d'agir avec si peu d'ordre et de raison. Je dirais plutôt qu'il y a une manière de ressemblance, non pas entière et pour ainsi dire *in terminis*, mais expressive, ou de rapport d'ordre, comme une ellipse et même une parabole ou hyperbole ressemblent en quelque façon au cercle dont elles sont la projection sur le plan, puisqu'il y a un certain rapport exact et naturel entre ce qui est projeté et la projection qui s'en fait, chaque point de l'un répondant suivant une certaine relation à chaque point de l'autre. C'est ce que les cartésiens ne considèrent pas assez et cette fois vous leur avez plus déferé, Monsieur, que vous n'avez coutume et que vous n'aviez sujet de faire. [...]

Il est vrai que la douleur ne ressemble pas au mouvement d'une épingle, mais elle peut ressem-

bler fort bien à des mouvements que cette épingle cause dans notre corps, et représenter ces mouvements dans l'âme, comme je ne doute nullement qu'elle ne fasse²².

Où Locke ne voit que « bon plaisir de Dieu », Leibniz paraît insister qu'il doit y avoir une connexion rationnelle. Les événements de l'âme doivent entretenir une relation quasi mathématique avec les « mouvements » du cerveau et du corps qui les produisent.

Nous pouvons poser le problème ainsi. Imaginez Dieu créant l'univers. Comment doit-il faire ? La doctrine suivante ne manquerait pas d'attrait : il doit créer le matériau physique et les lois de la physique, puis tout le reste s'ensuit. Dans cette optique, en fixant l'état *physique* de l'univers en tous temps, un Dieu créateur fixe tout à toutes les époques. S'il avait voulu faire un monde avec quelque chose de différent – où les piqûres d'épingle ne feraient pas mal –, il aurait tripatouillé les faits *physiques* pour que cela ne se produise pas. Il aurait dû organiser des nerfs et des circuits différents dans le corps et dans le cerveau. Il n'est pas de variation *indépendante* en vertu de laquelle le physique pourrait rester identique tandis que le mental serait différent. Telle est la position de Leibniz, du moins telle qu'elle apparaît dans ce passage. (Selon une autre interprétation, Leibniz pense qu'il y a une variation indépendante, mais que Dieu a, naturellement, choisi la *meilleure* manière d'associer les événements mentaux et physiques.)

Locke, en revanche, croit que Dieu a deux

choses différentes à faire. Premièrement, établir la physique et les lois de la physique. Mais ensuite décider comment « annexer » des événements mentaux aux événements physiques en fixant des relations psychophysiques. Tout se passe comme si le monde avait deux biographies – une d'événements physiques, une autre d'événements mentaux – et que Dieu avait dû décider comment les rattacher. De ce point de vue, il pourrait y avoir variation indépendante. Dieu aurait pu garder la physique en l'état, mais décider de ne pas annexer la douleur aux piqûres d'épingle.

Considérez maintenant une personne (vous) et un double physique de celle-ci (un jumeau). Si Locke a raison, il est en principe possible que le jumeau soit un zombie ou un mutant. Bien que son moi physique soit pareil au vôtre, il faudrait un exercice arbitraire de la bonté divine pour que leur vie mentale soit également semblable. C'est particulièrement évident dans la version « occasionnaliste » : pour des raisons impénétrables qui lui appartiennent, peut-être Dieu profite-t-il de l'occasion de mon orteil écrasé pour insérer la douleur dans ma biographie mentale, mais pas pour vous. Par ailleurs, si Leibniz a raison, cette possibilité est exclue. Si vous et votre jumeau vous cognez le pied avec la même force, et réagissez physiquement de la même façon, l'« expression » des événements physiques dans vos esprits doit être aussi la même, de même que les figures projetées par deux formes identiques sur un plan à un angle donné doivent être les mêmes.

Que Leibniz emploie une analogie mathématique ne manque pas d'intérêt. Le fait n'est pas simple-

ment qu'il fut meilleur mathématicien encore que Descartes et, entre autres choses, inventa le calcul différentiel. Il est plutôt que, pour Leibniz, tout l'ordre naturel doit finalement être transparent à la raison. Que les choses se passent de telle ou telle façon n'est pas fortuit. Si seulement on pouvait voir assez loin, il doit bien y avoir une raison à cela. Les choses doivent avoir un sens. Quand Leibniz dit que Dieu ne fait rien arbitrairement ni sans principe, il n'exprime pas tant son optimisme théologique qu'il insiste que nous devrions être capables de voir *pourquoi* les choses sont comme elles *doivent* être. C'est cette confiance dans ce qui devrait être possible à la raison qui fait de Leibniz, comme de Descartes, un « rationaliste ».

Dans la philosophie de l'esprit, le leibnizien doit nier la possibilité des zombies et des mutants. Si la biographie physique est fixée, la biographie mentale l'est aussi *ipso facto*. Il n'y a pas de variation indépendante réelle ni possible. Tout le problème *philosophique* est de comprendre pourquoi il en va ainsi. Il s'agit en fait de comprendre de quelle *manière* toute l'histoire physique rend l'histoire mentale vraie.

Locke croyait pouvoir laisser ouverte la question de savoir si c'est une « chose » immatérielle (un spectre) en nous qui pense ou si c'est le système physique lui-même puisque Dieu peut surajouter à tout ce qu'il lui plaît. Or il est on ne peut plus clair qu'il faut un esprit pour faire un esprit. Il faut une dispensation spéciale : la pensée ne saurait naître naturellement (ou, comme dit Leibniz, de façon rationnellement explicable) de la matière.

Car des parties de matière qui ne pensent point ont beau être étroitement jointes ensemble, elles ne peuvent acquérir par là qu'une nouvelle relation locale, qui consiste dans une nouvelle position de ces différentes parties ; et il n'est pas possible que cela seul puisse leur communiquer la pensée et la connaissance²³.

C'est ce genre de certitude *a priori* sur ce qui peut ou non causer d'autres choses qui fait foncièrement de Locke, comme tous les autres esprits de son temps, un rationaliste, même s'il est moins assuré des pouvoirs de la raison qu'un Descartes ou un Leibniz.

Les penseurs qui se sont intéressés à l'esprit et à la matière ne sont pas allés beaucoup plus loin que Locke et Leibniz. Aujourd'hui aussi il est des penseurs (parfois appelés les « nouveaux mystérieux ») convaincus qu'on ne comprendra jamais la relation entre l'esprit et la matière. Le problème reste en l'état où Locke l'avait laissé : une question rationnellement inexplicable, au bon plaisir de Dieu. Il est même des philosophes pour croire vraie une sorte de dualisme cartésien et penser que l'esprit est réellement un épiphénomène, qu'il ne cause jamais le moindre événement physique. Et cela parce que le physique est, pour eux, un *système fermé*. S'il existe un processus qui commence par une épingle qui vous pique et finit par une grimace, il y a, de l'épingle à la grimace, toute une chaîne physique qui explique celle-ci. Pour eux, il doit donc être *faux* que vous grimaciez parce que vous avez mal. Il faut en finir avec ce bon sens. Vous grimacez du fait de circuits physiques, non pas en

raison d'un mental surajouté. Ces penseurs sont en fait confrontés au même problème d'interaction que rencontre Locke. Nous y reviendrons dans le chapitre suivant.

Mais il est d'autres penseurs pour qui il est possible de mettre en évidence une relation rationnelle. J'introduirai deux grandes approches. La première essaie de donner une « analyse » du mental en des termes qui nous permettent d'y voir une expression leibnizienne du physique. La seconde vise une sorte de réduction scientifique du mental au physique, voire une identité entre les deux.

ANALYSE

L'analyse, telle que l'entendent les philosophes, essaie de dire ce qui rend vraies des propositions de nature mystérieuse en employant des termes d'une classe moins mystérieuse. On peut l'illustrer par un exemple familier. Supposez que quelqu'un soit dérouté par cette icône de la vie occidentale qu'est l'homme moyen, avec ses 2,4 enfants et 1,8 automobile. Comme ce chiffre qui tient de la plaisanterie peut-il avoir le moindre intérêt ? La réponse est donnée en montrant ce qui fait la vérité des propositions ainsi exprimées : en l'espèce, que le nombre d'enfants divisé par le nombre de familles est de 2,4, et le nombre d'automobiles par le nombre de propriétaires de 1,8. Cette information est succinctement présentée en termes d'homme moyen. Il est ce que Russell appelait une « construction logique » à partir d'un agrégat de faits. (Ce qui ne signifie pas que toutes les propositions sur la

moyenne soient censées ou utiles : comme on l'a dit, la personne moyenne a un seul testicule ou un seul sein.) Les philosophes parlent aussi de *réduction* d'un type de proposition à un autre. L'analyse pourvoit aux réductions²⁴.

L'analyse nous dit ce qu'il faut entendre par des propositions exprimées verbalement d'une certaine façon par des propositions exprimées en d'autres termes. Ses mérites, en tant qu'outil intellectuel, ont eux-mêmes suscité une vive controverse philosophique, et son statut a varié au fil des cent dernières années. D'aucuns, comme B. Russell et G.E. Moore (1873-1958), y ont vu l'objectif essentiel de la philosophie. Puis ses perspectives ont été contestées par le principal penseur américain du milieu du xx^e siècle, W.V.O. Quine (1908-2000), et par d'autres, dont le pessimisme a donné quelque crédibilité au fait déprimant que fort rares étaient les analyses philosophiques apparemment couronnées de succès. À l'heure actuelle, l'analyse jouit d'un timide renouveau. Mais, aux fins qui sont les nôtres, on peut laisser de côté ces questions méthodologiques. Le fait est que, si nous pouvons analyser les imputations mentales en termes physiques, se trouve ainsi justifié le rêve leibnizien qu'il y aurait une manière rationnelle ou *a priori* de voir *comment* le physique donne naissance au mental.

Prenons la douleur comme exemple d'état mental. Supposez que nous essayons d'analyser ce que signifie pour quelqu'un avoir mal. Nous identifions essentiellement la douleur en fonction de ce qu'elle nous fait (mais aussi de ce *à quoi* elle sert en termes d'évolution). La douleur nous fait faire diverses choses. Elle requiert l'attention, elle nous

fait immobiliser des parties du corps, nous détourne d'autres choses et, bien entendu, elle est désagréable. Supposez que nous puissions résumer ces conséquences en termes de tendances ou de dispositions. D'où la suggestion qu'avoir mal, c'est simplement être disposé de certaines façons. Telle est l'analyse de ce que veut dire la proposition que quelqu'un a mal, qui la rend vraie. Ce résultat serait un exercice *a priori* de la raison, produit par la réflexion sur ce que l'on entend vraiment par des propositions sur ce type d'événement mental. Ainsi se dissipe le mystère de la conscience. Puisque vous partagez les mêmes dispositions (on peut vérifier que vous avez tendance à vous conduire de la même façon), votre jumeau et vous partagez vos sensations, parce que les sensations se résument précisément à cela.

Cette doctrine a reçu le nom de « béhaviorisme logique ». Il y a, à mon sens, un fond de vérité, même si elle ne va pas sans difficulté. Ainsi pourrait-on objecter : nous savons bien que les gens peuvent partager la même sensation mais réagir un peu différemment. On peut se cogner le pied un jour et en faire toute une histoire, le refaire un autre jour, ressentir la même douleur mais sourire courageusement et suivre son chemin. Le comportement n'est pas un guide transparent des sensations, des pensées et des sentiments. (Tel est le sens de la blague sur deux béhavioristes au lit : « Pour toi, c'était super ! Et pour moi ? ») Il faut donc, à tout le moins, introduire des complications. Peut-être pourrions-nous sauver l'analyse en termes de dispositions au comportement en faisant valoir que si vous souriez vaillamment et continuez, vous êtes

encore en un sens *disposé* à des démonstrations plus expressives de douleur que vous réprimez pour une raison ou pour une autre. Il est presque impossible de supprimer entièrement les tendances comportementales dictées par la douleur, et certains excellent à faire la différence entre un enfant qui ne s'est pas fait mal et un enfant qui s'est fait mal mais fait bonne figure. Que la douleur dispose ainsi paraît être un de ses traits essentiels. Mais cela même est remis en cause par les cas de personnes victimes de certaines lésions cérébrales, et qui apparemment disent en toute sincérité que la douleur reste présente, mais qu'elles n'y prêtent plus attention. Il nous faut cependant observer qu'il est très dur de dégager le sens de ce constat. Soit un exemple de douleur carabinée : vous touchez une plaque chauffante ou vous vous cognez le pied dans le mur, il est très difficile d'imaginer ce même état mental sans imaginer qu'il est terriblement désagréable. Et il est difficile de l'imaginer sans sa propension à causer des manifestations comportementales typiques.

Les penseurs contemporains ont tendance à ne pas faire trop de crédit à ce type de béhaviorisme. Ils préfèrent une doctrine légèrement plus élaborée connue sous le nom de « fonctionnalisme ». Celui-ci aussi attache une grande attention à la fonction de l'état mental. Mais il identifie cette fonction d'une manière légèrement plus détendue. Il fait place à tout un réseau de relations physiques : non seulement des dispositions à se conduire de telle ou telle façon, mais aussi des causes typiques, voire des effets sur d'autres états mentaux – sous réserve que ceux-ci soient à leur tour convenablement

exprimés sous forme de dispositions physiques. Mais l'idée est foncièrement la même.

La douleur est un état mental qui se prête assez facilement au projet analytique, car il possède au moins une expression assez distinctive et naturelle en termes de comportement. D'autres états avec le même type d'expression naturelle pourraient comporter des émotions (tristesse, peur, colère et joie ont toutes des manifestations typiques en termes de comportement). Mais d'autres états mentaux ne se rattachent que très indirectement au comportement : voyez le goût du café, par exemple. Goûter le café procure une expérience singulière. Goûter le café ressemble pour nous à quelque chose (mais pas pour le zombie). Mais, typiquement, ça ne nous fait pas grand-chose. Les penseurs contemporains aiment à le constater en disant qu'il y a des *qualia*, des sensations brutes associées au café que l'on goûte. Et les amis des *qualia* sont assez sombres quant aux chances de réduire lesdits *qualia* à des dispositions comportementales. En fait, ils en reviennent à Locke. Ces *qualia* se surajoutent à divers événements physiques – dans mon cas, sinon dans le vôtre –, mais il aurait pu en aller autrement. Mais alors le scepticisme – peut-être êtes-vous un zombie ou un mutant – menace encore.

UN MODÈLE SCIENTIFIQUE

Il est important de remarquer une distinction dont le débat contemporain est friand. Jusqu'ici, nous avons présenté un Leibniz hostile à l'élément

de fortuité brute chez Locke au nom d'une relation rationnelle quasi mathématique entre le corps et l'esprit. On peut suggérer une voie médiane, qui s'oppose à la fortuité sans aller pour autant à une relation mathématique ou rationnellement transparente. On la formule habituellement en disant qu'il y a peut-être une identité *métaphysique* entre des faits ou événements mentaux et physiques, mais que ce n'en est pas forcément une qu'on puisse connaître *a priori*.

En voici une analogie courante. La physique classique identifie la température d'un gaz aux énergies cinétiques moyennes des molécules qui le composent. Ainsi, pour faire des gaz chauds, Dieu n'a qu'une chose à fixer : fixer le gaz et l'énergie cinétique moyenne de ses molécules, et ce faisant fixer la température. Il n'y a pas de variation indépendante. Il ne saurait y avoir de gaz zombie ou mutant, dans lequel l'énergie cinétique des molécules ne produit aucune température ou produit des températures différentes de celles associées à la même énergie dans d'autres gaz.

D'un autre côté, ce n'est pas *simplement* la raison, la pensée ou les mathématiques qui ont permis aux scientifiques d'assimiler la température à l'énergie cinétique moyenne. Il ne s'agit pas d'une percée *a priori*, d'une analyse en chambre de ce qu'on entend par température, mais du fruit d'expériences et d'observations ainsi que de considérations *théoriques* générales. Loin d'être purement *a priori*, ce résultat est au moins pour l'essentiel *a posteriori*. La relation n'est pas de celle qui pourrait être mise en évidence à l'avance par les mathématiques ou par des « idées claires et distinctes »,

comme le fait qu'un cercle sur un plan incliné projette une ellipse.

En général, dans les sciences, quand un terme ou une propriété théorique comme la température est identifié à un autre (ici l'énergie cinétique moyenne des molécules qui composent le gaz), le lien est assuré par des principes « passerelles » qui font partie des théories des sciences en question. Ainsi, par exemple, l'identification actuelle de gènes à des morceaux d'ADN se fait parce que, dans la biologie classique, les gènes sont définis par leur fonction dans l'héritabilité des caractères, et que la biologie moléculaire montre aujourd'hui que ce sont précisément les brins d'ADN qui ont cette fonction. Observez que l'analyse n'est pas *entièrement* absente. Il nous faut savoir ce que les gènes sont censés faire avant de pouvoir poser l'équation. Mais la grande découverte est la découverte scientifique contingente de ce qui fait ce qu'ils sont censés faire par définition.

Si, dans notre approche du problème esprit-cerveau, nous prenions modèle sur le type de réductions scientifiques décrit à l'instant, nous trouverions un état physique caractéristique des personnes partageant quelque état mental. Ainsi, par exemple, pourrions-nous découvrir que tous les gens qui ont mal, et eux seulement, partagent un état cérébral donné (on dit, d'une formule vague, qu'on a les fibres C à vif). Et l'on proposerait ensuite que c'est cela avoir mal, de même que certains bouts d'ADN sont des gènes. Une fois encore, on aurait une réduction complète du mental au physique.

Ce serait ce qu'on appelle une théorie de l'identité psychophysique.

Ses détracteurs affirment parfois qu'on ne peut croire cette théorie qu'en feignant une anesthésie permanente. Ils déplorent que soit laissé de côté tout ce qui est proprement mental. La bonne réfutation consiste à demander au contestataire ce qui, à son avis, a été négligé et de l'observer ensuite aux prises avec les difficultés du dualisme. Mais ce type de théorie de l'identité psychophysique crée d'autres difficultés. L'une est que, dans le cas des événements mentaux, c'est sa propre conscience qui tranche de la façon suivante. Dans la perspective du sujet, tout ce qui est ressenti comme douleur est douleur. Peu importe que ce soient les extrémités nerveuses ou tout autre chose. Si quelqu'un a subi une transplantation, où des fibres C organiques ont été remplacées par du silicium, lequel produit néanmoins les mêmes résultats, c'est encore de la douleur. Notre connaissance de la douleur n'est pas tributaire de la présence en nous de fibres C ou d'autres systèmes de génie biologique. Il existe une autorité de la première personne. De même, bien que nous ne puissions savoir si des candidats marginaux à l'expérience de la douleur, tels que la crevette, possèdent ou non des fibres C, nous pourrions être gênés de dire que, pour cette seule raison, ils ressentent ou non la douleur. L'identité ne paraît donc pas tout à fait aussi directe que dans d'autres cas scientifiques (ce qui pourrait être contesté).

Nous serions assez satisfaits si nous pouvions voir la relation entre les événements mentaux et les événements du cerveau et du corps aussi clairement

que nous pouvons voir la relation entre la température et l'énergie cinétique moyenne des gaz. Peut-être ne serait-il pas très important pour nous que le résultat ait été obtenu par la « pensée pure » davantage que par l'expérience. Nous pouvons donc apprécier l'objection de Leibniz à Locke sans partager totalement son rationalisme. Reste que, quand nous réfléchissons sérieusement à la relation entre le cerveau et le corps, d'un côté, l'esprit, de l'autre, il apparaît généralement que c'est moins notre pensée que notre simple ignorance scientifique qui nous laisse tomber. Dernièrement, de nombreux hommes de science se sont penchés sur la conscience et ont identifié toutes sortes d'états cérébraux impliqués dans le fonctionnement normal de la conscience. Par exemple, on a estimé vitales les ondes électromagnétiques d'une fréquence particulièrement basse dans le cerveau. Mais on ne voit pas très bien comment ce type de vérité peut contribuer à résoudre le problème – à nous permettre de prendre le parti de Leibniz contre Locke. Du point de vue lockéen, tout ce que l'homme de science a pu découvrir c'est que, *quand* le cerveau est dans un état spécifique, nous avons des symptômes de conscience. Mais *cela* pourrait tout aussi bien nous dire ce à quoi la conscience est annexée, de manière purement fortuite. La combinaison n'en devient pas intelligible pour autant. Et cela présuppose aussi le droit d'écarter de notre champ de préoccupation les possibilités du zombie et du mutant, sans quoi l'homme de science ne pourrait jamais établir la corrélation, si ce n'est, au mieux, dans son propre cas. Or, selon les « nouveaux mystériens », ni la

philosophie ni la science ne nous conduiront jamais plus loin. Jamais nous ne serons en mesure de nous ranger sans réserve avec Leibniz contre Locke.

SPECTRES INVERSÉS : LANGAGES PRIVÉS

Le cas de la couleur paraît spécialement laisser grande ouverte la possibilité au moins de mutants : de gens physiquement identiques, mais qui n'en perçoivent pas moins les couleurs tout autrement. Il pourrait même y avoir des mutants dont les spectres de couleur sont entièrement inversés les uns par rapport aux autres, au point que l'expérience que procure à l'un la lumière à l'extrémité rouge du spectre est l'expérience même que fait l'autre à l'extrémité bleue. Et il n'y aurait rien pour leur indiquer que c'est ainsi.

Le dualisme cartésien ouvre la possibilité de zombies et de mutants. Mais peut-être ouvre-t-il aussi une possibilité plus effrayante encore. Si nous croyons à la voie dualiste, nous pouvons être assurés au moins de savoir ce à quoi ressemble notre expérience *propre*. Sans doute l'esprit des autres est-il un peu conjectural, mais notre esprit nous est bien connu. Mais est-ce bien vrai ? Considérez non pas l'esprit des autres, mais *votre propre expérience passée*. Êtes-vous sûr que le monde vous paraît de la même couleur qu'hier ? Êtes-vous sûr qu'il vous semblait hier d'une couleur quelconque – autrement dit, que vous ayez bien fait l'expérience consciente que vous vous souvenez avoir faite ?

En posant ces questions, vous appliquez les pos-

sibilités du zombie et du mutant à votre propre passé. À première vue, bien entendu, les possibilités sont encore plus incongrues et absurdes qu'appliquées à d'autres esprits. Et nous sommes enclins à répondre que, bien entendu, nous savons parfaitement que les couleurs étaient largement les mêmes hier qu'aujourd'hui. Si, au réveil, le ciel avait aujourd'hui la couleur de l'herbe hier et inversement, nous ne manquerions pas de le remarquer.

J'accorde, naturellement, que nous remarquerions le changement. Mais, compte tenu du dualisme cartésien, cette sécurité est-elle garantie ? Tout dépend de notre façon d'envisager la mémoire et les événements mentaux. D'où vient notre certitude que les événements mentaux – jugés entièrement distincts, ne l'oublions pas, de tout ce qui est physique – laissent des traces fiables dans la mémoire ? Je puis *vérifier* que ma mémoire du monde physique est assez fiable. Je me souviens d'avoir mis la voiture au garage et le fait est que, lorsque je descends, elle y est. Je me souviens du chemin de la cuisine et, de fait, je m'y rends sans effort et sans la moindre erreur. Mais qu'est-ce qui vérifierait que ma mémoire du monde mental est exacte ? Dans les termes de Locke, pourquoi le « bon plaisir de Dieu » ne serait-il pas de m'annexer aujourd'hui certaines modifications mentales, en même temps que le souvenir trompeur que de semblables souvenirs m'étaient annexés hier ? Telle était la suggestion de Wittgenstein :

Débarrassez-vous toujours de l'idée d'un objet privé en supposant qu'il change constamment ;

*ce que vous ne remarquez point, parce que votre mémoire vous abuse constamment*²⁵.

C'est là le cœur de l'argument « anti-langue privée » de ses *Investigations philosophiques* (publiées à titre posthume en 1953) – l'un des arguments les plus célèbres de la philosophie du xx^e siècle. Wittgenstein essayait de montrer qu'il ne saurait y avoir de *pensée* significative sur la nature de sa vie mentale passée (ou future) si cette vie mentale est dissociée du monde physique ainsi que le propose le dualisme cartésien. Elle devient, pour ainsi dire, trop glissante ou même spectrale pour être un objet de nos souvenirs ou de nos intentions.

Appliquées à nos passés, les possibilités du mutant et du zombie sont certainement déroutantes. Mais, en vérité, elles ne devraient nous dérouter qu'à propos de l'image dualiste. Une fois de plus, pouvons-nous reculer de Locke vers quelque version leibnizienne ? Leibniz, rappelons-le, veut qu'il y ait une relation « rationnelle » entre le physique et le mental en sorte que l'événement mental qu'est la vision d'une couleur est une sorte d'*expression* de ce qui se passe physiquement, non pas une annexion accidentelle. Comment cela pourrait-il marcher dans le cas des couleurs ? L'idée leibnizienne est que si moi et mon jumeau (qui pourrait bien être moi-même tel que j'étais hier) fonctionnons physiquement de la même manière, il n'y a aucune possibilité que nos vies mentales soient différentes. Comment étayer cette suggestion ? Voici une esquisse de réponse.

Nombre des changements physiques sous-jacents à la perception de la couleur sont assez bien

compris²⁶. La perception de la couleur est le résultat de la stimulation des cônes qui remplissent la partie centrale de la rétine. La meilleure théorie actuelle suggère qu'il en est trois sortes, les cônes L, M et S (à longueur d'ondes longue, moyenne et courte). Les cônes L se « hérissent » ou envoient des message au nerf optique d'autant plus facilement qu'ils sont touchés par une lumière de longueur d'onde plus longue ; les cônes M sont davantage excités avec une lumière de longueur d'ondes moyenne ; et les cônes S avec une longueur courte. La couleur que nous percevons dépend donc d'abord d'une comparaison entre les niveaux d'excitation de ces trois sortes de cônes. Ainsi, par exemple, si S est beaucoup plus excité que L, c'est le code du bleu, la couleur située à l'extrémité des ondes courtes du spectre. Si L est plus excité que M, nous avons le rouge, et si M est plus excité que L, nous avons du vert. Tout se passe comme si les circuits étaient des « adversaires » et que le résultat dépendait de celui d'entre eux qui l'emporte sur les autres.

Songons maintenant que les couleurs ont quantité de propriétés intéressantes. En voici quelques-unes : on ne peut avoir une surface qui soit bleue tirant sur le jaune, de même qu'on ne peut en avoir une verte tirant sur le rouge. En revanche, on peut avoir des surfaces vertes tirant sur le bleu ou rouges tirant sur le jaune (orange). On ne peut avoir de brun vif ni de gris vif (il est difficile d'imaginer une flamme grise ou brune). Le jaune est une couleur plus claire que le violet. On peut avoir une gemme rouge, bleue ou verte transparente, mais on ne peut avoir de gemme blanche transparente – le

plus proche serait un blanc laiteux, comme une opale. On peut avoir une lumière blanche, mais pas une lumière noire.

Tout cela pourrait faire figure de faits bruts dans le royaume cartésien de l'esprit, où les couleurs sont censées avoir leur résidence. Mais nous pouvons commencer à y voir les expressions de divers faits physiques. On ne saurait voir une surface bleu-jaune, parce que le jaune et le bleu sont produits par des opposés mathématiques : nous avons du jaune quand $L > S$, et du bleu quand $S > L$. De même pour le rouge et le vert. On ne peut avoir du brun clair parce que le brun est du jaune assombri. Une surface est vue brune quand elle serait codée pour le jaune, sauf qu'il n'y a qu'un faible niveau énergétique général en comparaison de celui d'autres sources de lumière dans le contexte. De même pour le gris, qui est du blanc assombri. Le jaune est plus clair que le violet parce que la lumière jaune ($L > S$) est aussi plus proche de la fréquence à laquelle nos systèmes visuels ont leur sensibilité maximale. En comparaison le rouge, à une extrémité du spectre visuel, et le bleu, à l'autre, nous conduisent vers l'obscurité, où nous ne saurions avoir la moindre réponse. On ne saurait avoir de blanc transparent, car on ne voit un objet blanc que quand il diffuse la lumière.

Tout cela, bien entendu, ne fait qu'effleurer la surface de la science des couleurs. Mais au moins a-t-on un aperçu de la manière dont les choses « prennent sens ». Avec suffisamment de faits de ce type devant nous, nous pourrions être moins enchantés par la possibilité du spectre inversé. Commençons par le cas plus simple de la vision

monochromatique (noir et blanc). Soit la suggestion qu'un tel pourrait être mon double physique, mais voit foncé quand je vois clair, et inversement. Est-ce possible ? Instantanément, nous pourrions répondre que oui. Peut-être imaginons-nous que le monde lui apparaît tel qu'il apparaît sur un négatif photographique. Mais ça ne marche pas vraiment. Si j'éclaircis un bout de verre gris, je vois mieux à travers ; si je le vois plus foncé, je vois moins bien à travers lui. Puisqu'il est un double physique, cela doit être vrai de mon jumeau. Mais, pour lui, quand nous éclaircissons le verre, il lui « semble » que nous avons ajouté de la suie, puisqu'il devient subjectivement plus sombre. Et, inversement, nous ajoutons de la suie et il lui « semble » qu'il devient plus clair. Mais il nous faut alors imaginer que, pour lui, plus la vitre s'assombrit, mieux il voit à travers elle, tandis qu'il y voit de plus en plus mal à mesure qu'elle devient plus claire. Et c'est précisément cela qui paraît dénué de sens, qui ne marque pas une possibilité cohérente.

Prenons maintenant quelqu'un de physiquement identique à moi, mais qui soi-disant voit jaune quand je vois bleu, et inversement. Il n'est plus tout à fait aussi facile de l'imaginer. Il doit répondre de la même façon que moi, si bien qu'il ne peut clarifier que le jaune est une couleur foncée, par exemple. Cette différence de réponse et de comportement serait une différence physique. Il nous faut donc demander *comment* il voit le bleu comme une couleur vive et le jaune comme une couleur foncée. S'il voit vraiment le jaune ainsi, comme moi je vois le bleu, comment voit-il le brun ? Et l'orange ? Le brun est un jaune foncé, mais pour lui le jaune est

déjà sombre. Il est donc difficile d'imaginer comment ses discriminations physiques pourraient s'accorder aux miennes, compte tenu de cette totale disparité d'expérience mentale.

Bref, la possibilité devient beaucoup moins claire et nous pouvons nous sentir habilités à nier que ce soit une possibilité. Nous pourrions concocter une conception de l'esprit qui comble la brèche entre le physique et le mental, c'est-à-dire entre le système visuel du cerveau qui fonctionne et réagit parfaitement et les *qualia* « subjectifs » apparemment surajoutés de l'expérience des couleurs. Cette construction serait une justification de la position de Leibniz. L'expérience subjective de la couleur devient non pas un simple ajout bizarre, mais l'*expression* inévitable, rationnellement explicable des types de fonctionnement physique des créatures que nous sommes. Si on peut en faire autant pour tous les éléments de notre conscience, le problème est résolu.

PENSÉE

Nous en arrivons maintenant à un aspect légèrement différent de la conscience. Ce chapitre s'est concentré sur les sensations et les *qualia*. Or notre conscience se compose largement de pensées. Étranges choses que les « pensées ». Elles ont des pouvoirs de « représentation » : une pensée représente typiquement le monde comme étant d'une manière ou d'une autre. Une sensation, en revanche, paraît simplement se trouver là. Apparemment, elle n'indique rien au-delà d'elle-même

– un fait, par exemple, ou un fait présumé. (Il est des penseurs qui le nient. Ils croient, par exemple, qu'une sensation de douleur est une perception de blessure corporelle, et que cette perception représente le corps blessé, tout comme la pensée que demain c'est vendredi représente demain comme étant vendredi. Je laisse au lecteur le soin de réfléchir à la vraisemblance de cette idée.) La nature représentative des pensées – leur intentionnalité ou le caractère dirigé, comme on dit parfois – est en soi très déroutante. Si nous imaginons les pensées comme des sortes de « choses » présentes dans la conscience, la question devient : comment une « chose » peut en elle-même et d'elle-même indiquer autre chose (un fait ou un état de choses). Un panneau de signalisation, par exemple, peut assurément indiquer un village. Mais, c'est apparemment une question de chemin à suivre. Un panneau ne représente pas en soi et de lui-même la route du village. Encore faut-il savoir la prendre. Nous pourrions imaginer une culture dans laquelle le même objet physique qu'est pour nous un panneau de signalisation aurait une fonction toute différente : un panneau d'affichage, un totem, ou une œuvre d'art abstraite. C'est ce que nous voyons avec les animaux : si vous montrez du doigt quelque chose, le chien, typiquement, ne prête attention qu'au doigt tendu, à la plus grande irritation de son propriétaire. En revanche, il paraît incohérent d'imaginer une créature avec les mêmes *pensées* que nous, mais qui n'a pas appris à « prendre » ces pensées de la même façon que nous. C'est précisément ce « prendre » qui *fait* la pensée.

Probablement la bonne réaction consiste-t-elle à

nier que les pensées soient des choses. On donne parfois le nom de *réification* à l'erreur consistant à supposer qu'à chaque nom correspond une « chose ». Les penseurs s'accusent souvent les uns les autres de réification déplacée. Ce sont les gens qui pensent, et qu'ils le fassent n'est pas l'affaire d'une sorte de tache présente dans le cerveau ou dans l'esprit. C'est vrai même si la tache en question est conçue comme une petite phrase écrite dans le cerveau. La pensée consiste à prendre le monde d'une manière ou d'une autre, et la manière dont on le prend dépend de nos dispositions, plutôt que de l'agencement des choses en nous.

Que nous puissions penser à des états de choses absents – des états lointains, des états passés ou futurs – ne devrait peut-être pas nous dérouter davantage que le fait de pouvoir prêter attention au monde. C'est pourtant un tour de force qui nous sépare des autres animaux. Vraisemblablement peuvent-ils percevoir le monde, mais nous avons du mal à concevoir qu'ils puissent se représenter des états de choses lointains, passés et futurs. Nous, pourtant, nous en sommes certainement capables.

L'approche actuelle la plus populaire consiste à se concentrer sur la manière dont nous pouvons attribuer des pensées à la personne qui fonctionne bien. Quelque chose dans la conduite d'une personne doit nous permettre d'en déduire qu'elle pense à hier ou se concentre sur le temps prédit pour le week-end. Les pensées s'expriment dans un comportement tant linguistique que non linguistique, et peut-être pouvons-nous espérer une sorte de réduction : « X pense que *p* » si et seulement si les plans, les désirs ou le comportement de X sont

d'une certaine façon en accord avec le monde étant tel que *p*. Tout le problème serait de donner un contenu à ce « d'une certaine façon en accord ». On peut dire en toute justice que personne n'y est parvenu. Mais il est des suggestions sur la manière de s'y prendre. Nous disons qu'un système intelligent, un missile guidé par exemple, pense qu'il y a un avion à un mille de distance et à deux cents pieds de haut si ses systèmes le dirigent dans une direction *appropriée* à la présence d'un avion à cet endroit – compte tenu de son but, ou de sa fonction, qui est d'abattre les avions. De même pourrions-nous dire d'une personne qu'elle pense qu'il fera beau ce week-end si son comportement est approprié, compte tenu de ses buts (ou de ses fonctions) au temps de ce week-end. La difficulté serait de donner un contenu à cette pensée sans s'appuyer, par d'autres voies, sur d'autres états mentaux du sujet – ce que, précisément, personne ne sait faire.

Pour l'heure, je laisse de côté la pensée. Dans les deux chapitres qui suivent, j'examine plutôt deux autres éléments de notre vision du monde qui nourrissent également le dualisme cartésien. Le premier est un éventail de pensées sur notre propre liberté. Le second, un éventail de pensées sur notre propre identité.