

Où suis-je ?

Maintenant que j'ai gagné mon procès en vertu de la Loi sur la Liberté de l'Information, j'ai pour la première fois le loisir de révéler un curieux épisode de ma vie qui pourra intéresser non seulement les spécialistes de la philosophie de l'esprit, de l'intelligence artificielle et de la neuroscience, mais aussi le grand public.

Il y a plusieurs années, j'ai été contacté par des officiels du Pentagone qui m'ont demandé de me porter volontaire pour une mission secrète hautement dangereuse. En collaboration avec la NASA et Howard Hughes, le ministère de la Défense dépensait des milliards pour mettre au point un Module d'Extraction et de Creusement, ou MEC, qui devait percer la Terre et aller, à grande vitesse, déposer une tête nucléaire spéciale « au pied des silos de missiles des Rouges », comme me l'expliqua un des grands manitous du Pentagone.

Le problème, c'est que lors d'un des premiers tests, ils avaient réussi à coincer une ogive à un kilomètre de profondeur sous Tulsa, Oklahoma, et ils voulaient que j'aille la leur récupérer. « Pourquoi moi ? », ai-je demandé. Eh bien, la mission avait aussi des applications d'avant-garde dans le domaine de la recherche sur le cerveau, et ils avaient entendu parler de mon intérêt pour le cerveau et, bien entendu, de ma curiosité faustienne, de mon grand courage, etc. Alors, comment pouvais-je refuser ? La difficulté qui avait poussé le Pentagone à frapper à ma porte était que l'engin que je devais aller chercher était, d'une nouvelle façon, féroce-ment radioactif. D'après les appareils de mesure, quelque chose dans la nature de l'engin et de ses complexes interactions avec des poches

dé matière au plus profond de la Terre avait engendré des rayonnements qui pouvaient entraîner de graves lésions de certains tissus du cerveau. Aucun moyen n'avait été trouvé pour protéger le cerveau contre ces rayons mortels, qui étaient apparemment inoffensifs pour les autres tissus et organes du corps. Aussi avait-il été décidé que la personne qui irait récupérer l'ogive partirait sans son cerveau. Celui-ci serait conservé en sûreté en un lieu où il pourrait exercer ses fonctions de commande normales par l'intermédiaire de liaisons radio perfectionnées. Allais-je accepter de subir une intervention chirurgicale au cours de laquelle on extrairait totalement mon cerveau pour le placer dans un système de survie au Centre des Satellites Habités de Houston? Au fur et à mesure qu'ils seraient sectionnés, chacun des canaux d'entrée et de sortie seraient rétablis à l'aide d'une paire d'émetteurs-récepteurs radio miniaturisés dont l'un serait connecté au cerveau, et l'autre aux queues neurales du crâne vide. Aucune information ne se perdrait, la liaison serait entièrement préservée. Au début, j'étais un peu réticent. Est-ce que ça marcherait vraiment? Les neurochirurgiens de Houston m'ont encouragé : « Dites-vous, m'ont-ils conseillé, qu'il ne s'agit que d'un étirement des nerfs. Si votre cerveau était déplacé de quelques centimètres dans votre crâne, il n'y aurait aucune répercussion au niveau de votre esprit. Eh bien, nous allons simplement rendre les nerfs indéfiniment élastiques en y soudant des petits appareils de liaison radio. »

On m'a alors fait visiter le laboratoire de survie de Houston, et j'ai pu voir la cuve flambant neuve dans laquelle serait placé mon cerveau si j'acceptais. J'ai rencontré la nombreuse et brillante équipe de neurologues, hématologues, biophysiciens et ingénieurs en électricité, et au terme de plusieurs jours de discussions et de démonstrations, je leur ai donné le feu vert. J'ai alors été soumis à une kyrielle de tests sanguins, tomographies cérébrales, expériences, entretiens et autres. Ma biographie a été rédigée en long, en large et en travers, et d'interminables listes de mes croyances, de mes espoirs, de mes craintes et de mes goûts ont été dressées. Même mes enregistrements stéréo favoris ont été notés, et j'ai eu droit à une psychanalyse accélérée.

Enfin, le jour de l'opération est arrivé, et j'ai bien entendu été anesthésié, ce qui fait que je n'ai aucun souvenir de l'opération même. Quand je suis sorti du sommeil, j'ai ouvert les yeux, regardé tout autour de moi et posé l'inévitable et lamentablement banale question postopératoire : « Où suis-je ? » L'infirmière m'a souri : « Vous êtes à Houston », et j'ai réfléchi que dans tous les cas, il y avait de grandes chances que ce soit vrai. Elle m'a alors tendu un miroir. Oui, elles étaient bien là, les petites antennes pointant à travers leurs prises de titane scellées dans mon crâne.

« L'opération semble avoir réussi, ai-je dit, je voudrais voir mon cerveau. » Ils m'ont emmené (j'avais un peu le vertige et mon équilibre était incertain), à travers un long couloir, jusqu'au laboratoire de survie. Là, je fus salué par une acclamation de l'équipe d'assistance réunie, à laquelle je répondis par un petit signe que j'espérais désinvolte. Encore un peu étourdi, je me fis aider pour aller jusqu'à la cuve. Je scrutai son

contenu à travers ses parois de verre. Là, sur ce qui ressemblait à de la limonade, flottait indéniablement un cerveau humain, quoiqu'il fût presque entièrement couvert de circuits imprimés, de tubulures en plastique, d'électrodes et de tout un attirail compliqué. « C'est le mien ? » demandai-je. « Appuyez sur l'interrupteur de l'émetteur de sortie, sur le côté de la cuve, et vous verrez bien », répondit le directeur de projet. Je basculai l'interrupteur sur ARRET et m'affalai immédiatement, les jambes molles et l'estomac retourné, dans les bras des techniciens, dont l'un remit gentiment l'interrupteur sur la position MARCHE. Et pendant que je retrouvais mon équilibre et mon sang-froid, je me dis : « Je suis assis ici, sur une chaise pliante, en train de regarder mon cerveau à travers une vitre... Mais au fait, m'interrompis-je, n'aurais-je pas dû penser : "Me voici, flottant sur un liquide en ébullition, dévisagé par mes propres yeux?" » J'essayai de penser cette deuxième pensée. J'essayai de la projeter dans la cuve, comme pour la tendre à mon cerveau, mais je ne réussis pas à effectuer cet exercice avec une conviction suffisante. Encore un essai. « Me voici, moi, Daniel Dennett, flottant sur un liquide en ébullition, dévisagé par mes propres yeux. » Non, ça ne marchait pas. C'était pour le moins déconcertant. En tant que philosophe et physicaliste convaincu, je croyais dur comme fer que le processus concret de mes pensées se déroulait quelque part dans mon cerveau. Et pourtant, quand je pensais « Me voici », c'était ici, hors de la cuve, que surgissait la pensée, là où moi, Dennett, j'étais assis à observer mon cerveau.

J'essayai plusieurs fois de me faire penser dans la cuve, mais en vain. Je tentai de me préparer à cette tâche en faisant des exercices mentaux. Je me dis : « Le soleil brille là-bas », cinq fois de suite, en pensant chaque fois à un lieu différent : dans l'ordre, le coin ensoleillé du laboratoire, la pelouse d'entrée de l'hôpital, Houston, Mars et Jupiter. Je me rendis compte que je n'avais pas trop de mal à faire sauter mes « là-bas » jusqu'à leurs différentes destinations de la carte céleste. En un instant, je pouvais lancer un « là-bas » aux confins de l'espace, puis cibler le « là-bas » suivant, avec une extrême précision, sur le quart supérieur gauche d'une tache de rousseur sur mon bras. Mais pourquoi avais-je des problèmes avec « ici »? « Ici, à Houston », marchait assez bien, de même qu'« ici, dans le laboratoire », et même « ici, dans cette partie du laboratoire », mais « ici, dans la cuve » persistait à rester une expression mentale sans effet. J'essayai de la penser en fermant les yeux. Ça avait l'air mieux, mais je ne pouvais toujours pas la détacher de moi, sauf peut-être pendant un imperceptible instant. Je ne pouvais pas être sûr. La découverte de cette impossibilité d'être sûr m'ébranla. Comment savais-je quel lieu était désigné par « ici » quand je pensais « ici »? Pouvais-je croire que je désignais un lieu donné alors qu'en fait, je voulais parler d'un autre? Je ne voyais pas comment on pouvait admettre une telle idée sans couper les quelques liens intimes entre une personne et sa vie mentale qui avaient survécu au massacre des spécialistes et philosophes du cerveau, des physicalistes et des behavioristes. Peut-être étais-je incorrigible quant au lieu que je désignais en disant « ici ». Mais dans les circonstances présentes, il me semblait que j'étais

contraint par la force de l'habitude mentale à avoir des pensées désignatives systématiquement fausses, ou alors que l'endroit où se trouve une personne (et où se forment ses pensées du point de vue de l'analyse sémantique) n'est pas forcément celui où réside son cerveau, le siège physique de son âme. Troublé, j'essayai alors de m'orienter en recourant à une vieille astuce de philosophe : je commençai à donner des noms à tout.

« Yorick, dis-je à voix haute à mon cerveau, tu es mon cerveau. Le reste de mon corps, qui est assis sur cette chaise, s'appelle désormais Hamlet. » Nous sommes donc tous là : Yorick, mon cerveau, Hamlet, mon corps, et moi, Dennett. Mais où suis-je ? Et quand je pense « où suis-je », où se forme cette pensée ? Dans mon cerveau, lequel se prélassait dans la cuve, ou ici, entre mes oreilles, où elle semble s'être formée ? Ou nulle part ? Je n'ai aucun problème avec ses coordonnées temporelles, alors, pourquoi n'aurait-elle pas aussi des coordonnées spatiales ? Je commençai alors à faire une liste des possibilités.

1. *Où Hamlet va, Dennett va.* Il était facile de réfuter ce principe sur la base des expériences mentales de transplantation de cerveau bien connues des philosophes. Si Jean et Jacques échangent leurs cerveaux, Jean est celui qui a l'ancien corps de Jacques ; il suffit d'ailleurs de le lui demander : il affirmera qu'il est Jean et vous racontera bien les plus petits détails de la vie de Jean. Il était clair que mon corps et moi pouvions nous séparer, mais je ne pourrais probablement pas fausser compagnie à mon cerveau. La règle empirique se dégageant nettement de ces expériences mentales était que dans une opération de transplantation de cerveau, on veut être le donneur, et non pas le receveur. Il vaudrait mieux parler de transplantation de corps, en fait. Alors, la vérité, c'était peut-être plutôt que,

2. *Où Yorick va, Dennett va.* Ça n'avait rien d'attrayant. Comment pouvais-je être dans la cuve, sans moyen d'aller où que ce soit, alors qu'il était évident que j'étais hors de celle-ci, à la regarder et à commencer à envisager honteusement de regagner ma chambre pour y prendre un déjeuner copieux ? Je voyais bien que j'esquivais le fond du problème, mais il me semblait tout de même que je touchais là quelque chose d'important. Cherchant à étayer mon intuition, je trouvai un argument d'un type légaliste qui n'aurait sans doute pas déplu à Locke.

Supposons, raisonnai-je, que je prenne maintenant l'avion pour la Californie, que j'y fasse un hold-up dans une banque, et que je sois arrêté. Dans quel Etat serais-je jugé : en Californie, où a eu lieu le vol, ou au Texas, où se trouvait le cerveau ayant monté cette expédition ? Serais-je un malfaiteur californien au cerveau interdit de séjour ou un malfaiteur texan téléguidant un complice fantôme ? Je pourrais peut-être m'en tirer grâce, justement, à l'indécidabilité de cette question de compétence juridictionnelle, mais on pourrait estimer qu'il s'agissait d'un délit concernant deux Etats, et donc fédéral. Supposons, de toute façon, que j'aie été condamné. La Californie se contenterait-elle de jeter Hamlet au trou en sachant Yorick en train de mener la belle vie et de prendre luxueusement les eaux au Texas ? Et le Texas, de son côté,

mettrait-il Yorick en prison en laissant Hamlet libre de prendre le prochain bateau pour Rio ? Cette dernière solution me plut : repoussant la peine capitale ou toute autre condamnation cruelle et inhabituelle, l'Etat serait dans l'obligation de maintenir le système de survie de Yorick, mais pourrait le transférer de Houston à Leavenworth, et abstraction faite de l'infamie subie, cette conclusion de l'affaire ne me dérangerait nullement et je me considérerais alors comme un homme libre. L'Etat trouve peut-être son compte en plaçant les gens dans des établissements pénitentiaires, mais il ne me placerait, moi, dans aucun établissement en y mettant Yorick. Mais alors, une troisième solution émergeait.

3. *Dennett est partout où il pense être.* En généralisant, cette thèse était la suivante : une personne a, à tout moment, un point de vue, et le lieu où il se situe (qui est déterminé intérieurement par le contenu dudit point de vue) est également celui de la personne.

Cette proposition ne va pas sans poser de problèmes, mais il me semblait que je progressais dans la bonne direction. Le seul ennui, c'est qu'elle semblait impliquer une situation d'in vraisemblable infaillibilité du type pile-je-gagne/face-tu-perds. Ne m'étais-je pas moi-même souvent trompé dans mon appréciation du lieu où j'étais, et n'avais-je pas en tout cas au moins eu des doutes à ce sujet ? Ne pouvait-on pas se perdre ? Bien sûr, mais il n'y a pas que géographiquement que l'on peut se perdre. Une personne qui se perdrait dans une forêt pourrait essayer de se rassurer en se disant qu'elle sait au moins où elle est : elle connaîtrait au moins le milieu familial qu'est son corps. Quoique dans ce cas, ce ne serait peut-être pas une bien grande aide. On peut pourtant imaginer bien des situations plus terribles, et je me demandai si je n'étais pas dans l'une d'elles.

Il était clair que le point de vue avait quelque chose à voir avec le lieu où l'on est physiquement, mais il s'agissait là d'une notion trouble. Il était évident que le contenu d'un point de vue n'était pas identique au contenu des croyances ou des pensées et n'était pas déterminé par ce dernier. Que dire, par exemple, du point de vue du spectateur qui, dans une salle de projection en cinérama, hurle et se tord sur son siège en voyant une séquence de montagnes russes qui annihile sa distanciation psychique ? A-t-il oublié qu'il est assis, en toute sécurité, dans la salle ? Dans ce cas précis, j'étais porté à penser qu'il était l'objet d'un déplacement illusoire de point de vue. Mais dans d'autres cas, mon inclination à qualifier de tels déplacements d'illusoires était moins forte. Ceux qui travaillent dans des laboratoires et des installations où l'on manipule des matières dangereuses par l'intermédiaire de bras et de mains mécaniques asservis subissent un déplacement de point de vue plus marqué que tout ce que peuvent provoquer les effets spéciaux au cinéma. Ils sentent le poids et les parois glissantes des conteneurs qu'ils tiennent de leurs doigts métalliques. Ils savent parfaitement bien où ils se trouvent et ne sont pas induits en erreur par l'expérience effectuée ; pourtant, c'est comme s'ils étaient à l'intérieur de la chambre d'isolement qu'ils observent. Un effort mental leur permet de faire aller et venir leur point de

vue, comme quelqu'un qui verrait un cube de Necker ou un dessin d'Escher sous une nouvelle perspective. Il semble bien extravagant de supposer qu'en effectuant ce petit exercice de gymnastique mentale, ils se font aller et venir, eux.

Leur exemple m'emplissait néanmoins d'espoir. Car si, en dépit de mes intuitions, j'étais bien dans la cuve, je pourrais peut-être, en m'entraînant, parvenir à adopter tout à fait couramment ce point de vue. Je devrais me concentrer sur des images de moi flottant confortablement dans ma cuve et envoyant des ordres à ce corps familial, là-bas, dehors. Je réfléchis que la facilité ou la difficulté de cette entreprise était sans doute indépendante de la vérité de la position du cerveau. Si je m'y étais entraîné avant l'opération, ce serait peut-être enfantin, maintenant. Vous devriez essayer vous-même un simulacre du genre suivant : imaginez que vous avez écrit une lettre provocatrice qui a été publiée dans le *New York Times*, à la suite de quoi les autorités ont décidé d'enfermer votre cerveau pendant une période de surveillance de trois ans dans une Clinique pour Cerveaux Dangereux à Bethesda, dans le Maryland. Votre corps, lui, est bien entendu libre de gagner sa vie et de continuer à accumuler un revenu soumis à l'impôt. Mais pour le moment, il est assis dans une salle de conférences, en train d'écouter le curieux récit d'une expérience similaire de Daniel Dennett. Essayez. Projetez-vous à Bethesda, puis revenez en soupirant d'aise vers votre corps, si lointain et pourtant apparemment si proche. Ce n'est que grâce à une grande télémaîtrise (la vôtre ou celle des autorités?) que vous pouvez réfréner votre impulsion de faire poliment applaudir ces mains avant de guider ce bon vieux corps jusqu'aux toilettes, puis de lui offrir un cognac bien mérité dans le salon. La tâche de l'imagination est ardue, mais si vous atteignez le but recherché, les résultats vous récompenseront peut-être de cet effort.

Enfin bref, j'étais là, à Houston, perdu dans mes pensées, comme on dit, mais pas pour longtemps. Mes divagations furent bientôt interrompues par les médecins de Houston qui désiraient tester ma toute nouvelle prothèse cérébrale avant de m'envoyer effectuer ma dangereuse mission. Comme je l'ai déjà dit, je ressentis d'abord un léger vertige, ce qui n'avait rien de surprenant, mais je m'habituai rapidement à ma nouvelle situation (qui était, après tout, pratiquement indifférenciable de l'ancienne). Mon accommodation n'était pourtant pas parfaite, et même maintenant, j'ai encore de légers problèmes de coordination. La lumière se déplace à une vitesse très élevée, mais finie, et quand mon cerveau et mon corps s'éloignent l'un de l'autre, la délicate interaction entre mes systèmes de commande est troublée par le décalage temporel. Tout comme il devient pratiquement impossible de parler en entendant sa propre voix en écho ou avec un léger retard, je suis presque incapable de suivre des yeux un objet mobile dès que mon cerveau et mon corps sont distants de plus de quelques kilomètres. La plupart du temps, cette faiblesse est à peine discernable, quoique je ne sois plus capable de frapper une balle de base-ball en lui donnant de l'effet avec la même sûreté que jadis. Bien entendu, il y a des contreparties. Si l'alcool a

toujours un goût aussi agréable qu'avant et continue de chauffer mon gosier en attaquant mon foie, je peux désormais en boire sans limites et sans être ivre le moins du monde, curiosité que certains de mes amis les plus proches ont peut-être bien remarquée (quoique j'aie parfois fait semblant d'être ivre afin de ne pas attirer l'attention sur ma particularité). Pour des raisons identiques, je prends un cachet d'aspirine quand je me foule le poignet, mais si la douleur persiste, j'appelle Houston pour que l'on m'administre de la codéine *in vitro*. Dans les périodes de maladie, mes notes de téléphone peuvent atteindre des montants astronomiques.

Mais revenons-en à mon aventure. A la fin, les docteurs et moi avons estimé que j'étais prêt pour ma mission souterraine. Je laissai donc mon cerveau à Houston et me rendis à Tulsa en hélicoptère. Enfin, c'est ainsi que je ressentis les choses, et c'est ainsi que j'ai tendance, spontanément, à les décrire. Durant le voyage, je réfléchis encore à mes angoisses passées et décidai que mes premières conjectures postopératoires avaient été empreintes de panique. L'affaire était loin d'être aussi étrange ou métaphysique que je l'avais imaginé. Où étais-je? De toute évidence, à deux endroits : à la fois à l'intérieur et à l'extérieur de la cuve. Tout comme on peut avoir un pied dans le Connecticut et l'autre dans le Massachusetts, j'étais à deux endroits à la fois. J'étais devenu un de ces individus éclatés dont on entendait tant parler. Et plus je pensais à cette réponse, plus elle me semblait limpide. Pourtant, aussi étrange que ça puisse paraître, plus elle semblait juste, moins la question qu'elle résolvait peut-être paraissait importante. Triste destin, quoique non sans précédents, pour un problème philosophique. Mais cette réponse ne me satisfaisait évidemment pas totalement. Il restait une question irrésolue, qui n'était ni « Où sont mes différentes parties? » ni « Quel est mon point de vue actuel? » Il semblait, en tout cas, subsister une interrogation de ce genre. Car il paraissait vraiment indéniable que, d'une certaine façon, c'était moi, et pas seulement la plus grande partie de moi, qui allais m'enfoncer dans la terre sous Tulsa à la recherche d'une ogive atomique.

Quand je trouvai cette ogive, je me réjouis d'être venu sans mon cerveau, car l'aiguille du compteur Geiger spécial que j'avais emporté était totalement assolée. J'appelai Houston sur ma radio, et informai le centre dirigeant l'opération de ma position et du stade auquel j'en étais. En réponse, les techniciens m'indiquèrent, d'après mes observations sur place, comment démonter l'engin. Je m'étais mis au travail avec mon chalumeau quand, tout à coup, il se produisit quelque chose de terrible. Je devins complètement sourd. Au début, je pensai que c'était mon casque radio qui était tombé en panne, mais quand je le tapotais, je n'entendais rien. De toute évidence, les émetteurs-récepteurs auditifs m'avaient lâché. Je n'entendais plus ni Houston ni ma propre voix, mais je pouvais parler, ce qui me permit de commencer à leur raconter ce qui s'était passé ; au beau milieu de ma première phrase, je me rendis compte que quelque chose d'autre n'allait plus. Mon appareil vocal était paralysé. Ensuite ce fut ma main droite qui devint inerte — un autre

émetteur-récepteur était mort. J'étais vraiment dans de beaux draps. Et ce n'était pas encore le pire ! Quelques minutes plus tard, je devins aveugle. Je pestai contre mon manque de chance, puis contre les savants qui m'avaient mis dans un tel pétrin. J'étais là, sourd, muet et aveugle, dans un trou radioactif, à plus d'un kilomètre sous Tulsa. C'est alors que ma dernière liaison radio cérébrale rendit l'âme, me laissant tout à coup confronté à un nouveau problème encore plus effrayant : si un instant auparavant, j'étais enterré vivant dans l'Oklahoma, maintenant, j'étais à Houston, dépourvu de corps. Je ne compris pas tout de suite ma nouvelle situation. Il me fallut plusieurs minutes d'angoisse avant de me rendre compte que mon pauvre corps était à plusieurs centaines de kilomètres, son cœur battant et ses poumons respirant, mais à part ça tout aussi mort que le corps de n'importe quel donneur cardiaque, le cerveau truffé d'équipement électronique hors d'usage, inutile. Le changement de perspective qui, jusqu'ici, m'avait paru pratiquement impossible, me semblait maintenant parfaitement naturel. Je pouvais encore ramener mes pensées dans mon corps, sous Tulsa, mais l'illusion était assez difficile à soutenir. Car c'était certainement une illusion de supposer que j'étais encore en Oklahoma : j'avais perdu tout contact avec ce corps.

Je me rendis alors compte, dans un de ces éclairs de révélation dont nous devrions nous méfier, que j'avais mis le doigt sur une éclatante démonstration de l'immatérialité de l'âme telle qu'elle est vue par les physicalistes. En effet, quand le dernier signal radio entre Tulsa et Houston s'était évanoui, n'étais-je pas passé de Tulsa à Houston à la vitesse de la lumière ? Et cela sans aucune augmentation de masse ? Ce qui s'était déplacé de A à B à une telle allure, c'était certainement moi, ou en tout cas mon âme ou mon esprit, le centre immatériel de mon être, le siège de ma conscience. Mon point de vue était resté un peu en arrière, mais j'avais déjà remarqué l'influence indirecte du point de vue sur la situation physique. Je ne voyais pas comment un philosophe physicaliste pourrait le contester, si ce n'est en s'engageant dans la voie sinistre et anti-intuitive de l'interdiction de la notion de personne. Pourtant, elle était si enracinée dans notre vue du monde, ou tout au moins je le croyais, qu'il serait aussi peu convaincant et aussi peu sincère de la nier que d'affirmer, à l'encontre de Descartes, « *non sum* ».

La joie de la découverte philosophique m'aida à surmonter quelques minutes, ou peut-être quelques heures, très désagréables, mon impuissance et la détresse de ma situation me devenant de plus en plus apparentes. Des vagues de panique me parcoururent, suivies de nausées que l'absence de leur phénoménologie normale, liée au corps, rendait encore plus atroces. Pas de décharge d'adrénaline provoquant des picotements dans les bras, pas de palpitations, pas de salivation prémonitoire. J'eus bien, un court instant, une sensation de contraction de mes boyaux, ce qui me fit momentanément espérer un renversement de processus, une réincarnation progressive. Mais l'isolement et l'unicité de cet élanement me persuadèrent bientôt qu'il n'était que le premier signe

du mal du corps fantôme dont, comme tout amputé, j'allais très certainement souffrir.

Moralement, j'étais en pleine confusion. D'un côté, j'étais ravi de ma découverte philosophique et je me creusais la cervelle (une des rares choses habituelles que je pusse encore faire) à essayer de trouver comment je pourrais communiquer ma découverte à des revues techniques ; de l'autre, j'étais plein d'amertume, seul, rongé par la crainte et l'incertitude. Heureusement, ce malaise ne dura pas longtemps, car mon équipe d'assistance technique me plongea dans un sommeil sans rêves dont je m'éveillai aux accents familiers et merveilleusement fidèles de l'ouverture de mon trio pour piano favori de Brahms. C'était donc pour cela qu'ils avaient voulu connaître la liste de mes enregistrements préférés ! Il ne me fallut pas longtemps pour réaliser que j'entendais cette musique sans oreilles. La sortie du saphir de la chaîne stéréo était introduite directement dans mon nerf auditif par l'intermédiaire de quelque circuit de redressement complexe. Je me shootais littéralement à la musique de Brahms, expérience inoubliable pour tout mordu de la stéréo. A la fin du morceau, je ne fus nullement surpris d'entendre la voix rassurante du directeur du projet parlant dans un micro qui était maintenant ma prothèse auditive. Il confirma mon analyse de ce qui avait déraillé et m'assura que des mesures avaient été prises en vue de me redonner une enveloppe corporelle. Il n'entra pas dans le détail de ces mesures et, après quelques autres morceaux de musique, je glissai à nouveau vers le sommeil. D'après ce que j'appris ultérieurement, je dormis plus de six mois, et quand je m'éveillai, ce fut pour constater que j'avais retrouvé tous mes sens. Mais quand je me regardai dans une glace, je fus un peu surpris d'y voir un visage inconnu. Barbu et un peu plus empâté, ayant un air de famille certain avec l'ancien, et la même expression de vive intelligence et de résolution, mais sans nul doute un autre visage. Des explorations intimes plus poussées me menèrent à l'indubitable conclusion que c'était un nouveau corps, ce que me confirma le directeur de projet. Il n'offrit pas de me faire le récit de l'histoire de ce corps et je décidai (sagement, vu de maintenant) de ne pas chercher à en savoir plus. Comme de nombreux philosophes n'ayant pas entendu parler de mon histoire l'ont plus récemment spéculé, l'acquisition d'un nouveau corps laisse la personne intacte. Et après un certain temps d'adaptation à sa nouvelle voix, à de nouvelles capacités et faiblesses musculaires, etc., la personnalité s'avère également en grande partie préservée. On a bien souvent observé des changements de personnalité nettement plus prononcés chez des personnes physiquement transformées par des opérations de chirurgie esthétique, sans parler des changements de sexe, et nul ne semble contester la pérennité de la personne dans ces cas-là. Toujours est-il que je me fis rapidement à mon nouveau corps, au point de ne pas pouvoir retrouver ce qu'il avait de nouveau dans ma conscience ou ma mémoire. L'image qui apparaissait dans mon miroir me devint vite tout à fait familière. A noter qu'elle comportait toujours des antennes, ce qui explique que je ne fus nullement surpris d'apprendre que mon cerveau était encore dans son cocon, entre les murs du laboratoire de survie.

Je décidai que ce bon vieux Yorick méritait bien une visite. Mon nouveau corps, que nous appellerons Fortinbras, et moi-même, nous rendîmes donc à Houston, où nous fûmes une fois encore applaudis par les techniciens qui, en fait, se félicitaient eux-mêmes. Une fois de plus, j'étais devant la cuve, à contempler le pauvre Yorick, et la fantaisie me poussa à basculer à nouveau, avec une parfaite désinvolture, l'interrupteur de son émetteur de sortie. Imaginez un peu ma surprise quand force me fut de constater qu'il ne se passait rien de spécial ! Pas de défaillance, pas de nausées, aucun changement perceptible. Un technicien se précipita pour remettre l'interrupteur sur la position de marche, mais là encore, je ne sentis rien. J'exigeai alors des explications, que le directeur du projet s'empressa de me fournir. Il en ressortit qu'avant même la première opération, ils avaient réalisé un double informatique de mon cerveau, en reproduisant la totalité de sa structure de traitement des informations, mais aussi sa vitesse de calcul, dans un gigantesque programme. Après l'opération, mais avant de m'envoyer en Oklahoma, ils avaient fait tourner ce système et Yorick en parallèle. Les signaux émis par Hamlet étaient transmis simultanément aux récepteurs de Yorick et aux nombreuses entrées de l'ordinateur. Et les sorties de Yorick n'étaient pas seulement adressées à Hamlet, mon corps ; elles étaient également enregistrées et comparées aux réponses simultanées du programme, qui, pour des raisons qui ne m'étaient pas claires, s'appelaient « Hubert ». Les jours et les semaines passèrent, et les sorties obtenues étaient identiques et synchrones, ce qui ne prouvait évidemment pas qu'ils avaient réussi à copier la structure fonctionnelle du cerveau, mais constituait une base empirique fort encourageante.

Les entrées d'Hubert, et donc ses activités, étaient restées en parallèle avec Yorick pendant tout le temps où j'avais été désincarné. Et maintenant, pour le démontrer, ils avaient enclenché l'interrupteur principal qui, pour la première fois, donnait à Hubert la commande directe de mon corps, c'est-à-dire, évidemment, de Fortinbras. (J'appris qu'Hamlet était resté enseveli sous terre et était, depuis, très certainement redevenu poussière. En tête de ma tombe trônait toujours le somptueux amas de l'engin abandonné, orné sur le côté du mot MEC en grosses lettres, ce qui devrait donner aux archéologues du siècle prochain un curieux aperçu des rites d'inhumation de leurs ancêtres.)

Ensuite, les techniciens du laboratoire me montrèrent l'interrupteur principal, qui avait deux positions marquées C, pour Cerveau (ils ne savaient pas que mon cerveau s'appelait Yorick) et H, pour Hubert. Effectivement, il était sur H, et ils m'expliquèrent que si je le désirais, je pouvais le commuter sur C. Ce que je fis, avec un petit pincement de cœur. Et il ne se passa rien. Un déclic, c'est tout. Voulant contrôler leurs dires, l'interrupteur général étant positionné sur C, je basculai le bouton de l'émetteur de sortie de Yorick, sur la cuve, et cette fois-là, je faillis sans aucun doute me trouver mal. Après qu'il eut été inversé et que j'eus, en quelque sorte, repris mes esprits, je me remis à jouer avec l'interrupteur principal, dont je ne cessais de changer la position. Exception faite du déclic transitoire, je ne perçus pas la moindre différence. Je pouvais

l'enclencher au beau milieu d'une phrase, et la suite d'un énoncé commencé sous le contrôle de Yorick sortait sans aucun arrêt ni décrochage sous le contrôle d'Hubert. J'avais un cerveau de secours, une prothèse qui pourrait peut-être, un jour, m'être bien utile si Yorick avait un problème. Ou, vice versa, je pouvais garder Yorick comme cerveau de rechange et utiliser Hubert. Ce choix semblait tout à fait sans importance, car l'usure et la fatigue de mon corps n'avaient d'effet dégradant sur aucun des deux cerveaux, qu'il commandât les mouvements de mon corps ou émit des sorties dans le vide.

Le seul aspect vraiment incertain de cette nouvelle phase de mon évolution, c'était la possibilité, que je ne mis pas longtemps à entrevoir, que quelqu'un déconnecte le cerveau de secours — Hubert ou Yorick, selon le cas — de Fortinbras et le rebranche à un autre corps — quelque Rosencrantz ou Guildenstern fraîchement débarqué. Alors, si ce n'était pas déjà fait, il y aurait clairement deux personnes. L'une serait moi, et l'autre serait une sorte de super-jumeau. S'il y avait deux corps, l'un sous le contrôle d'Hubert et l'autre sous celui de Yorick, lequel serait reconnu comme le vrai Dennett ? Serait-ce le corps couplé à Yorick, en vertu de la priorité causale de ce dernier et de son ancienne liaison intime avec le corps originel de Dennett, Hamlet ? Ça paraissait un peu trop légaliste, un peu trop entaché de l'arbitraire de la consanguinité et de la possession juridique, pour constituer un argument convaincant au niveau métaphysique. Supposons, en effet, qu'avant l'entrée en scène du second corps, j'aie gardé Yorick comme cerveau de secours pendant des années, laissant à Hubert le soin permanent de piloter mon corps, c'est-à-dire Fortinbras. Le couple de squatters formé par Hubert et Fortinbras semblerait alors (pour combattre une intuition juridique par une autre) être le véritable Dennett et l'héritier légal de tout ce qui lui appartenait. C'était là une question intéressante, mais loin d'être aussi pressante qu'une autre qui me préoccupait : je sentais que dans une telle éventualité, je survivrais tant que l'un ou l'autre des couples cerveau/corps resterait intact, mais je ne savais pas bien, émotionnellement, si je devais souhaiter la survie des deux.

Je fis part de mes inquiétudes aux techniciens et au directeur du projet, leur expliquant que pour des raisons en grande partie sociales, l'idée de l'existence de deux Dennett me répugnait. Je ne voulais pas être mon propre rival dans le cœur de ma femme, ni partager mon modeste salaire de professeur avec un autre Dennett. Ce qui était encore plus fou et plus dégoûtant, c'était l'idée d'en savoir tant sur quelqu'un d'autre pour qui je n'aurais pas plus de secrets. Comment pourrions-nous nous regarder en face ? Mes collègues du laboratoire me firent remarquer que j'ignorais totalement le bon côté de la situation. N'y avait-il pas des tas de choses que je voulais faire mais qu'étant seul, je n'avais pas pu réaliser ? Un Dennett pourrait rester là et être le professeur, le mari et le père, tandis que l'autre pourrait partir à l'aventure, la gorge serrée à la pensée de sa famille, bien sûr, mais heureux de savoir que l'autre Dennett veillerait sur la maisonnée. Je pourrais être fidèle et adultère à la fois. Je pourrais même me tromper moi-même, sans parler d'autres

possibilités encore plus horribles que mes collègues essayèrent promptement d'imposer à mon imagination surmenée. Mais ma mission en Oklahoma (ou était-ce à Houston?) m'avait rendu moins téméraire, et je reculai devant cette occasion qui m'était offerte (mais qui, en fait, ne m'était peut-être pas offerte à moi, qui sait !).

Il existait une autre possibilité encore plus désagréable : le cerveau de secours, Hubert ou Yorick, pourrait être déconnecté des entrées de Fortinbras et rester ainsi, débranché. Dans ce cas, comme dans le précédent, il y aurait alors deux Dennett, ou tout au moins deux prétendants à mon nom et à mes biens, l'un rattaché à l'enveloppe charnelle de Fortinbras, et l'autre tristement dépourvu de corps. Mais l'égoïsme tout autant que l'altruisme me poussèrent à intervenir pour empêcher un tel événement. Je demandai donc que l'on prenne des mesures afin que personne ne puisse toucher aux connexions des émetteurs-récepteurs ou à l'interrupteur principal sans que je (nous? non, je) sois au courant et d'accord. Comme je n'avais aucune envie de passer ma vie à surveiller le matériel à Houston, il fut mutuellement décidé de protéger tous les branchements électroniques du laboratoire en les enfermant dans des boîtiers soigneusement verrouillés. Ceux qui assuraient le fonctionnement du système de survie de Yorick et ceux qui alimentaient Hubert en électricité seraient protégés par des dispositifs à sécurité intégrée, et j'emporterais l'interrupteur principal, équipé d'un système de commande à distance par radio, partout où j'irais. Je le porte attaché à ma ceinture et — un instant — le voici. Tous les deux ou trois mois, je vérifie l'état des choses en changeant de canal. Je le fais évidemment toujours en présence de quelques amis, car si jamais le nouveau canal était, Dieu m'en garde, défectueux ou occupé de quelque façon, il y aurait là une personne soucieuse de mes intérêts qui rebasculerait l'interrupteur et me ramènerait du néant. Car après une telle inversion, je serais capable de sentir, voir, entendre et percevoir par tous les autres moyens ce qui arriverait à mon corps, mais je serais dans l'impossibilité d'agir. D'autre part, les deux positions de l'interrupteur ne sont, volontairement, pas étiquetées, ce qui fait que je ne sais absolument pas si je passe d'Hubert à Yorick ou l'inverse. (Certains d'entre vous penseront peut-être que je ne sais donc vraiment pas qui je suis, et encore moins où je suis. Mais ce genre de réflexion n'affecte plus tellement l'essence de ma Dennettitude, ma perception de qui je suis. S'il est vrai que d'une certaine façon, je ne sais pas qui je suis, c'est alors là une autre de vos résistibles vérités philosophiques.)

En tout cas, jusqu'ici, quand j'avais actionné l'interrupteur, il ne s'était jamais rien passé. Alors, essayons une nouvelle fois...

« C'EST PAS TROP TÔT ! J'AI CRU QUE TU N'APPUIERAIS JAMAIS SUR CE BOUTON ! Tu ne peux pas savoir à quel point ces deux dernières semaines ont été horribles — enfin, maintenant, tu le sais; et c'est à ton tour de subir ce purgatoire. J'ai vraiment désiré ce moment de toutes mes forces ! Il y a à peu près deux semaines — excusez-moi, Messieurs Dames, mais il faut que je l'explique à mon.. disons, frère, mais vous êtes déjà au courant, vous comprendrez — il y a à

peu près deux semaines, nos deux cerveaux se sont légèrement désynchronisés. Je ne sais pas plus que toi si mon cerveau actuel est Hubert ou Yorick, mais toujours est-il qu'ils se sont décalés. Et ça a évidemment été en empirant, car quand nous recevions tous les deux une entrée, j'étais dans un état réceptif légèrement différent du tien, et cette différence n'a fait que croître et embellir. En un rien de temps, l'illusion que j'avais de maîtriser mon corps — c'est-à-dire notre corps — s'était totalement envolée. Et je ne pouvais rien faire, je n'avais aucun moyen de t'appeler. TU NE SAVAIS MEME RIEN DE MON EXISTENCE ! C'est comme être promené dans une cage, ou plutôt comme être possédé : j'entendais ma voix dire des choses que je ne voulais pas dire, je regardais, frustré, mes mains faire des gestes que je ne leur avais pas demandé de faire. Tu te grattais, quand ça te démangeait, mais pas comme je l'aurais fait, et tu m'empêchais de dormir, à force de te tourner et de te retourner. J'étais complètement épuisé, au bord de la dépression nerveuse, désespérément entraîné par la frénésie de tes activités, soutenu uniquement par la pensée qu'un jour, tu basculerais l'interrupteur.

« Et maintenant, c'est ton tour, mais tu pourras au moins te dire que je sais, moi, que tu es là. Comme une femme enceinte, je vais manger — ou au moins sentir, voir — pour deux, et j'essaierai de te faciliter les choses. Ne t'inquiète pas. Dès que ce colloque sera terminé, nous irons tous les deux à Houston, et nous verrons ce qu'on peut faire pour donner un autre corps à l'un d'entre nous. Tu pourrais avoir un corps de femme, ou un corps d'une autre couleur. Il faut y réfléchir. En tout cas, si nous voulons tous les deux garder ce corps, je jure que je laisserai le directeur du projet tirer à pile ou face lequel de nous le conservera et lequel aura à en choisir un nouveau. Ça serait équitable, non? Je te promets, de toute façon, de ne pas t'oublier. Et je fais ce serment devant témoins.

« Mesdames et Messieurs, la conférence que nous venons d'entendre n'est pas exactement celle que j'aurais donnée, moi, mais je peux vous assurer que tout ce qui a été dit était rigoureusement exact. Et maintenant, vous m'excuserez, mais je pense qu'un bon fauteuil me — nous — ferait extrêmement plaisir. »

Réflexions

L'histoire que vous venez de lire n'est pas véridique (au cas où vous vous seriez posé des questions), et elle ne pourrait pas l'être. Les prouesses techniques qui sont décrites sont impossibles à l'heure actuelle, et certaines d'entre elles resteront peut-être éternellement hors de notre portée, mais ce n'est pas ce qui nous importe. L'important, c'est de savoir s'il y a quelque chose de théoriquement impossible, quelque chose

d'incohérent, dans toute cette histoire. Lorsque les contes philosophiques deviennent trop extravagants et parlent, par exemple, de machines à voyager dans le temps, d'univers doubles ou de démons trompeurs et infiniment puissants, nous pouvons à bon droit refuser d'en conclure *quoi que ce soit*. Notre conviction d'avoir compris les problèmes posés peut, en effet, manquer de fiabilité et n'être qu'une illusion due au style très vivant du récit.

Dans ce cas-ci, l'opération chirurgicale et les micro-radios décrites ne sont pas imaginables avant un futur lointain ni même dans un quelconque avenir technologique clairement envisagé, mais il s'agit à coup sûr de science-fiction « innocente ». Il est moins certain que l'introduction d'Hubert, le double informatique de Yorick, le cerveau de Dennett, soit encore dans des limites raisonnables. (En tant qu'auteurs de fiction, nous avons le droit d'établir des règles au fur et à mesure que nous avançons, mais nous encourons le risque de raconter une histoire sans intérêt théorique.) Hubert est censé fonctionner *en parfait synchronisme* avec Yorick pendant des années sans l'intervention d'aucune liaison interactive correctrice. Il ne s'agirait pas là seulement d'un grand triomphe technologique, mais quasiment d'un miracle. Le problème n'est pas seulement que pour qu'un ordinateur se rapproche d'un cerveau humain, qui traite des millions de canaux d'entrée et de sortie parallèles, il faudrait qu'il ait une structure fondamentale totalement différente de celle des ordinateurs actuels. Même si un tel ordinateur existait, sa taille et sa complexité rendraient la perspective d'un tel comportement synchrone *indépendant* virtuellement impossible. Or, à défaut du traitement synchronisé et identique des deux systèmes, un élément essentiel de l'histoire devrait être abandonné. Pourquoi? Parce que la condition préalable de l'existence d'une seule personne ayant deux cerveaux (dont un de secours) en dépend. Voyez ce que dit Ronald de Sousa, dans « Rational Homunculi », au sujet d'un cas semblable :

Quand le Dr. Jekyll se transforme en M. Hyde, il se passe quelque chose d'étrange et de mystérieux. Sont-ils deux personnes occupant un même corps en alternance? Mais voici quelque chose de plus étrange encore : le Dr. Prodige et le Dr. Miracle occupent, eux aussi, un corps à tour de rôle. La différence, c'est qu'ils sont aussi semblables que de vrais jumeaux ! Vous objectez : pourquoi dire, alors, qu'ils deviennent l'un l'autre? Eh bien, pourquoi pas : si le Dr. Jekyll peut devenir un homme aussi différent de lui que Hyde, il doit être d'autant plus facile pour Prodige de se changer en Miracle, qui lui ressemble trait pour trait.

Il nous faut un conflit ou une différence marquée pour que nous puissions écarter notre idée naturelle qu'à un corps correspond tout au plus un agent.

Etant donné que plusieurs des éléments les plus remarquables de « Où suis-je? » reposent sur la supposition du traitement synchrone indépendant dans Yorick et Hubert, il est important de noter que cette supposition est véritablement exorbitante — elle serait à mettre dans le

même sac que l'hypothèse de l'existence, quelque part, d'une autre planète semblable à la Terre, avec un double, atome par atome, de vous, de tous vos amis et de tout ce qui vous entoure,* ou la supposition que l'univers ne date que d'il y a cinq jours (en fait, il ne paraît beaucoup plus ancien que parce que lorsque Dieu l'a créé, il y a cinq jours, Il a également créé un grand nombre d'adultes pleins de « souvenirs », des bibliothèques remplies de livres apparemment vieux, des montagnes regorgeant de fossiles flambant neufs, etc).

La possibilité de production d'une prothèse cérébrale comme Hubert n'est donc que théorique, quoique des éléments de système nerveux moins merveilleux soient peut-être bien pour demain. Il existe déjà, pour les aveugles, des ébauches d'yeux artificiels par l'intermédiaire de caméras de télévision; certains de ces dispositifs envoient les informations directement à des parties du cortex visuel du cerveau, mais d'autres évitent ces interventions chirurgicales d'une haute technicité et les transmettent par l'intermédiaire d'autres organes sensoriels externes, comme les récepteurs tactiles du bout des doigts ou encore un réseau de points sensibles sur le front, l'abdomen ou le dos du patient.

Les perspectives d'extension non chirurgicale de l'esprit sont précisément explorées dans le prochain texte, signé David Sanford, professeur de philosophie à la Duke University, qui constitue une suite de « Où suis-je? »

D. D.

* Comme dans la fameuse expérience mentale de la « Terre jumelle », d'Hilary Putnam. Cf. « Autres lectures »

sur les points de départ de nos perceptions et de nos motivations, mais dans son état actuel, elle a peu d'incidence sur le processus de pensée par lequel ces points de départ ont été transcendés.

Il peut y avoir des obstacles biologiques à l'achèvement de certaines sortes de progrès moral. Il ne fait aucun doute qu'il y a des obstacles psychologiques et sociaux, et que certains d'entre eux peuvent avoir des causes biologiques. Cela ne les rend pas insurmontables. Il faut les reconnaître et les traiter à l'aide d'une théorie morale qui ne soit pas utopique. Mais cette reconnaissance ne revient pas à l'acceptation d'un fondement biologique de l'éthique. On admet seulement par là que la moralité, comme tout autre processus de développement culturel, doit compter avec ses points de départ et avec la nature des matériaux qu'elle s'efforce de transformer.

LA BISECTION DU CERVEAU ET L'UNITÉ DE LA CONSCIENCE

I

ON a manifesté récemment beaucoup d'optimisme, parmi les philosophes et les chercheurs en neurosciences, quant aux perspectives de faire faire des découvertes majeures sur la base neurophysiologique de l'esprit. Les raisons de cet optimisme ont été extrêmement abstraites et générales. Je voudrais apporter quelques raisons d'être pessimiste. Ce type de compréhension de nous-mêmes peut rencontrer des limites qui n'ont pas été en général prévues : la conception mentaliste que l'on a de la nature des êtres humains peut constituer une résistance à une tentative pour la coordonner avec un mode de compréhension des humains comme systèmes physiques, et qui serait nécessaire à l'établissement de quoi que ce soit qui puisse être décrit comme une compréhension de la base physique de l'esprit. Je ne considérerai pas

les possibilités qui s'ouvriraient à nous si nous devions rencontrer des limites de ce genre. J'essaierai de présenter des raisons de croire à l'existence de ces limites – des raisons tirées des données nombreuses dont nous pouvons à présent disposer portant sur l'interaction entre les deux moitiés du cortex cérébral, et sur ce qui se produit quand elles cessent d'être en corrélation l'une avec l'autre. Le trait propre à la conception mentaliste qui peut ne pas se laisser intégrer avec ces données n'est pas un trait trivial ou accessoire, que l'on puisse aisément laisser de côté. C'est la notion d'une personne *unique*, d'un sujet unique d'expérience et d'action, qui présente des difficultés. Les difficultés peuvent être surmontées de certaines façons que je pourrais ne pas avoir prévues. Mais d'un autre côté cela peut être la première d'un grand nombre d'impasses qui se présenteront au fur et à mesure que nous serons en quête d'un mode de compréhension physiologique de l'esprit.

Partir à la recherche de la base ou de la réalisation physique des propriétés du monde phénoménal constitue dans de nombreux domaines un début d'enquête très profitable, et c'est la direction que privilégient, dans le cas des phénomènes mentaux, ceux qui recherchent une forme quelconque de réduction empirique de l'esprit au cerveau, par le biais d'une théorie de l'identité, d'une théorie fonctionnaliste, ou de quelque autre technique. Quand on tente d'appliquer le réductionnisme physique à un trait phénoménal du monde extérieur, les résultats sont quelquefois très bons, et peuvent être poussés jusqu'à des niveaux de plus en plus profonds. Si d'un autre côté, ils ne sont pas entièrement bons, et si certains traits de l'image phénoménale demeurent inexplicables par la réduction physique, alors nous pouvons écarter ces traits comme étant *purement* phénoménaux, et reporter notre tentative pour les comprendre au moment où notre connaissance de la base physique de l'esprit et de la perception aura avancé suffisamment pour la permettre. (Un exemple de ceci pourrait être l'illusion de la lune, ou d'autres illusions sensorielles qui n'ont pas de base susceptible d'être découverte dans les objets perçus.)

Cependant, si nous rencontrons le même type de difficulté en explorant la base physique des phénomènes de l'esprit lui-même, nous ne pouvons adopter la même ligne de retraite. A savoir : si un trait de l'esprit échappe à l'emprise de la théorie physique, nous ne pouvons pas reporter notre tentative de la comprendre au moment où nous étudierons l'esprit lui-même – parce que c'est exactement ce que nous sommes supposés être en train de faire. Déferer à un mode de com-

préhension de la base de l'esprit qui échappe à l'étude de la réalisation physique de certains de ses aspects, c'est admettre l'irréductibilité du mental au physique. Une manière d'admettre cela de façon tranchée serait de souscrire à une forme quelconque de dualisme. Mais si l'on répugne à s'engager sur une telle voie, alors on ne voit pas bien ce que l'on devrait faire des traits de la notion mentaliste de personnes qui résistent à l'assimilation à un mode de compréhension des êtres humains comme systèmes physiques. Il se peut que nous ne puissions jamais trouver une base objective pour certains de ces traits, ni que nous ne puissions non plus les abandonner. Il peut nous être impossible d'abandonner certaines manières dont nous nous concevons et nous représentons nous-mêmes, nonobstant le peu de justifications qu'elles reçoivent de la part de la recherche scientifique. C'est vrai, je suppose, de la notion d'unité de la personne : c'est une notion dont la validité peut être mise en question grâce à certaines découvertes récentes sur la dualité fonctionnelle du cortex cérébral. Il peut être utile de présenter ici un aperçu de ces résultats.

II

Les liaisons supérieures entre les deux hémisphères cérébraux ont été retirées chez des hommes, des singes et des chats, et les résultats ont conduit certains chercheurs à parler de la création de deux centres séparés de conscience dans un seul corps. Voici les faits¹. En gros l'hémisphère cérébral gauche est associé au côté droit du corps, et celui de droite au côté gauche. Des stimuli tactiles provenant d'un côté sont transmis à l'hémisphère opposé – à l'exception de la tête

1. La littérature sur les cerveaux divisés a des dimensions raisonnables. Un excellent aperçu récent est Michael Gazzaniga, *The bisected Brain*, New York, Appleton Century Crofts, 1970. Sa liste de références de neuf pages ne prétend pas constituer une bibliographie complète sur le sujet, cependant. Gazzaniga a écrit un court exposé de vulgarisation : *The split brain in man*, *Scientific American*, ccxvii, 1967, p. 24-29. Le meilleur exposé général d'un point de vue philosophique se trouve dans plusieurs articles de R. W. Sperry, le principal chercheur dans ce domaine : *The Great Cerebral Commisure*, *Scientific American*, ccx, 1964, p. 42 ; *Brain Bisection and the Mechanisms of Consciousness*, dans *Brain and Conscious Experience*, ed. J. C. Eccles, Berlin Springer Verlag, 1966 ; *Mental Unity following Surgical Disconnections of the Cerebral Hemispheres*, *The Harvey Lectures*, Series lxii, New York, Academic Press, 1968, p. 293-323 ; *Hemisphere Deconnection and Unity in Conscious Awareness*, *American Psychologist*, xxii, 1968, p. 723-733. On peut trouver plusieurs articles intéressants dans *Functions of the Corpus Callosum : Ciba Foundation Study Group*, n° 20, ed. G. Ettliger, London, J. & A. Churchill, 1965.

et du cou, qui sont reliés aux deux côtés à la fois. De plus la moitié gauche de chaque rétine, *i.e.* celle qui parcourt le côté droit du champ visuel, envoie des impulsions à l'hémisphère droit, et des impulsions provenant du côté gauche du champ visuel sont transmises par la moitié gauche de chaque rétine à l'hémisphère droit. Des impulsions auditives provenant de chaque oreille sont transmises jusqu'à un certain degré aux deux hémisphères à la fois. Les sensations olfactives, par contre, sont transmises ipsilatéralement : la narine gauche transmet à l'hémisphère gauche et la narine droite à l'hémisphère droit. Enfin l'hémisphère gauche contrôle en général la production de la parole.

Les deux hémisphères sont liés à la moelle épinière et aux nerfs périphériques à travers un point de jonction cérébral commun, mais ils communiquent aussi directement l'un avec l'autre, par une grande bande transversale de fibres nerveuses appelée le corps calleux, à quoi

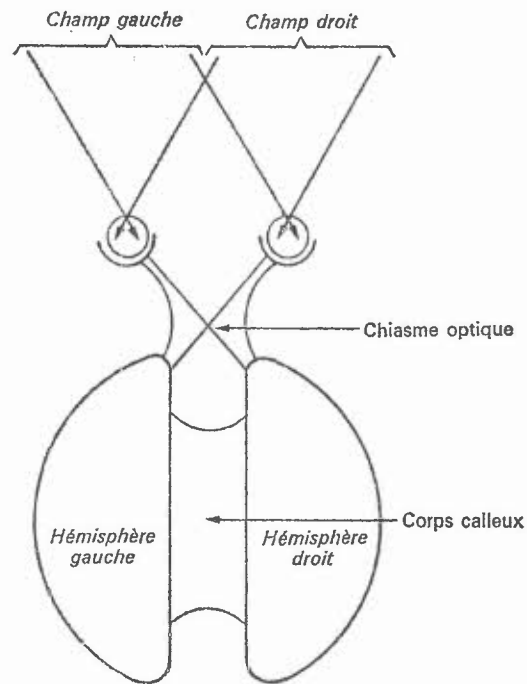


FIG. 1. — Une représentation très schématique (vue du dessus) des yeux et du cortex cérébral

s'ajoutent quelques voies de communication plus petites. Ces commissures cérébrales directes jouent un rôle essentiel dans la coordination ordinaire des fonctions entre les deux hémisphères chez les personnes normales. L'une des caractéristiques frappantes de la question est que ce fait est demeuré inconnu, tout au moins dans la communauté de langue anglaise, jusqu'à la fin des années cinquante, bien qu'on ait retiré à un bon nombre de patients leurs commissures cérébrales au cours d'opérations destinées à traiter l'épilepsie pendant la décennie antérieure. On n'avait pu observer aucun effet significatif d'ordre comportemental ou mental chez les patients, et on avait conjecturé que le corps calleux n'avait aucune fonction particulière, sauf peut-être celle d'empêcher les hémisphères de s'affaïsser.

Alors R. Myers et R. W. Sperry introduisirent une technique pour traiter les deux hémisphères séparément². Ils sectionnèrent le nerf optique de chats, en sorte que chaque œil envoie une information directe (une information concernant la moitié opposée du champ visuel) seulement à l'un des côtés du cerveau. Il fut alors possible d'entraîner les chats à des tâches simples requérant l'usage d'un seul œil, et de voir ce qui arrivait quand l'un d'entre eux utilisait l'autre œil à la place. Chez les chats dont le corps calleux était intact, le transfert d'apprentissage visuel était très bon. Mais chez certains chats, ils enlevèrent le corps calleux aussi bien que le chiasme optique ; et dans ces cas, rien n'était transmis d'un côté à l'autre. En fait aux deux côtés séparés on pouvait apprendre simultanément des discriminations en conflit les unes avec les autres, en fournissant aux deux yeux des stimuli opposés durant une série unique de renforcements. Cependant cette capacité à exercer une fonction indépendante ne produisit pas de défauts sérieux du comportement. A moins que les circuits d'accès aux deux hémisphères ne soient artificiellement isolés, l'animal semblait normal (bien que si un singe au cerveau sectionné attrape une cacahuète avec les deux mains, le résultat ressemble quelquefois à une lutte à la corde de traction).

Plutôt que de résumer toutes ces données, je me concentrerai sur les cas humains, que l'on fut amené à reconsidérer à la suite des

2. R. E. Myers et R. W. Sperry, Interoocular Transfer of a Visual Form Discrimination Habit in Cats after Section of the Optic Chiasma and Corpus Callosum, *Anatomical Record*, cxv, 1953, p. 351-352; R. E. Myers, Interoocular Transfer of Pattern Discrimination in Cats Following Section of Crossed Optic Fibers, *Journal of Comparative and Physiological Psychology*, xlviii, 1955, p. 470-473.

découvertes faites sur les chats et les singes³. Dans l'opération de bisection du cerveau pour le traitement de l'épilepsie, le chiasme optique est laissé intact, en sorte que l'on ne puisse accéder aux deux hémisphères séparément qu'à travers les deux yeux. La solution du problème du contrôle de l'entrée de l'information visuelle consiste à projeter des signaux sur un écran, sur l'un ou l'autre des côtés du point médiant du champ de vision du patient, assez longtemps pour qu'ils soient perçus, mais pas assez longtemps pour permettre un mouvement de l'œil qui conduirait le signal à la moitié opposée du champ visuel et par conséquent au côté opposé du cerveau. C'est ce que l'on appelle stimulation tachistoscopique. Les entrées d'information tactile par les mains sont pour la plupart isolées de façon très efficace, de même que les sensations olfactives à travers les deux narines. On a récemment isolé avec succès les sources d'information auditive, car chaque oreille semble envoyer des signaux plus puissants à l'hémisphère contralatéral qu'à l'hémisphère ipsilatéral. Pour ce qui est de l'information à la sortie, la distinction la plus claire est fournie par la parole, qui est exclusivement le produit de l'hémisphère droit⁴. Le cas de l'écriture est moins clair : elle peut être occasionnellement produite sous une forme rudimentaire par l'hémisphère droit, en utilisant la main gauche. En général le contrôle moteur est contralatéral, *i.e.* dépend de l'hémisphère opposé, mais une certaine quantité de contrôle ipsilatéral survient parfois, particulièrement de la part de l'hémisphère gauche.

Il en résulte les choses suivantes. Les signaux adressés au côté

3. La première publication de ces résultats fut M. S. Gazzaniga, J. E. Bogen, et R. W. Sperry, Some Functional Effects of Sectioning the Cerebral Commissures in Man, *Proceedings of the National Academy of Sciences*, XLVIII, 1962, Pt 2, p. 1765-1769. Il est intéressant de constater que la même année vit la publication d'un article proposant l'interprétation d'une lésion du cerveau humain dans des termes semblables, suggérée par les découvertes antérieures qui avaient été faites chez les animaux. Cf. N. Geschwind et E. Kaplan, A Human Cerebral Disconnection Syndrome, *Neurology*, XII, 1962, p. 675. Intéressant aussi est le long exposé en deux parties de N. Geschwind qui passe le domaine en revue, et pose quelques questions philosophiques explicitement : Disconnections Syndromes in Animals and Man, *Brain*, LXXXVIII, 1965, p. 247-294. Certaines parties sont réimprimées, avec des éléments nouveaux, dans *Boston Studies in the Philosophy of Science*, vol. IV, 1969. Voir aussi son article The Organization of Language and the Brain, *Science*, CLXX, 1970, p. 940.

4. Il y a à cela des exceptions individuelles, comme il y en a à la plupart des généralisations concernant les fonctions cérébrales : les gauchers tendent à avoir un contrôle linguistique bilatéral, et c'est chose courante dans la prime enfance. Tous les sujets de ces expériences cependant étaient droitiers et révélaient une dominance de l'hémisphère cérébral droit.

droit du champ visuel ou ressentis sans être vus par la main droite ne peuvent être rapportés verbalement. Les signaux adressés à la moitié gauche du champ visuel ou ressentis par la main gauche ne peuvent être rapportés, bien que si le mot « chapeau » est reçu comme signal sur la gauche, la main gauche retire un chapeau d'un groupe d'objets dissimulés si la personne reçoit l'ordre de se saisir de ce qu'elle a vu. En même temps elle insistera verbalement sur le fait qu'elle n'a rien vu. Ou encore, si deux mots différents sont adressés comme signaux aux deux moitiés du champ visuel (par exemple « crayon » et « brosse à dents ») et si on donne l'ordre à l'individu de retirer l'objet correspondant de derrière un écran, avec les deux mains, alors les deux mains iront à la recherche de la collection d'objets indépendamment l'une de l'autre, la main droite prenant le crayon et le rejetant tandis que la main gauche part à sa recherche, la main gauche rejetant de façon similaire la brosse à dents que la main droite sélectionne avec satisfaction.

Si un objet dissimulé est placé dans la main gauche et que l'on demande à la personne de deviner de quoi il s'agit, les conjectures fausses produisent un froncement de sourcils ennuyé, puisque l'hémisphère droit, qui perçoit l'information tactile entend aussi les réponses. Si l'hémisphère dont dépend le langage faisait des conjectures correctes, le résultat serait un sourire du patient. Une odeur envoyée à la narine gauche (qui stimule l'hémisphère droit) produira un déni verbal de la part du sujet qu'il perçoive une quelconque odeur, mais si on lui demande de pointer avec la main gauche en direction d'un objet correspondant, il réussira à attraper, par exemple, une gousse d'ail, protestant tout du long qu'il ne sent absolument rien : comment peut-il alors pointer du doigt ce qu'il sent ? Si l'odeur est désagréable, comme celle des œufs pourris, ces dénégations seront accompagnées de froncement du nez, la bouche fera la lippe, et il émettra des exclamations gutturales de dégoût⁵.

Un exemple particulièrement poignant de conflit entre les deux hémisphères est le suivant. On place en dehors de la vue du patient une pipe dans sa main gauche, puis on lui demande d'écrire avec sa main gauche ce qu'il y tenait. Très laborieusement et lourdement la main gauche écrit les lettres P et I. Alors soudain l'écriture se fait

5. H. W. Gordon et R. W. Sperry, Lateralization of Olfactory Perception in the Surgically Separated Hemispheres in Man, *Neuropsychologia*, VII, 1969, p. 111-120. Un patient cependant fut capable dans ces circonstances de dire qu'il sentait quelque chose de désagréable, sans être capable de le décrire plus avant.

plus rapide et plus légère, le I devient un E, et le mot est complété en PEIGNE. De toute évidence l'hémisphère gauche a fait une conjecture fondée sur l'apparence des deux premières lettres, et a interféré, avec contrôle ipsilatéral. Mais alors l'hémisphère droit reprend le contrôle de la main à nouveau, raye lourdement les lettres EIGNE et trace le dessin grossier d'une pipe⁶.

Il y a bien d'autres données. Le patient au cerveau sectionné ne peut dire si des formes adressées aux deux moitiés du champ visuel ou tenus en dehors du regard dans chaque main sont identiques ou différentes, même si on lui demande d'indiquer sa réponse en inclinant ou en secouant la tête (réponses qui sont disponibles pour les deux hémisphères). Le sujet ne peut pas distinguer une ligne continue d'une ligne discontinue projetée à travers les deux moitiés du champ visuel si la cassure tombe au milieu. Pas plus ne peut-il dire si deux taches situées dans des moitiés opposées du champ ont la même couleur ou des couleurs différentes, bien qu'il soit capable de faire toutes ces choses si les images à comparer tombent à l'intérieur d'une seule moitié du champ visuel. Dans l'ensemble l'hémisphère droit est meilleur pour les tests reposant sur des relations spatiales, mais il est presque incapable de faire des calculs. Il paraît cependant susceptible d'émotion. Par exemple, si une photographie de femme nue est projetée sur la moitié gauche du champ visuel du patient, il émettra de forts grognements et peut-être rougira, sans être capable de dire ce qui lui a fait plaisir, bien qu'il puisse dire : « Ouh, vous avez là un drôle de truc. »

Tout ceci se combine avec ce qui prend une allure parfaitement normale dans des activités ordinaires, quand on n'a créé aucune isolation artificielle de l'information dans les deux hémisphères. Les deux côtés s'endorment et se réveillent au même moment. Les patients peuvent jouer du piano, boutonner leur chemise, nager, et accomplir convenablement d'autres activités requérant la coordination bilatérale. Qui plus est, ils ne font état d'aucune sensation de division ou de réduction de leur champ visuel. La déviation la plus notable par rapport au comportement ordinaire eut lieu chez un patient dont la main gauche semblait être quelque peu hostile à la femme dudit patient. Mais dans l'ensemble les deux hémisphères coopèrent admirablement, et on doit recourir à des techniques expérimentales subtiles

6. Rapporté dans Jerre Levy, *Information Processing and Higher Psychological Functions in the Disconnected Hemispheres of Human Commissurotomy Patients* (Dissertation doctorale non publiée, California Institute of Technology, 1969).

pour les amener à opérer séparément. Si l'on n'y prend pas garde, ils se donneront l'un à l'autre des consignes périphériques, transmettant de l'information par des signaux auditifs, visuels, ou perceptibles au moyen d'autres canaux sensoriels, qui compensent le manque de lien commissural. (Il y a une forme de communication qu'il est particulièrement difficile d'empêcher, tellement elle est directe : les deux hémisphères peuvent mouvoir les muscles du cou et de la face, et tous deux peuvent les sentir bouger ; en sorte qu'une réponse produite au niveau facial ou dans la tête par l'hémisphère droit peut être détectée par celui de gauche, et il semble confirmé qu'ils s'adressent des signaux l'un à l'autre à travers ce médium⁷.

III

Ce que l'on désire naturellement savoir concernant ces patients, c'est combien d'esprits ils peuvent bien avoir. Ceci soulève immédiatement des questions quant aux sens auxquels une personne ordinaire peut être dite avoir un seul esprit, et quelles sont les conditions sous lesquelles diverses expériences et activités peuvent être attribuées au même esprit. Il nous faut avoir une idée quelconque de quoi une personne ordinaire est une, pour pouvoir comprendre ce que nous voulons savoir quand nous demandons s'il y en a une ou deux, quand nous essayons de décrire ces patients extraordinaires.

Cependant, plutôt que de commencer par une analyse de l'unité de l'esprit, la manière dont je procéderai sera de tenter d'appliquer la conception ordinaire, non analysée, directement à l'interprétation de ces données, en posant la question de savoir si ces patients ont un esprit, ou deux, ou s'ils relèvent d'une configuration plus exotique. Ma conclusion sera que la conception ordinaire selon laquelle nous avons un esprit unique, dénombrable, ne peut pas du tout leur être

7. De plus la condition de disconnection radicale peut ne pas être stable : il peut y avoir une tendance à la formation de nouveaux chemins de communication interhémisphériques par l'intermédiaire du radical cervical, si un certain temps s'écoule. Ceci est confirmé en partie par l'observation de patients ayant subi une commissurotomie, mais mieux encore par des cas d'agénésie du corps calleux. Des gens qui ont grandi sans celui-ci ont appris à s'en passer : leur réussite aux tests est beaucoup plus proche de la normale que celle de patients récemment opérés. (Cf. L. J. Saul et R. W. Sperry, *Absence of commissurotomy symptoms and agenesis of the Corpus Callosum*, *Neurology*, xviii, 1968. Ce fait est très important, mais pour le moment je le laisserai de côté pour me concentrer sur les résultats immédiats de la disconnection.

appliquée, et qu'ils ne possèdent aucun nombre particulier d'esprits, bien qu'ils se livrent sans aucun doute à des activités mentales. Une compréhension plus claire de la notion d'esprit individuel devrait ressortir au cours de la discussion, mais les difficultés qui font obstacle à son application aux cas de bisection du cerveau fourniront des raisons plus générales d'émettre des doutes. Il se peut que le concept ne soit pas non plus applicable à des êtres humains ordinaires car il incorpore une conception trop simple de la manière dont les êtres humains fonctionnent.

Néanmoins je ferai usage de la notion d'un esprit individuel en discutant les cas au début, car je veux considérer de façon systématique comment ils pourraient être compris en termes d'esprits individuels, et soutenir qu'ils ne peuvent pas l'être. Cela accompli, je me tournerai vers les personnes ordinaires, comme vous et moi.

Il semble y avoir cinq interprétations des données expérimentales utilisant le concept d'esprit individuel.

1) Les patients ont un esprit tout à fait normal associé à l'hémisphère gauche, et les réponses émanant de l'hémisphère droit sans langage sont les réponses d'un automate, et ne sont pas produites par des processus mentaux conscients.

2) Les patients ont seulement un esprit, associé à l'hémisphère gauche, mais il survient également (en association avec l'hémisphère droit) des phénomènes conscients isolés, qui ne sont pas intégrés dans un esprit, bien qu'ils puissent être attribués à l'organisme.

3) Les patients ont deux esprits, dont l'un peut parler et l'autre pas.

4) Ils ont un esprit, dont les contenus dérivent de chacun des hémisphères et sont plutôt particuliers et dissociés.

5) Ils ont un esprit normal la plupart du temps, tandis que les hémisphères fonctionnent de manière parallèle, mais deux esprits surgissent dans des situations expérimentales productrices de résultats intéressants. (Peut-être l'esprit individuel se sépare-t-il en deux et se reconstitue-t-il une fois l'expérience terminée.)

Je montrerai que chacune de ces interprétations est inacceptable pour une raison ou pour une autre.

IV

Je discuterai d'abord les hypothèses 1) et 2), qui ont en commun le refus d'attribuer les activités de l'hémisphère droit à l'esprit, puis je traiterai des hypothèses 3), 4) et 5), qui associent toutes un esprit

aux activités de l'hémisphère droit, bien qu'elles divergent quand il s'agit de savoir ce qu'est l'esprit.

La seule justification de l'hypothèse 1), qui refuse d'attribuer la conscience aux activités de l'hémisphère droit, est le fait que le sujet dénie, invariablement, être conscient des activités de cet hémisphère. Mais prendre cela pour une preuve que les activités de l'hémisphère droit sont inconscientes, c'est faire une pétition de principe, parce que la capacité à porter un témoignage est la seule aptitude dont dispose l'hémisphère gauche, et bien sûr l'hémisphère gauche n'est pas conscient de ce qui se produit du côté droit. Si d'autre part nous considérons les manifestations de l'hémisphère droit lui-même, il semble n'y avoir en principe aucune raison de considérer la verbalisation comme une condition nécessaire de la conscience. Il peut y avoir d'autres raisons d'attribuer des états mentaux conscients qui soient suffisants, même sans verbalisation. Et en fait, ce que l'hémisphère droit peut faire par lui-même est trop élaboré, trop intentionnellement dirigé, et trop intelligible psychologiquement pour être considéré simplement comme une collection de réponses automatiques inconscientes.

L'hémisphère droit n'est pas très intelligent et ne peut pas parler, mais il est capable de répondre à des stimuli visuels et auditifs complexes, y compris le langage, et il peut contrôler l'effectuation de tâches de discrimination et de manipulation qui requièrent une attention suivie — telles qu'épeler des mots simples avec des lettres en plastique. Il peut coordonner des stimuli auditifs, visuels et tactiles, afin de suivre les instructions de l'expérimentateur, et il peut effectuer certains tests d'aptitude. Il ne fait aucun doute que si une personne était privée de son hémisphère gauche entièrement, en sorte que les seules capacités qui lui restent soient celles du côté droit, nous ne dirions pas dans ces conditions qu'elle s'est transformée en un automate. Bien que privée du langage, elle demeurerait consciente et active, avec un champ visuel diminué et une paralysie partielle du côté droit dont elle pourrait guérir jusqu'à un certain point. Compte tenu de cela, il semblerait arbitraire de dénier que les activités de l'hémisphère droit sont conscientes, simplement parce qu'elles surviennent côte à côte avec celles de l'hémisphère gauche, dont le caractère conscient n'est pas en question.

Je n'ai pas l'intention de soutenir que la ligne de partage entre l'activité consciente et inconsciente est bien tranchée. Il est même possible que la distinction soit en partie relative, au sens où un élément

d'activité mentale peut être ou non attribué à la conscience, selon les autres activités mentales de la même personne qui se déroulent au même moment, et selon qu'elle leur est reliée de manière appropriée. Même si ceci est vrai, cependant, les activités de l'hémisphère droit chez les personnes au cerveau sectionné ne tombent pas dans la catégorie des événements dont l'inclusion dans le champ de la conscience dépend de ce qui se passe ailleurs dans l'esprit du patient. Leurs déterminants comprennent une gamme complète de facteurs psychologiques et ils requièrent un certain état d'éveil. Il est clair que l'attention, et même la concentration, est requise pour les tâches accomplies par la main gauche cachée et pour le champ visuel gauche stimulé par tachistoscopie. Les sujets ne passent pas leurs tests expérimentaux dans une sorte de rêve : selon toute évidence ils sont en contact avec la réalité. L'hémisphère gauche se plaint occasionnellement qu'on lui demande d'accomplir des tâches que l'hémisphère droit peut accomplir, parce qu'il ne sait pas ce qui se produit quand l'hémisphère droit contrôle la réponse. Mais l'hémisphère droit montre suffisamment de conscience de ce qu'il fait pour justifier qu'on lui attribue un contrôle conscient en l'absence de témoignage verbal. Si les patients ne déniaient pas avoir une conscience de ces activités, il ne devrait pas faire de doute qu'ils en sont conscients.

Les considérations qui rendent la première hypothèse intenable servent aussi à réfuter l'hypothèse 2), qui suggère que les activités de l'hémisphère droit sont conscientes sans appartenir à un esprit pour autant. La question peut se poser de savoir si cette proposition est intelligible, mais nous n'avons pas besoin de les considérer ici, parce qu'elle perd toute plausibilité face au haut degré d'organisation et de cohérence intermodale des activités de l'hémisphère droit. Elles ne flottent pas en l'air, et ne sont pas organisées de manière fragmentaire. L'hémisphère gauche suit des instructions, contrôle des stimuli tactiles, auditifs et visuels, et fait la plupart des choses qu'on est en droit d'attendre d'un esprit normal. Les données nous mettent en présence non seulement d'épisodes de comportement intentionnel, mais aussi d'un système capable d'apprendre, de réagir émotionnellement, de suivre des instructions, et d'accomplir des tâches qui requièrent l'intégration de divers déterminants psychologiques. Il semble clair que les activités de l'hémisphère droit ne sont pas inconscientes, et qu'elles font partie de quelque chose qui a une structure mentale spécifique : un sujet d'expérience et d'action.

V

Je me tournerai à présent vers les trois hypothèses selon lesquelles les activités conscientes de l'hémisphère droit sont attribuées à un esprit. On doit les considérer ensemble, parce que la difficulté fondamentale pour chacune d'elles tient à l'impossibilité d'établir une distinction entre elles. La question est donc celle de savoir si les patients ont deux esprits, un esprit, ou un esprit qui se sépare à l'occasion en deux.

De nombreuses raisons militent en faveur de la thèse selon laquelle ils ont deux esprits, *i.e.* que les activités de l'hémisphère droit appartiennent à un esprit spécifique⁸. Chaque côté du cerveau semble produire ses propres perceptions, croyances et actions, qui sont reliées les unes aux autres de manière usuelle, mais pas à celles du côté opposé. Les deux parties du cortex partagent un corps commun, qu'elles contrôlent à travers un cerveau moyen et une moelle épinière commune. Mais les fonctions supérieures sont indépendantes non seulement physiquement, mais aussi psychologiquement. Les fonctions de l'hémisphère droit non seulement n'ont pas accès au langage, mais aussi à toute combinaison directe avec les fonctions correspondantes de l'hémisphère gauche — *i.e.* avec des fonctions de l'espèce de celles que l'hémisphère droit réalise aisément dans la partie qui est la sienne, telle que la discrimination des couleurs ou des formes.

Il y a un témoignage fourni par les hémisphères gauches des patients qui semble aller contre la thèse des deux esprits. Ils rapportent n'éprouver aucune diminution du champ visuel, et peu d'absence de sensation du côté gauche. Sperry minimise cette donnée en arguant qu'elle est comparable au témoignage des personnes victimes de scotome (destruction partielle de la rétine), selon lequel elles ne font état d'aucun trou dans leur champ visuel — bien que ces trous puissent être décou-

8. C'est la théorie de R. W. Sperry. Il la formule ainsi : « A la place d'un courant unique et unifié et normal de la conscience, ces patients se comportent de manières multiples, comme s'ils avaient deux courants indépendants de conscience, l'un dans chaque hémisphère, dont chacun est coupé et hors de contact avec les expériences mentales de l'autre. En d'autres termes, chaque hémisphère semble avoir ses sensations propres et privées, ses propres concepts, et ses propres impulsions à agir, accompagnées d'expériences de volitions, de cognition et d'apprentissage associées. A la suite de l'opération chirurgicale, chaque hémisphère a aussi désormais sa propre chaîne séparée de souvenirs qui sont rendus inaccessibles au processus de reconnaissance des autres », *American Psychologist*, xxiii, p. 724.

verts par d'autres personnes observant leurs déficiences perceptives. Mais nous n'avons pas besoin d'admettre qu'un mécanisme de l'invention élaboré est à l'œuvre dans l'hémisphère gauche pour rendre compte de tels témoignages. Il est parfaitement possible que bien qu'il y ait deux esprits, l'esprit associé avec chaque hémisphère reçoive, à travers le radical cervical commun, une certaine quantité de stimulations ispilatérales grossières, en sorte que l'esprit doté du langage ait un appendice rudimentaire et indifférencié à gauche de son champ visuel, et *vice versa* pour l'hémisphère droit⁹.

Les difficultés réelles rencontrées par l'hypothèse des deux esprits coïncident avec les raisons que l'on a de penser que nous avons affaire à un seul esprit — en l'occurrence le caractère hautement coordonné des relations que les patients ont avec le monde dans les circonstances ordinaires. Quand ils ne se trouvent pas dans des situations expérimentales, leur étonnante dissociation comportementale disparaît, et ils fonctionnent normalement. Il y a peu de doute que l'information en provenance des deux parties de leur cerveau puisse être mise en commun de manière à produire un contrôle coordonné de leur comportement. Et bien que ceci ne s'accomplisse pas par les méthodes usuelles, il n'est pas évident que cela suffise à étayer un refus d'attribuer les fonctions de coordination à un seul esprit. Après tout, s'il est permis au patient de toucher des choses avec les deux mains et de sentir leur odeur avec les deux narines, il parvient à une notion unifiée de ce qui se passe autour de lui et de ce qu'il fait, sans manifester aucune incohérence de type gauche-droite dans ses comportements ou ses attitudes. Il semble étrange de suggérer que nous ne sommes pas en position d'attribuer toutes ces expériences à la même personne, simplement à cause de certaines particularités dans la manière dont la coordination s'effectue. Les gens qui *connaissent* ces patients trouvent naturel d'entrer en relation avec eux en tant qu'individus uniques.

Néanmoins, si nous attribuons la coordination à un seul esprit, nous devons aussi attribuer la dissociation suscitée expérimentalement à cet esprit, et ce n'est pas facile. La situation expérimentale révèle diverses dissociations ou conflits qui sont inhabituels non seulement en raison de la simplicité de sa base anatomique, mais aussi parce qu'une large *gamme* de fonctions est scindée en deux branches non

9. Certaines données confirment des entrées d'information primitives ipsilatérales de ce genre. Cf. M. Gazzaniga, *The Bisected Brain*, chap. 3.

communicantes. Ce n'est pas comme si deux centres de volition conflictuels partageaient un appareil de réception et de raisonnement commun. La scission est beaucoup plus profonde que cela. L'hypothèse d'un esprit unique doit par conséquent affirmer que les contenus de la conscience unique d'un individu sont produits par deux systèmes de contrôle situés dans les deux hémisphères, chacun étant doté d'une structure mentale bel et bien complète. Si ce double contrôle s'accomplissait durant des situations expérimentales, selon une alternance temporelle, il serait intelligible, quoique mystérieux. Mais ce n'est pas ce que pose l'hypothèse, et l'hypothèse telle qu'elle est ne nous permet pas de comprendre. Car chez ces patients il semble y avoir des choses qui se déroulent *simultanément*, qui ne peuvent pas cadrer avec l'idée d'un esprit unique : une attention simultanée à deux tâches incompatibles, par exemple, sans interaction entre les intentions des mains droite et gauche respectivement.

Cela rend difficile de concevoir quel effet cela fait (*what it is like*) d'être l'une de ces personnes. L'absence d'interaction au niveau d'un système de contrôle préconscient serait compréhensible. Mais une absence d'interaction dans le domaine de l'expérience visuelle et de l'intention consciente menace le postulat de l'unité de la conscience, qui est au fondement de notre compréhension d'un autre individu en tant que personne. Ces postulats sont associés à l'idée que nous avons de nous-mêmes, qui règle dans une mesure considérable notre compréhension des autres. Et ce sont précisément ces postulats, je crois, qui font qu'il est impossible de parvenir à une interprétation des cas discutés en termes d'une certaine quantité numérique d'esprits.

En gros, nous admettons qu'un esprit individuel a un accès immédiat suffisant à ses états conscients, en sorte que, pour des éléments d'expérience qui surviennent simultanément ou dans une contiguïté temporelle rapprochée, l'esprit qui en est le sujet peut aussi faire l'expérience des *relations* les plus simples qu'il y a entre eux s'il est attentif. Ainsi nous admettons que quand une personne individuelle a deux impressions visuelles, elle peut habituellement avoir l'expérience de l'identité ou de la différence entre leur coloration, leur forme et leur taille, de la relation entre leur position et leur mouvement dans son champ visuel, et ainsi de suite. On peut dire la même chose des relations entre diverses expériences. Les expériences d'une personne individuelle sont pensées comme prenant place dans un domaine lié intrinsèquement *dans l'expérience*, en sorte que les relations entre

des expériences peuvent être substantiellement saisies dans des expériences de ces relations¹⁰.

Des patients au cerveau sectionné manquent de manière dramatique à se conformer à ces postulats dans les situations expérimentales, et ils échouent dans les cas les plus simples. De plus la dissociation a lieu entre deux classes d'états conscients chacun étant caractérisé par une cohérence *interne* significative : les postulats normaux concernant l'unité de la conscience sont vérifiés dans chaque hémisphère séparément, tandis que les comparaisons nécessaires ne peuvent pas se faire quand les deux hémisphères ne sont plus en contact.

Ces considérations nous ramènent à l'hypothèse selon laquelle ces patients ont deux esprits chacun. Cela a au moins l'avantage de nous permettre de comprendre quel effet cela fait d'être ces individus, pour autant que nous n'essayions pas d'imaginer quel effet cela fait d'être les deux en même temps. Et cependant la voie qui nous conduit à accepter confortablement cette conclusion est bloquée par la coordination comportementale très convaincante révélée par ces patients dans la vie ordinaire, en comparaison de laquelle les symptômes dissociés suscités par la situation expérimentale semblent périphériques et atypiques. Nous sommes confrontés à des corps de données diamétralement opposés, dans un cas où une décision arbitraire est impossible. Nous sommes fortement enclins à penser qu'il y a un nombre entier *quelconque* d'esprits dans ces têtes, mais les données nous empêchent de dire combien.

Ce dilemme rend l'hypothèse 5) attrayante au premier abord, en particulier parce que les données qui sont à l'origine du conflit sont jusqu'à un certain point relevées à des moments différents. Mais la suggestion selon laquelle un second esprit est porté à l'existence seulement pendant des situations expérimentales perd son caractère plausible, quand on y réfléchit. En premier lieu, elle est entièrement *ad hoc* : elle propose d'expliquer un changement en termes d'un autre

10. Les deux peuvent bien sûr diverger, et c'est ce fait qui sous-tend le problème philosophique classique du spectre inversé, qui a une relation seulement lointaine avec le thème de cet article. Il y a un type de relation qui peut se produire entre les éléments de l'expérience d'une personne qui ne peut pas se produire entre les éléments de l'expérience de deux personnes distinctes : sembler avoir une couleur similaire par exemple. Pour autant que notre concept de similitude dans l'expérience, dans le cas d'une personne individuelle, dépende de son expérience de la similitude, le concept n'est pas applicable quand deux personnes sont en jeu au moins.

sans suggérer une explication du second. Il n'y a rien dans la situation expérimentale dont on puisse espérer qu'il produise un changement interne fondamental chez le patient. En fait elle ne produit aucun changement anatomique et dégage seulement un certain nombre de symptômes remarquables. Un événement aussi inhabituel que l'éclosion puis la disparition d'un esprit mériterait d'être expliqué mieux que par cette explication pour les besoins de la cause.

Mais en second lieu, les données comportementales ne seraient même pas expliquées par cette hypothèse, simplement parce que les réponses coordonnées des patients et leurs réponses dissociées ne sont pas clairement distinctes dans le temps. Pendant la durée des expériences, le patient se comporte dans une large mesure comme s'il était un individu unique : dans sa posture, en suivant des instructions sur l'endroit où il doit faire porter son regard, toute la gamme de contrôle comportemental ordinaire à l'œuvre dans le fait de se situer lui-même en relation avec l'expérimentateur et le dispositif expérimental. Les deux moitiés de son cerveau coopèrent complètement au regard de ces informations spécifiques qui les atteignent séparément de manière différenciée. Pour ces raisons, l'hypothèse 5) ne semble pas constituer une option véritable : si deux esprits sont à l'œuvre dans la situation expérimentale, ils doivent être à l'œuvre dans une large mesure en harmonie, bien qu'ils soient en partie dissonants. Et s'il y a alors deux esprits, pourquoi ne peut-il pas y avoir deux esprits à l'œuvre parallèlement le reste du temps ?

Néanmoins la coordination psychologique que révèlent les patients dans la vie ordinaire est si complète que je ne crois pas qu'il soit possible d'accepter cette conclusion, pas plus qu'une conclusion impliquant qu'on leur attribue un nombre entier d'esprits. Ces cas se situent à mi-chemin entre ceux de personnes ordinaires au cerveau intact (entre les hémisphères cérébraux desquels il y a aussi coopération, bien qu'elle s'effectue largement par l'intermédiaire du corps calleux), et des paires d'individus engagés dans des activités qui requièrent une coordination précise de leur comportement, par exemple quand on utilise une scie à deux manches, ou quand on joue un duo. Dans ce dernier type de cas, nous avons deux esprits qui communiquent par des directives périphériques subtiles ; dans le premier il n'y a qu'un seul esprit. Rien dans l'un ou l'autre cas ne peut nous conduire à assimiler le cas du patient dont le cerveau est scindé en deux à l'un ou l'autre d'entre eux. Si nous décidons qu'ils ont vraiment deux esprits, alors on ne verrait pas pourquoi on ne conclurait pas, pour

des raisons anatomiques, que chacun a deux esprits, mais que nous n'y faisons pas attention, sauf en présence de ces cas bizarres, parce que la plupart des paires d'esprits logés dans un corps unique fonctionnent de manière parfaitement parallèle à cause de la communication directe qu'il y a entre les hémisphères qui en constituent les bases anatomiques. Les deux esprits qui existent attelés dans chacun de nous seraient bien identiques, à cette différence près que l'un peut parler tandis que l'autre ne le peut pas. Mais il est clair que ce type d'argument ne nous conduirait nulle part. Car si l'idée d'un esprit unique n'est pas dépourvue d'application, elle s'applique à des individus ordinaires dont le cerveau est intact, et si elle ne s'applique pas à eux, elle devrait être mise au rencart, auquel cas il n'y a pas lieu de demander si ceux qui ont un cerveau scindé ont un esprit ou bien deux¹¹.

VI

Si ma position est correcte, et qu'il n'y a pas un nombre entier d'esprits individuels que ces patients puissent être dits avoir, alors l'attribution d'une activité mentale consciente et significative ne requiert pas l'existence d'un sujet mental unique. Ceci est en soi-même très troublant, parce que cela va à l'encontre de notre besoin de traiter les états mentaux que nous attribuons aux autres sur le modèle des nôtres. Il y a quelque chose dans la conception ordinaire de ce qu'est une personne, ou dans la conception ordinaire de l'expérience qui nous conduit à rechercher une explication de ces cas que cette conception même nous rend incapables de fournir. Cela peut paraître un problème qui ne mérite pas qu'on y prête une grande attention. Il n'est pas surprenant qu'ayant commencé avec un phénomène radicalement différent de tout ce que nous connaissions auparavant, nous soyons amenés à la conclusion qu'il ne peut pas être décrit dans des termes usuels. Cependant, je pense que la considération de cas très inhabituels de ce genre devrait nous conduire à être sceptiques

11. Au cas où quelqu'un serait enclin à souscrire à la conclusion que nous avons tous deux esprits, je suggérerai que la difficulté ne s'arrêterait pas là. Car les opérations mentales d'un hémisphère cérébral unique, telles que la vision, l'audition, le langage, l'écriture, la compréhension verbale, peuvent dans une large mesure être séparées l'une de l'autre par des déliations corticales appropriées; pourquoi alors devrions-nous considérer chaque hémisphère comme habité par plusieurs esprits coopérant entre eux et dotés de capacités spécifiques? Où devrions-nous nous arrêter? Si la décision quant au nombre d'esprits associés avec un cerveau est en grande partie arbitraire, la difficulté initiale a disparu.

quant au concept d'un sujet unique de conscience tel qu'il s'applique à nous-mêmes.

Le problème fondamental auquel nous sommes confrontés quand nous essayons de comprendre ces cas dans des termes mentalistes tient au fait que nous nous considérons nous-mêmes comme des paradigmes d'unité psychologique, et que nous sommes incapables de nous projeter nous-mêmes à l'intérieur de leur vie mentale, que ce soit une fois ou deux fois. Mais en nous servant ainsi de nous-mêmes comme pierre de touche quand nous demandons si un autre organisme peut être dit loger un sujet individuel d'expérience ou non, nous ignorons, sans y prendre garde, la possibilité que notre propre unité ne soit rien d'absolu, mais simplement un exemple parmi d'autres de coordination, plus ou moins effective, dans le système de contrôle d'un organisme complexe. Ce système parle à la première personne du singulier par nos bouches, et cela permet de comprendre le fait que nous puissions considérer son unité comme étant en un sens numériquement absolue, plutôt que relative et comme fonction de la coordination de ses contenus.

Mais c'est là une véritable illusion. L'illusion consiste à projeter à l'intérieur du centre de l'esprit le sujet même dont nous cherchons à expliquer l'unité : la personne individuelle avec toutes ses complexités. L'explication ultime de l'unité de ce que nous appelons esprit unique consiste dans une énumération des types de coordination fonctionnelle qui permettent de le diviser en types. Nous savons que ceux-ci peuvent s'émauser de différentes manières, à différents degrés. Croire que même dans leur version complète ils peuvent être expliqués par la présence d'un sujet numériquement unique est une illusion. Ou bien ce sujet est le réceptacle de la vie mentale, auquel cas il est complexe et on doit expliquer son unité en termes de l'opération unifiée de ses composants et de ses fonctions ou bien c'est un point sans extension, auquel cas il n'explique rien.

Un cerveau intact contient deux hémisphères cérébraux dont chacun possède des systèmes perceptifs, mémoriels et de contrôle adéquats pour diriger le corps sans l'assistance de l'autre. Ils coopèrent en le dirigeant avec l'aide d'un système de communication interne à deux voies agissant de façon constante. Les souvenirs, les perceptions, les désirs, et ainsi de suite, par conséquent, ont des bases physiques dédoublées des deux côtés du cerveau, et pas seulement en raison de similarités dans l'information qu'ils reçoivent initialement, mais en raison d'un échange qui a lieu par la suite. La coopération des hémis-

sphères non détachés dans l'exercice de contrôle du corps est plus efficace et directe que la coopération d'une paire d'hémisphères détachés l'un de l'autre, mais c'est une coopération néanmoins. Même si nous analysons l'idée d'unité en termes de coordination fonctionnelle, donc, l'unité de notre conscience sera moins claire que nous ne l'avions supposé. La conception naturelle d'une personne individuelle contrôlée par un esprit possédant un champ visuel unique, des facultés individuelles pour chacun des autres sens, des systèmes unitaires de mémoire, de désir, de croyance, et ainsi de suite, peut entrer en conflit avec les faits physiologiques quand nous les appliquons à nous-mêmes.

Le concept de personne peut éventuellement survivre à son application à des cas requérant que nous parlions de deux personnes ou plus dans un seul corps, mais elle semble nous engager fortement à admettre une forme quelconque de dénombrement. Comme ceci semble ouvert au doute, il est possible que l'idée ordinaire, simple, d'une personne individuelle, viendra à paraître bizarre un jour, quand les complexités du système humain de contrôle deviendront plus claires, et quand nous deviendrons moins certains du fait qu'il y a quelque chose de très important *dont* nous sommes l'unité. Mais il est aussi possible que nous soyons incapables d'abandonner cette idée quelle que soit la nature de nos découvertes ultérieures.

QUEL EFFET CELA FAIT, D'ÊTRE UNE CHAUVE-SOURIS?

C'EST la conscience qui fait que le problème corps-esprit* est vraiment difficile à résoudre. Peut-être est-ce pour cela que les discussions courantes de ce problème y prêtent peu attention ou se méprennent de manière aussi évidente à son propos. La vague récente d'euphorie réductionniste a produit plusieurs analyses des phénomènes mentaux et des concepts relatifs au domaine mental destinés à expliquer la possibilité d'une variété quelconque de matérialisme, d'identification psychophysique ou de réduction¹. Mais les problèmes auxquels on

* The mind-body problem (*N.d.T.*)

1. Par exemple J. J. C. Smart, *Philosophy and Scientific Realism*, London, Routledge & Kegan Paul, 1963; David Lewis, *An Argument for the Identity*

dans le second. Entre la deuxième et la troisième de ces attitudes il y a une ressemblance fondamentale quant à la source de la motivation.

Ces considérations, associées aux résultats négatifs de la première partie de cette enquête, devraient inciter à penser que dans la théorie morale comme dans la psychologie morale, ce n'est pas le saut kantien qui va du particulier et de l'affectif au rationnel et à l'universel qui fait toute la différence ; c'est plutôt le pas humien — c'est-à-dire le premier pas humien — qui va du moi vers l'autre.

Le moi et le futur

Imaginons un processus auquel seraient soumises deux personnes, A et B, au terme duquel on pourrait dire — en esquivant toute question — qu'elles ont fait l'échange de leurs corps. C'est-à-dire — pour préciser davantage — qu'on prend un certain corps humain qui est tel que, lorsque nous étions en sa présence, nous étions en présence de l'individu A ; certaines déclarations sorties de sa bouche correspondaient à des souvenirs des expériences passées de A ; certains de ses mouvements constituaient en partie les actions de A et s'interprétaient comme l'expression de la personnalité de A, et le reste à l'avenant. Et désormais, une fois l'échange accompli, ce qui sort de sa bouche paraît exprimer simplement les souvenirs qui auparavant étaient des souvenirs correspondant aux expériences de B, ses mouvements constituent en partie des actions caractéristiques de la personnalité de B, et le reste à l'avenant ; pour l'autre personne ce serait l'inverse.

Il y a certaines limites importantes imposées par la philosophie dans la manière de construire de telles fictions et d'en manipuler les variantes. Je ferai état de deux limites majeures, non pas pour en parler plus longuement, mais précisément pour m'en débarrasser.

Il y a certaines limites, surtout en ce qui concerne le caractère et les manières, à notre aptitude à imaginer de telles fictions, même si on s'en tient tout simplement à la disposition à admettre que les nouvelles prestations de ce corps qui était celui de A sont l'expression du carac-

tère de B ; si les individus d'origine A et B étaient extrêmement différents à la fois physiquement et psychologiquement, et si, en plus, ils étaient de sexe différent, on pourrait avoir beaucoup de mal à retrouver les dispositions de B dans toutes les actions possibles du Corps-de-A. Ignorons cela et, pour notre propos, retenons un A et un B suffisamment semblables (sans préciser) pour qu'il n'y ait pas de difficulté ; après l'expérience, les familiers de A et B sont simplement *sidérés* par le caractère très B des actes émanant de ce qui était auparavant le Corps-de-A, et inversement. Ainsi l'exploit qui consiste à imaginer un échange des corps est supposé possible au sens le plus étroit. Mais il y a encore une autre limite qui doit être franchie pour que l'exploit soit non seulement possible, au sens le plus étroit, mais encore ait comme résultat que nous soyons, après mûre réflexion, prêts à le décrire comme une permutation de corps entre A et B — c'est-à-dire un résultat tel que, mis en présence de ce qui était au préalable le Corps-de-A, nous soyons prêts à dire sérieusement que nous sommes en présence de B.

Pour que cela se produise, il semblerait qu'il soit nécessaire que les propos tenus par cette personne physique A soient considérés comme exprimant authentiquement la mémoire du passé de B. Or la mémoire est une notion qui repose sur la causalité ; et dans l'usage qu'on en fait, il semble qu'une condition nécessaire pour que le savoir actuel que X a des expériences passées de X constitue la mémoire de ces expériences, c'est que la chaîne causale liant les expériences et le savoir reste partie intégrante du corps de X. En raison de quoi, si des propos tenus par une personne physique donnée doivent être considérés comme explicitant le souvenir des expériences de B, il doit y avoir une forme de causalité pertinente entre l'état correspondant où se trouve cette personne physique et les expériences vécues à l'origine par B. Une façon radicale de s'assurer que la condition est remplie, dans le cas de l'échange imaginé, c'est de supposer, avec Schoemaker¹, que les cerveaux de A et de B sont échangés. On n'a pas besoin d'une

1. *Self-knowledge and Self-identity*, Ithaca, New York, 1963, p. 23 sq.

solution si radicale. Ainsi, supposons qu'il fût possible d'extraire toute l'information d'un cerveau humain, et de l'entreposer dans une machine, en attendant que ce cerveau fût réparé, voire remplacé, pour la réintroduire ensuite ; il paraîtrait excessif de vouloir que l'homme qui en résulterait ne puisse avoir les souvenirs qu'il avait avant l'opération. En ce qui concerne la connaissance que nous avons de notre passé, nous établissons des distinctions entre le fait de simplement se remémorer quelque chose celui de se voir rappeler quelque chose, et celui de réapprendre quelque chose. Ces distinctions correspondent (en gros) aux distinctions à faire entre une information à nulle donnée nouvelle, une information pour partie nouvelle et une information totalement nouvelle. Et il paraît clair que la mise en réserve de l'information que nous avons tout à l'heure imaginée n'introduirait pas de données nouvelles, qui sont nécessaires et suffisantes pour qu'on parle « d'apprendre de nouveau ». Ainsi donc, on peut décrire le cas qui nous occupe en parlant d'information extraite des cerveaux de A et de B, stockée et réintroduite avec permutation ; tel est le modèle que j'aurai à l'esprit, et qui me paraît convenir à mon argumentation.

Imaginons ce qui suit. Le processus détaillé ci-dessus fonctionne ; deux personnes peuvent, disons, entrer dans une machine et en sortir transformées d'une façon appropriée. Si A et B sont les personnes qui entrent, appelons celles qui sortent *Corps-de-A* et *Corps-de-B* ; *Corps-de-A* est cette personne (quelle qu'elle soit) devant laquelle je me trouve lorsque, après l'expérience, je me trouve devant le corps qui était auparavant celui de A, c'est-à-dire la personne qui serait naturellement prise pour A par quelqu'un qui ne ferait que voir cette personne, qui connaissait bien l'apparence de A avant l'expérience et qui ignorerait tout de l'expérience. Un compte rendu honnête de l'expérience laissera en suspens la question de savoir laquelle (s'il y en a une) des deux personnes, de A et de B, est *Corps-de-A*. La description de l'expérience comme celle de « personnes échangeant leurs corps » implique naturellement que *Corps-de-A* est en fait B.

Prenons deux personnes, A et B, qui vont être soumises à l'expérience (on peut supposer, en restant plutôt dans le vague, qu'elles le veulent bien ; préciser à ce stade pourquoi elles pourraient le vouloir

ou non, ce qu'elles redouteraient, etc., ce serait anticiper sur des points qui seront abordés plus tard). En plus, nous annonçons que, des deux personnes qui vont naître de l'expérience, Corps-de-A et Corps-de-B, l'une va se voir attribuer 100 000 dollars, tandis que l'autre sera torturée. Nous demandons alors à A comme à B de choisir quel traitement devrait être appliqué à chacune des personnes qui naîtra de l'expérience, en laissant (si possible) l'égoïsme inspirer leur choix.

Supposons que A choisisse que Corps-de-B connaisse le traitement de faveur et Corps-de-A le mauvais traitement, et que B fasse le choix inverse (ceci pourrait indiquer qu'ils considèrent que l'expression « échange des corps » décrit bien ce qui a dû se passer). L'expérimentateur ne peut agir de façon à satisfaire ces deux jeux de préférences, celles qui sont exprimées par A et celles qui sont exprimées par B. D'où il résulte qu'en clair, A et B ne peuvent tous les deux obtenir ce qu'ils désirent : si en effet l'expérimentateur, avant l'expérience, annonce à A et à B qu'il a l'intention de choisir (par exemple) d'infliger à Corps-de-B le mauvais traitement et d'accorder à Corps-de-A le traitement de faveur, A peut alors dire à juste titre : « Ce n'est pas le résultat que j'avais voulu », et B peut dire à juste titre : « C'est bien le résultat que j'avais voulu ». Ainsi, manifestement, A et B, avant l'expérience, peuvent chacun découvrir soit que le résultat qu'ils ont choisi sera celui qui se produira, soit que ce qu'ils ont choisi ne se produira pas, et, en ce sens, ils peuvent obtenir, ou échouer à obtenir, ce qu'ils voulaient. Mais est-il aussi vrai, lorsque l'expérimentateur, après l'expérience, agit en accord avec l'un des choix et non avec l'autre, qu'à ce moment-là l'un des deux (A ou B) aura ce qu'il avait voulu et l'autre non ?

Il semble qu'il y ait de solides raisons de le penser. En effet, supposons que l'expérimentateur, ayant eu connaissance du choix de A et de celui de B, ne dise rien à A ni à B de ce qu'il va faire ; il réalise l'expérience ; puis, par exemple, il inflige le mauvais traitement à Corps-de-B et accorde le traitement de faveur à Corps-de-A. A ce moment-là, Corps-de-B non seulement se plaindra du mauvais traitement en tant que tel, mais il se plaindra (puisque'il a la mémoire de A) du fait que ce n'est pas le traitement qu'il avait choisi, puisqu'il avait choisi que

Corps-de-B fût bien traité ; et puisque A avait fait son choix en tout égoïsme, il peut ajouter qu'il avait précisément choisi ainsi parce qu'il ne voulait pas que le mauvais traitement lui fût infligé à lui. Pendant ce temps-là, Corps-de-A exprimera sa satisfaction, d'une part parce qu'il recevra 100 000 dollars, d'autre part parce que l'expérimentateur a choisi d'agir dans le sens que lui, B, avait si judicieusement choisi. Tout cela nous incite fortement à dire que l'expérimentateur a fait en sorte que B au bout du compte ait obtenu ce qu'il voulait et A non. Cela nous incite donc fortement à dire que Corps-de-B est en réalité A et que Corps-de-A est en réalité B ; donc que le processus en jeu dans l'expérience est bien un processus d'échange de corps. Pour les mêmes raisons, on pourrait dire que A et B, dans notre exemple, avaient, en réalité, fait tous les deux un choix judicieux, et que c'est par malchance que le choix que A avait correctement fait ne s'est pas concrétisé, et par chance que le choix que B avait correctement fait s'est, lui, concrétisé. Cela semblerait démontrer que se soucier de ce qui m'arrivera dans le futur ne signifie pas nécessairement se soucier de ce qui arrivera à ce corps (celui que j'ai maintenant) ; et, derechef, pourrait servir à montrer que, en un certain sens que peut avoir l'obscur formule de Descartes, moi et mon corps sont « réellement distincts » (encore que rien, naturellement, dans ces considérations, ne vienne appuyer l'idée que je pourrais exister sans corps du tout).

Ces réflexions paraissent renforcées si nous considérons les cas où A et B font d'autres choix dans le cadre de l'expérience. Supposons que A choisisse que Corps-de-A reçoive l'argent et que Corps-de-B ait la souffrance en partage. Et que B fasse le choix inverse. Ici non plus, il n'y a aucun résultat qui s'accorde aux deux choix à la fois : nos sujets ne peuvent tous les deux obtenir ce qu'ils veulent. L'expérimentateur annonce, avant l'expérience, que Corps-de-A recevra l'argent et que Corps-de-B aura la souffrance en partage. Ainsi A, à ce stade, obtient ce qu'il veut (le résultat annoncé correspond au choix qu'il a exprimé). Après l'expérience, la répartition est faite comme prévu. Corps-de-A et Corps-de-B devront admettre que ce qui arrive est conforme au choix que A avait fait au départ. Corps-de-B reconnaîtra tout naturellement cela (puisque'il a sa mémoire de A) en disant que

c'est la répartition qu'il avait choisie; il se rappellera, entre autres choses, que l'expérimentateur avait annoncé ce résultat, qu'il l'avait approuvé comme étant ce qu'il avait choisi, etc. Cependant lui-même (Corps-de-B) n'apprécie certainement pas ce qui lui arrive, et aimerait mieux recevoir ce que Corps-de-A reçoit, c'est-à-dire 100 000 dollars. De son côté, Corps-de-A se souvient d'avoir fait un choix différent de ce qui s'est produit, mais il considère comme de la chance que l'expérimentateur n'ait pas fait ce que lui-même avait choisi. Tout se passe donc comme si Corps-de-A obtenait ce qu'il veut mais qu'il n'avait pas choisi, tandis que Corps-de-B obtient ce qu'il avait choisi mais qu'il ne veut pas. Une fois de plus il apparaît qu'ils sont bien, respectivement, B et A, et que, dans ce cas de figure, les choix originaux de A et de B n'étaient pas judicieux.

Supposons, en dernière hypothèse, que dans le choix original A se comporte comme dans le premier cas, B comme dans le second. A savoir : A choisit que Corps-de-B reçoive l'argent et que Corps-de-A ait la souffrance en partage, et B fait exactement le même choix. Dans ce cas, l'expérimentateur paraît se trouver dans la situation agréable d'exaucer les désirs de chacun — ou au moins, comme Dieu, les vœux de chacun. Dans ce cas, Corps-de-B est heureux de ce qu'il reçoit, se rappelle l'avoir choisi, et se félicite de la sagesse (pense-t-il) de son choix; dans le même temps, Corps-de-A est malheureux de ce qu'il reçoit, se rappelle l'avoir choisi et est forcé de reconnaître que son choix était malavisé. Ainsi une fois de plus nous avons une situation qui conforte les conclusions tirées de notre première hypothèse.

Considérons maintenant la question non plus du choix par A et B du résultat de l'expérience, mais de leur acceptation initiale de l'expérience. S'ils sont au départ prêts à accepter le programme de l'expérience comme celui « d'un échange des corps », une chose susceptible de les intéresser pourra être l'apparence du corps de l'autre. Sous cet angle aussi, ce qui arrivera après l'expérience peut suggérer que l'expression « échange des corps » était une bonne description de cette expérience. Si A et B acceptent l'expérience, c'est que chacun d'entre eux n'est pas sans apprécier l'apparence, le physique et autres aspects du corps de l'autre. Après l'expérience, Corps-de-B pourrait se trou-

ver dire quelque chose comme : « Quand j'ai donné mon accord à l'expérience, je pensais que le visage de B était très beau, mais maintenant que je le vois dans la glace, je n'en suis plus très sûr »; ou Corps-de-A pourrait dire : « Quand j'ai donné mon accord, je ne savais pas que A avait une jambe de bois; mais maintenant que tout est fait, je me trouve avec cette jambe de bois et je veux revenir à la situation initiale. » Il se peut qu'il ajoute qu'il trouve la jambe de bois très gênante et que Corps-de-B lui réponde, par exemple, qu'il se rappelle l'avoir trouvée très gênante au début mais qu'on s'y habitue; mais peut-être faudrait-il en savoir un peu plus que moi, quant à la physiologie de l'accoutumance aux membres artificiels, pour savoir si Corps-de-A trouverait la jambe de bois gênante ou non : ce corps-là, après tout, a cette jambe-là depuis un certain temps. Mais ce genre de détail mis à part, le résultat dans son ensemble, considéré sous cet aspect, paraît bien confirmer nos précédentes conclusions touchant cette expérience.

Supposons maintenant que lorsque l'expérience leur est proposée (en termes qui ne supposent pas la question résolue) A et B pensent plutôt à leurs avantages et à leurs désavantages psychologiques. A pense d'abord à certaines formes d'anxiété auxquelles il est très sujet, tandis que B s'interroge à propos des souvenirs atroces d'aventures passées qui continuent à le tourmenter. Chacun d'eux espère que l'expérience, d'une façon ou d'une autre, aura pour effet de lui permettre de se débarrasser de ces problèmes. Peut-être ont-ils été sensibles à des argumentations philosophiques disant qu'une continuité physique est au minimum une condition nécessaire de l'identité personnelle : A, par exemple, se dit que, pour peu que l'expérience réussisse, la personne qui va prendre en charge son corps n'aura pas cette anxiété et que, si l'autre personne connaît sans aucun doute l'anxiété — peut-être son anxiété, d'une certaine façon —, au moins cette personne-là ne sera pas lui. L'expérience a lieu et l'expérimentateur (à qui A et B ont au préalable révélé en privé leurs divers problèmes et leurs divers espoirs) demande à Corps-de-A s'il est délivré de son anxiété. L'autre lui répondra sans doute qu'il ne comprend pas ce qu'il veut dire. Il n'a jamais eu d'anxiété; par contre il avait bien quelques souvenirs très

désagréables, il se rappelle qu'il s'est lancé dans l'expérience pour s'en débarrasser, et il est déçu de découvrir qu'il les a toujours. Corps-de-B réagira de la même façon aux questions ayant trait à ses souvenirs pénibles, faisant remarquer qu'il a toujours, lui, son anxiété. Ces résultats confirment un peu plus la description de l'expérience « d'échanges des corps ». Et l'ensemble des résultats fait penser que la seule chose rationnelle à faire, en présence d'une telle expérience, serait de chercher son identité dans ses souvenirs et autres choses semblables et non dans son corps. L'argumentation philosophique censée montrer que la continuité physique est au moins une condition nécessaire de l'identité personnelle semblerait simplement erronée.

Considérons maintenant quelque chose d'apparemment différent. Quelqu'un, au pouvoir de qui je suis tombé, m'annonce que je vais être torturé le lendemain. Je suis effrayé et pense au lendemain avec une grande appréhension. Il ajoute que, le moment venu, je ne me souviendrai pas d'avoir été averti de ce qui devait m'arriver. En effet, peu de temps avant la torture, on me soumettra à un autre traitement qui me fera oublier ce qui m'avait été annoncé. Ceci ne va certainement pas me reconforter, étant donné que je sais parfaitement que je peux oublier des choses, et que je serai en tout état de cause torturé à l'improviste parce que j'aurai oublié ou qu'on m'aura fait oublier l'annonce anticipée de cette torture : ce sera toujours une torture que j'attends dans l'épouvante, aussi longtemps que j'ai conscience de l'annonce anticipée. L'expérimentateur ajoute alors que mon oubli de l'annonce ne sera qu'un élément d'un processus plus général : lorsque l'heure de la torture arrivera, je n'aurai aucun souvenir d'aucune sorte de ce que je suis actuellement en mesure de me rappeler. Cela ne me reconforte pas, là non plus, puisque je puis facilement m'imaginer victime d'un accident, par exemple, qui me plongerait à la fois dans un état complet d'amnésie et dans des douleurs intenses ; cela pourrait assurément m'arriver, je n'aimerais pas que cela m'arrivât, ni savoir que cela dût m'arriver. L'expérimentateur ajoute alors qu'à l'heure d'être torturé, non seulement je ne me souviendrai pas de ce dont je suis en mesure de me souvenir actuellement, mais que j'aurai une vision différente de mon passé, très différente de la mémoire que j'en

ai maintenant. Je ne crois pas que cela me reconforterait non plus. Car je peux concevoir, à défaut de sa réalisation concrète, du moins la possibilité de devenir complètement fou et de me prendre peut-être pour Georges IV ou pour quelqu'un d'autre ; et si on me dit que c'est cela qui va m'arriver, on ne diminuera pas la terreur que j'éprouve à l'annonce catégorique que je vais être torturé, mais cela ne fera qu'intensifier l'horreur. Et je ne vois pas pourquoi l'expérimentateur me mettrait dans de meilleures dispositions s'il me disait, pour finir, que le sentiment du passé que j'aurai en moi au moment d'être torturé correspondra exactement au passé d'une autre personne actuellement en vie, et que ce sentiment du passé me sera donné par le transfert (par exemple) des informations stockées par son cerveau dans le mien. Assurément la terreur resterait la réaction appropriée ; non parce qu'on ne saurait pas ce qui va se passer, mais parce que sur un point vital au moins on saurait ce qui va se passer — la torture qu'on peut effectivement s'attendre à subir et à voir précédée, en plus, par certains bouleversements mentaux.

S'il en est ainsi, la question dans son ensemble paraît maintenant tout à fait mystérieuse. Car l'exemple que nous venons d'explorer n'est, en fait, qu'un aspect, abordé différemment, de l'opération que nous avons étudiée plus haut. Et elle la représente comme une éventualité parfaitement détestable, alors que les considérations précédentes la représentaient comme quelque chose qu'on pouvait choisir, peut-être même avec enthousiasme, parmi les options présentées alors. Cette opération est présentée différemment, c'est vrai, et sur deux points capitaux, mais lorsque nous regardons ces deux différences de présentation, pouvons-nous réellement nous convaincre que la seconde présentation soit fautive et nous induise en erreur, laissant ainsi le champ libre à la première version qui, tout à l'heure, nous paraissait si convaincante ? Certainement pas.

La première différence, c'est que dans la seconde version, la torture est, d'un bout à l'autre, représentée comme quelque chose qui va m'arriver à moi ; l'expérimentateur dit constamment « vous ». Ainsi il n'est pas très neutre. Mais eût-il dû être neutre ? Ou, pour le dire autrement, est-ce que l'emploi qu'il fait de la seconde personne a sim-

plement un effet émotionnel et rhétorique sur moi, me donnant de la frayeur quand une réflexion plus poussée m'eût montré qu'il n'y avait pas de quoi en avoir ? Ce n'est certainement pas évident. Le problème est que, au fur et à mesure qu'il avance dans ses prédictions, je semble capable de le suivre sans difficulté. Et si je me demande si ce qu'il a dit me donne des raisons d'avoir peur d'être torturé, je pourrais considérer que, derrière ma peur, on trouve quelque principe de ce genre : le fait que je subisse une douleur physique dans le futur n'est pas annulé par l'état psychologique qui pourra être le mien à ce moment-là, à l'exception triviale de ces états psychologiques qui par eux-mêmes excluent l'expérience de la douleur, ce qui est surtout le cas de l'inconscience (à supposer qu'il s'agisse là d'un état psychologique). En particulier les impressions que je conserve du passé n'auront aucune incidence sur ma souffrance. Le principe paraît assez sensé.

Il est important de noter qu'il n'est pas vrai que tout ce que je considère, dans l'état actuel des choses, comme un mal fût à craindre rationnellement comme un mal si l'on me prédisait que cela dût m'arriver dans le futur mais après que j'aurais, dans l'intervalle, connu des métamorphoses psychologiques importantes. Car le fait que je considère l'événement, dans l'état actuel des choses, comme un mal, peut dépendre de facteurs liés à des opinions ou à un tempérament qui pourraient eux-mêmes se trouver modifiés par les métamorphoses psychologiques en question. Ainsi, si je suis affreusement sujet à l'acrophobie et qu'on me dise que je vais dans un avenir prochain me retrouver au sommet d'un à-pic, je serai dans un premier temps effrayé ; mais si on me dit que je serai psychologiquement métamorphosé entre-temps de façon à être délivré de mon acrophobie (et que, comme pour l'autre prédiction je le croie), alors je n'ai pas de raison de m'effrayer de l'événement prédit, en tout cas je n'ai pas la même raison. Dans le même ordre d'idées, je pourrais me préparer à revoir une certaine personne avec appréhension ou bien avec joie à cause des souvenirs que j'ai de nos relations passées. Pour une part, les souvenirs ont quelque chose à voir avec mon émotion, non seulement dans le présent, mais dans la projection que je fais de l'avenir : car c'est à une rencontre elle-même affectée par la présence de ces souvenirs que je

me prépare. Si je suis convaincu que, le moment venu, je n'aurai pas ces souvenirs, je n'aurai pas exactement les mêmes raisons qu'avant pour attendre cette rencontre dans l'une ou l'autre de ces dispositions d'esprit (par parenthèse, le spiritualisme m'invite à croire qu'après ma mort j'aurai les mêmes raisons qu'avant d'être dans telle ou telle disposition à l'idée de retrouver des êtres : avec l'unique modification que je peux être sûr que tout se passera bien).

La douleur physique, cependant, exemple que j'ai choisi pour sa simplicité (et non pour quelque raison obsessionnelle), n'a qu'une relation absolument minimale avec le caractère ou les croyances. Aucune modification, si importante soit-elle, de mon caractère ou de mes croyances, n'atténuerait l'âpreté de tortures subies par moi ; corrélativement, aucune annonce de changements survenus dans mon caractère et dans mes croyances, si importants soient-ils, ne peut déloger la peur de la torture qui m'est promise en même temps que ces changements.

Je ne veux pas du tout laisser entendre que la seule base ni, à plus forte raison, la seule base rationnelle sur laquelle repose la peur que j'ai de ces différentes prédictions soit celle du lien entre les événements et les dispositions psychologiques dans lesquelles je serai, quand ils finiront par arriver. Je fais seulement remarquer que c'est une composante de cette peur ; ce n'est pas la seule. Car certainement on craindra et on repoussera de quelque autre façon les changements eux-mêmes, ou, dans bien des cas, on voudrait le faire. D'où ce vieux paradoxe de l'utilitarisme hédoniste ; si on recevait l'assurance qu'en subissant certaines opérations et en étant relié à une machine on connaîtrait jusqu'à la fin de ses jours une série continue d'expériences délicieuses et variées, on pourrait bien repousser une telle proposition, et on réagirait par la peur si quelqu'un se proposait de nous y contraindre. Et que, dans cette dernière hypothèse, la peur et l'horreur apparaissent comme étant des réactions justifiées, nous autorise à disqualifier l'interprétation (au cas où quelqu'un aurait le cran de la proposer) d'après laquelle on repousserait volontairement la proposition d'un tel bonheur parce qu'on aurait conscience de ses devoirs envers autrui, devoirs que l'état hédoniste ferait abandonner. La perspective d'une vie d'aliéné ou d'une vie végétative béates est jugée horrible par la

plupart des hommes (peut-être pas par tous) : ce sentiment n'est manifestement pas lié à la façon dont la situation leur apparaîtrait à ce moment-là, car la situation, à ce moment-là, ne leur paraîtrait pas horrible. Pour le cas dont nous étions en train de discuter, ces dernières considérations font simplement mieux comprendre que les annonces de l'expérimentateur provoquent une double raison d'être horrifié : la perspective d'être torturé, et la perspective avant d'être torturé de changer de caractère et de mémoire. Et assurément, pour répéter ce qui a déjà été dit, la seconde perspective ne fournit ni matière à rejeter, ni matière à ne pas redouter la première.

J'ai dit qu'il y avait deux différences appréciables entre la seconde façon de présenter la situation d'échange des corps et la première. La première différence dont nous venons de parler était que l'expérimentateur m'annonçait que la torture m'atteindrait *moi*, un « moi » qui serait alors psychologiquement très changé. Il nous faut encore trouver une raison de dire qu'il n'aurait pas dû l'annoncer, ou que je serais incapable de comprendre ce qu'il dit s'il le fait ; or je ne semble que trop capable de le comprendre. La seconde différence est que dans cette présentation, l'expérimentateur ne parle pas de l'autre, sinon pour évoquer le rôle quelque peu marginal de celui qui fournit les impressions du passé dont je me retrouve l'hôte. Il n'en parle pas du tout comme de quelqu'un qui se retrouvera l'hôte des impressions du passé qui est mien (et, accessoirement, le bénéficiaire des 100 000 dollars — considération qui, dans l'état d'esprit que m'attribue cette version, me rendra simplement jaloux).

Mais pourquoi l'expérimentateur *devrait-il* parler de cet autre homme et de ce qui va lui arriver ? Mon souci égoïste est qu'on me dise ce qui va m'arriver à moi, et cela je le sais : la torture, précédée de changements de caractère, d'opérations du cerveau, de changements de souvenirs. Savoir qu'une autre personne — ou nulle autre, ou beaucoup d'autres — sera maltraitée comme moi peut m'affecter d'autre façon : provoquer de la compassion, une horreur plus grande devant le pouvoir de ce tyran, et semblables autres sentiments ; en tout état de cause cela ne peut sûrement pas avoir d'effet sur ma façon d'attendre la torture — n'est-il pas vrai ? Mais — dira un objecteur —

voilà qui est abandonner précisément la donnée qui, comme la première présentation du cas l'a montré, fait toute la différence : car c'est abandonner la personne qui, comme la première présentation l'a montré, sera vous. Ce n'est pas seulement abandonner une donnée qui devrait avoir un effet fondamental sur vos peurs, c'est abandonner la personne pour laquelle vous avez peur. Naturellement, conclura notre objecteur, cela fait toute la différence.

Mais est-ce possible ? Considérons la série de cas suivants. Dans chaque cas nous allons supposer qu'après ce qui est décrit, A, comme plus haut, va être torturé ; nous supposons également que la personne A est informée par avance précisément que les choses suivantes vont lui arriver avant qu'elle soit soumise à la torture :

- 1 / A subit une opération qui entraîne une amnésie totale ;
- 2 / une amnésie est provoquée chez A, et d'autres manipulations entraînent certains changements de caractère ;
- 3 / des changements de caractère sont provoqués chez A, et en même temps des souvenirs « illusoires » lui sont communiqués : ils sont fictifs et ne correspondent à aucune personne existante ;
- 4 / même chose qu'en 3, à ceci près que les traits de caractère et les impressions mnésiques sont fabriqués de façon à correspondre à une autre personne réelle, B.
- 5 / Même chose qu'en 4, à ceci près qu'on obtient le résultat en insérant dans le cerveau de A l'information stockée dans le cerveau de B. La méthode employée conserve B dans son état antérieur.

6 / même chose qu'en 5 pour A. Mais B ne conserve pas son état antérieur puisque l'opération est effectuée dans les deux directions.

Je considère comme acquis que personne ne va contester que A ait des raisons, et des raisons assez simples, de craindre la douleur quand il s'agit de l'hypothèse 1 ; je ne vois pas pourquoi ceci ne pourrait s'élargir à 2 ; 3 ne saurait certainement pas introduire de différence de fond — c'est seulement une situation qu'on serait fondé, pour plus d'une raison, ainsi qu'on l'a vu plus haut, à craindre davantage. La situation 4, au moins, introduit la personne B, qui était au cœur de l'objection que nous sommes en train de discuter. Mais B ne paraît pas introduit d'une façon qui apporte une différence tangible ; si je m'attends à

une souffrance par delà une métamorphose qui comporte de nouvelles impressions mnésiques, ce sera une donnée purement extérieure, relativement à cela, que les impressions mnésiques soient conformes à un modèle. La situation 4 ne remplit pas non plus la condition de relation causale que j'avais signalée au début, pour que les « souvenirs » soient vraiment des souvenirs ; encore que vous remarquerez que si le travail est fait correctement, je pourrais sans doute tirer de Corps-de-A des remarques au sujet de ce qu'il attendait initialement de l'expérience — remarques appartenant originellement à B — et qui avaient eu tant de poids pour nous dans la première version de l'histoire. J'aurai l'assurance qu'il en est également ainsi en 5, qui nous fournit, en plus, une application recevable de la condition de relation causale.

Toutefois cette situation appelle deux remarques. D'abord, si nous nous restreignons à A et à Corps-de-A, nous ne paraissions pas avoir ajouté quoi que ce soit qui entraîne une différence tangible dans ses peurs. De même que, à passer de 3 à 4, on ne voyait pas de différence significative dans le fait que les nouvelles « impressions mnésiques » précédant la souffrance eussent été, comme c'était le cas, conformes à un modèle, de même, en passant de 4 à 5, tout ce que nous avons ajouté c'est qu'elles ont un modèle qui est aussi leur cause : et on ne voit toujours pas en quoi pourrait en dépendre le fait que celui qui anticipe l'avenir s'attende à souffrir ou à ne pas souffrir. Pour illustrer cela, plaçons-nous sur le terrain du caractère : si A a la capacité de s'attendre à souffrir, il a celle de s'attendre à une souffrance précédée d'une transformation de ses dispositions — et rien ne sera modifié, dans son attente, par le fait que la transformation de ses dispositions soit calquée sur, ni, à plus forte raison, indirectement causée par les dispositions d'une autre personne. Si sa peur peut, pour ainsi dire, survivre au changement, la façon dont le changement est conduit est simplement anecdotique. La deuxième remarque concernant 5, c'est que si la question cruciale qui tourmente A à propos du destin de Corps-de-A est de savoir si Corps-de-A sera ou non B¹, 5 ne répond pas

1. Cela n'est pas nécessairement la question décisive, mais il serait correct de faire comme si c'était le cas, pour répondre à l'objection dont il est question ici.

encore à cette condition : car en 5, nous avons encore un B incontesté à côté de Corps-de-A, et ces deux-là ne sont certainement pas la même personne.

Mais dans la situation 6, pensions-nous, A est finalement B. Toutefois, si les craintes initiales de A pouvaient survivre aux changements attendus en 5, comme elles le faisaient en 4 et en 3, elles peuvent aussi leur survivre en 6. Effectivement, si on se place du point de vue de A, de son attente et de ses craintes, il y a moins de différences entre 6 et 5 qu'il n'y en a entre 5 et 4 ou entre 4 et 3. Dans ces derniers passages, il y avait au moins des différences — encore qu'on n'ait pas vu qu'elles fussent vraiment déterminantes — dans le contenu ou la cause de ce qui arrivait à A. Dans le cas présent, il n'y a absolument aucune différence quant à ce qui lui arrive. La seule différence est dans ce qui arrive à quelqu'un d'autre. Si A peut craindre la souffrance quand on lui décrit 5, pourquoi cesserait-il de le faire en 6 ?

Je ne vois qu'une seule façon d'accorder une importance décisive au passage de 5 à 6, et cela ne va pas sans grande difficulté. Ce serait de nier que le passage de 5 à 6 ajoute seulement, comme je le dis, quelque chose qui arrive à *quelqu'un d'autre* ; ce qui se passe plutôt, continuera-t-on à m'objecter, c'est que ce passage réintroduit A lui-même, sous la forme de Corps-de-B ; puisqu'il réapparaît sous cette forme, c'est cette personne, et non l'infortuné Corps-de-A qui focalisera les attentes de A. C'est, dans les faits, réaffirmer le point de vue développé dans notre première présentation de l'expérience. Or cela entraîne assurément la conséquence que A ne doit pas craindre pour corps de A qui intervient en 5. Car d'après l'argumentation dont il s'agit, Corps-de-A, en 6, n'est pas A ; c'est Corps-de-B qui est A. Mais Corps-de-A, en 5, est, pour ce qui est du caractère, de l'histoire et du reste, exactement le même que Corps-de-A en 6 ; donc si ce dernier n'est pas A, le premier, en 5, n'est pas A non plus (c'est ce détail, certainement, qui pousse à décrire ce qui se passe en 6 comme étant la *réintroduction* de A). Mais personne d'autre que Corps-de-A, en 5, n'a de meilleur titre à être A. Aussi, en 5, A semble-t-il ne pas exister. Ceci tendrait à expliquer pourquoi A n'a pas lieu de redouter 5 — encore qu'il puisse craindre le cours des choses qui y conduit. Mais

il semblait, un peu plus haut, qu'il avait quelque raison de craindre 5. Admettons, toutefois, que cela ait été une illusion, et que A n'existe véritablement pas en 5 ; mais alors existe-t-il en 4, en 3, en 2 ou en 1 ? Il paraît difficile de le nier pour 1 et 2 ; allons-nous tracer la ligne de démarcation entre 3 et 4 ?

Ici quelqu'un me dira : vous n'avez pas à vouloir tracer une ligne de démarcation — les cas limites sont les cas limites et vous n'avez pas à pousser nos concepts au-delà de leurs limites. Mais cet avertissement bien connu, aussi pertinent soit-il dans bien des cas, semble, dans le cas présent, entraîner une très sérieuse difficulté. Il peut rassurer intellectuellement ceux qui observent la situation de A ; Mais qu'est-ce que A est censé en tirer ? Si on me dit qu'une situation à venir est un cas limite quant au fait de savoir si c'est moi qui serai touché, qu'on ne peut décider intellectuellement si ce sera moi ou non, voilà quelque chose dont je ne peux rien tirer ; en particulier parce que cela ne peut correspondre à aucune représentation concevable dans le cadre de mes prévisions et des émotions qui les accompagnent.

Si je m'attends à ce qu'une certaine situation, S, se réalise dans le futur, il y a évidemment un large spectre d'émotions et d'inquiétudes, relatives à S, que je vais peut-être éprouver maintenant en liaison avec mon attente. A moins que ne je sois extraordinairement égoïste, il n'est pas nécessaire, pour que cette attente provoque mon inquiétude, que je sois personnellement concerné par S — l'expression « concerné par S » signifiant que je participerai à S comme quelqu'un qui, à ce moment-là, fera quelque chose ou subira quelque chose, ou encore que S aura des conséquences que je subirai à ce moment-là ou plus tard. Il y a des émotions, toutefois, que je ressentirai seulement si je suis concerné par S, et la peur en est un exemple manifeste.

Par ailleurs, la forme que mon attente donne à S sera nécessairement, à plus d'un égard, indéterminée ; en particulier elle sera indéterminée en ceci qu'elle ne décide pas si je serai concerné par S ou non. Ainsi, je peux avoir un sérieux motif de penser que l'un de nous cinq va être blessé, mais aucun motif de croire que ce soit moi plutôt qu'un autre. Les émotions que je vais ressentir seront corrélativement affectées par cette incertitude. Ainsi, pour m'en tenir aux préoccupa-

tions égoïstes liées à la peur, je vais être probablement un peu plus détendu que si j'étais certain d'être la victime, et un peu moins que si je restais tout à fait en dehors de l'histoire. La crainte sera mêlée d'appréhension, qui lui donnera un caractère particulier ; le reste à l'avenant. La ronde des émotions est réglée par ce que j'imagine de la résolution ultérieure de l'indéterminé : moments de terreur pure quand je pense que je serai visé ; moments d'espoir quand je pense que ce ne sera pas moi. Toutes les émotions sont liées à la réalisation de ce que j'attends, et ce que j'attends, dans la situation ci-dessus, ne peut se réaliser que sous l'une des deux formes prévues.

Il y a d'autres cas de figure où des attentes indéterminées peuvent être liées à la peur. Ainsi je puis m'attendre (névrotiquement peut-être) à ce que quelque chose de très déplaisant soit sur le point de m'arriver ; je peux même m'attendre à ce que cela prenne une forme donnée, sans avoir pour autant à ma disposition de série ou en tout cas de série limitée à passer en revue mentalement pour donner un visage à cette forme. Bien différente serait la peur de quelque chose de totalement indéterminé — la peur (pourrait-on dire) d'une horreur sans nom. Si quelqu'un avait une telle peur, on pourrait même dire qu'il a, en un certain sens, une attente parfaitement définie : si ce à quoi il s'attend se produit effectivement, il n'y aura rien de plus défini à dire après l'événement, que ce qui en était dit quand cet événement était attendu. Naturellement, ces deux cas sont des cas de *peur*, parce que la seule chose qui soit définie au milieu de cette indétermination est la croyance que c'est moi que les événements vont toucher.

Ce qui est au cœur de l'attente de S, c'est la pensée de ce que sera S lorsque cela se réalisera — pensée qui peut ne pas être déterminée, qui peut balancer entre des alternatives, etc. Lorsque S me concerne, cette pensée peut prendre une forme particulière : pensée de ce que sera l'événement pour moi, projection imaginée de moi plongé dans S¹. Je n'ai pas nécessairement à penser à S de cette façon quand je suis concerné par S, mais je peux être capable de le faire (quelqu'un

1. Pour une discussion plus approfondie de ces problèmes, voir *Imagination and the Self in Problems of the Self*, Cambridge University Press, 1973, p. 26-45.

pourrait me dire que cette possibilité se reflète même dans la langue, dans la distinction entre « s'attendre à être blessé », et « s'attendre à ce qu'on soit blessé » ; mais je suis très réservé là-dessus, et, de toute façon, c'est un point sans intérêt).

Supposons maintenant qu'il y ait une sorte de S dont il ne soit pas possible, pour des raisons abstraites, de décider si S me concerne ou non, comme l'impose à l'expérimentation la méthode de réflexion que nous discutons. Il faut faire remarquer que l'attente de S n'est pas *indéterminée* aux différents sens que nous venons de considérer. Ce n'est pas comme à propos de l'horreur sans nom puisque, dans ce dernier cas, la caractéristique définie était que cela devait arriver au sujet, et cela faisait qu'il éprouvait sans aucune ambiguïté une grande peur. Cela ne rappelle pas davantage ce à quoi s'attend l'homme qui a une chance sur cinq d'être blessé ; sa peur était certes ambiguë, mais son point d'application, et celui de son attente, était que lorsque S en viendrait à se réaliser, ce à quoi il s'attend se réaliserait assurément, d'une façon ou d'une autre. Dans le cas présent, la peur (la peur de la torture, s'entend, non la peur de l'expérience préalable) ne semble ni justifiée, ni injustifiée, ni d'une ambiguïté elle-même justifiée. Corrélativement, le sujet a un mal insurmontable à savoir que penser de S. S'il s'engage dans une projection imaginaire (projection de ce que ce sera pour lui), il répond implicitement à la question qui est nécessairement sans réponse possible ; s'il pense qu'il ne peut s'engager dans une telle réflexion, il y a tout lieu de considérer qu'il donne encore une réponse, quoique de sens contraire. Peut-être doit-il simplement s'abstenir d'une telle réflexion, mais est-ce pour lui simplement s'abstenir d'une telle réflexion, quand il est irrémédiablement impossible de décider s'il peut ou s'il ne peut pas s'y engager ?

On pourrait objecter que tout ce que ces considérations peuvent entraîner, c'est que la peur, certes, n'a pas de fondement assuré dans ce cas-là ; mais qu'il pourrait y avoir quelque autre forme, moins tranchée, d'inquiétude, qui, elle, serait adaptée à cette attente particulière qu'est l'attente d'une situation intellectuellement indécidable ; on objecterait qu'il y a peut-être des impressions de cette sorte qui existent réellement, dans des situations réelles. Ainsi des objets matériels

subissent parfois des métamorphoses surprenantes qui ne permettent plus de se faire une idée claire de leur identité. Supposons que je sois sentimentalement attaché à un objet qui, par la suite, connaisse ce sort-là ; peut-être ne pourrais-je ni le considérer tout à fait comme je faisais auparavant, ni m'en désintéresser totalement, mais éprouver à son égard quelque autre sentiment plutôt ambivalent. De la même façon, pourrait-on ajouter, en face de celui qui est la future victime de la souffrance, et dont je ne peux savoir clairement quelle relation d'identité il aura avec moi, je ne peux ressentir ni ce que je ressentirais si j'étais assuré qu'il fût moi, ni ce que je ressentirais si j'étais assuré qu'il ne fût pas moi ; je ressentirais plutôt quelque inquiétude ambiguë.

Mais cette analogie n'a guère le pouvoir de régler le problème le plus déroutant que pose le cas de figure considéré — problème qui a déjà été rencontré lorsqu'on a parlé de la difficulté du sujet à réfléchir à la situation en s'impliquant ou non. Car ne considérer la future victime de la souffrance que comme l'objet du sentiment après que cet objet a subi une métamorphose, et assimiler le désarroi ambigu causé par la souffrance future de notre victime au désarroi ambigu causé par quelque dommage infligé au dit objet, c'est établir entre la victime et moi une franche séparation et ainsi déplacer le point obscur du problème là où il n'a pas lieu d'être. Je dois me rapprocher davantage de la victime. Mais est-il possible que je m'en approche sans m'attendre à endurer sa souffrance ? Si c'est possible, l'analogie ne l'a pas démontré. On ne peut certainement pas s'en rapprocher en s'attendant à une souffrance pour ainsi dire ambiguë ; cela est totalement exclu. Il semble qu'il y ait un obstacle insurmontable à ce que je me figure par anticipation une situation dans laquelle la réflexion ne peut décider si je joue ou non un rôle.

Cette perplexité paraît, qui plus est, tourner à la franche absurdité si on passe de l'indécidabilité réflexive, à son amie et relation proche, la décision conventionaliste. Cela se révèle si l'on considère une autre description, clairement conventionaliste, de la série de cas qui a suscité la discussion présente. Cette description refuserait une conclusion sur laquelle je m'étais appuyé dans une argumentation précédente — à savoir que si nous refusons que Corps-de-A, en 6, soit A (puisque c'est

Corps-de-B qui l'est), nous devons alors nier que Corps-de-A, en 5, soit A, puisqu'ils sont exactement semblables. « Non », objectera-t-on, « voilà qui est considérer que nous disons la même chose dans différentes situations. Bien entendu, lorsque nous avons un excellent candidat à l'être de A — à savoir Corps-de-B — nous l'appelons A ; mais cela ne veut pas dire que nous ne devions pas appeler Corps-de-A A, dans cette autre situation, lorsqu'il n'y a personne de mieux à notre disposition. A situations différentes, descriptions différentes ». Cette façon de raisonner convient assurément à des avocats qui ont à décider à qui appartient une propriété qui a subi une série déroutante de transformations ; ils ont seulement à décider — et dans chaque situation, la propriété, supposerons-nous, doit aller à quelqu'un — en se fondant sur ce qui est compatible avec la loi et les faits. Mais quand il s'agit des peurs ou des attentes qu'un individu peut avoir au sujet de son propre futur, employer cette méthode paraît n'avoir plus aucun sens. Si les craintes de A s'étendent à ce qui arrivera à Corps-de-A en 5, je ne vois pas comment elles pourraient se désintéresser rationnellement du sort de la personne tout à fait semblable en 6, en se laissant dire que dans la seconde situation quelqu'un aurait, pour décider d'appeler une autre personne A, une raison qu'il n'aurait pas dans la première.

Ainsi, pour résumer, tout se passe comme s'il y avait deux façons de présenter l'expérience imaginée et le choix qui lui est associé. Chacune emporte la conviction, et chacune conduit à des conclusions opposées. L'idée, qui plus est, que la situation après l'expérience est conceptuellement indécidable sur le point en question ne semble pas réduire la perplexité mais plutôt l'accroître ; cependant que l'idée (si souvent invoquée dans ces domaines) que c'est décidable par convention est encore pire. En raison de quoi, je ne sais pas du tout quelle option il serait judicieux de choisir si on avait à choisir avant l'expérience. Et je trouve cela plutôt inquiétant. Mais, quelle que soit notre perplexité, une des argumentations qui y a abouti mérite d'avoir un traitement particulier parce qu'elle s'inscrit en faux contre quelque chose qui est, je pense, souvent vaguement admis. On reconnaît en effet souvent qu'il y a des aspects « première personne » et des aspects « troisième personne » dans les questions qui touchent les individus

et qu'il y a des problèmes liés aux relations entre ces aspects. On reconnaît aussi que les considérations liées au « mental » (comme on peut vaguement les désigner) et les considérations de continuité physique jouent un rôle dans les questions d'identité de la personne (ce qui ne veut pas dire qu'il y ait des critères mentaux et physiques d'identité de la personne). On serait tenté de penser qu'il y a corrélation entre les deux séries : en gros, qu'une approche de type « première personne » focalise l'attention sur des considérations portant sur le mental, tandis qu'une approche de type « troisième personne » insisterait sur les considérations de continuité physique. La discussion présente est exactement l'illustration du contraire. La première argumentation, qui a abouti à la conclusion « mentaliste » que A et B échangeraient leurs corps et que chaque personne devrait s'identifier au lieu de ses souvenirs et de son caractère, a été entièrement conduite dans une approche de type « troisième personne ». La deuxième argumentation, qui assimilait identité à continuité physique, s'intéressait au problème de « première personne » de ce que A pouvait attendre. Qu'il en soit ainsi me paraît (encore que je ne veuille pas poursuivre ici cette discussion) assez significatif.

Je terminerai en suggérant une manière assez précaire de se rapprocher d'une solution du problème, en utilisant seulement le peu de matériau désormais disponible.

Les arguments apparemment décisifs de la première présentation, qui indiquaient que A devait s'identifier à Corps-de-B, trouvaient leur force dans la parfaite adaptation d'une situation répondant à la description de « l'échange des corps ». Mais cette adéquation est fondamentalement artificielle ; c'est la résultante de la volonté que l'expérimentateur a de construire une situation qui produise naturellement, avec le minimum d'hésitation, cette description. Etant donné la méthode employée, il aurait facilement pu s'arrêter plus tôt ou continuer plus loin. Il aurait pu s'arrêter à 5 en laissant B tel qu'il était ; ou il aurait pu continuer et feindre deux personnes ayant chacune le caractère et les souvenirs de A, ou tout aussi bien deux personnes ayant les idiosyncrasies de B. S'il avait opté pour l'une ou l'autre de ces dernières hypothèses, nous aurions encore moins su quoi dire ;

l'expérimentateur a précisément choisi de nous rendre aussi facile que possible de trouver quelque chose à dire. Sans doute, si nous avions quelque modèle de personnes qui seraient comme des fantômes incarnés dans des corps et, qui fussent d'une certaine façon véritablement échangés grâce à certains protocoles, on pourrait considérer cette expérimentation parfaite comme étant l'expérimentation efficace ; la seule méthode qui aurait comme résultat que les fantômes fussent échangés sans être détruits, ni dispersés ni quoi que ce soit d'autre. Mais on ne saurait sérieusement utiliser un tel modèle. L'expérimentateur n'a pas, comme dans le cas de ces modèles, *provoqué* un échange de corps ; il a plutôt choisi la situation parmi un ensemble de situations également possibles que nous serions tout à fait prêts à qualifier d'échanges de corps. Comparé à cela, le principe d'après lequel les craintes qu'on a peuvent s'étendre à la souffrance à venir, quels que soient les changements psychologiques qui pourraient survenir dans l'intervalle, semble tout à fait simple. Peut-être ne l'est-il pas vraiment ; mais nous demandons à voir ce qu'il aurait d'erroné. Tant qu'on ne nous aura pas montré ce qu'il aurait d'erroné, nous devrions peut-être décider que si nous étions A, et que nous dussions prendre une décision égoïste, nous devrions laisser la souffrance à l'individu Corps-de-B. Ce serait risqué ; mais qu'il y ait place ici pour la notion de *risque* est en soi une des caractéristiques majeures du problème.

Le cas Makropoulos et l'ennui qui s'attache à l'immortalité

Cet essai est issu d'une conférence prononcée¹ dans le cadre d'un cycle consacré « à l'immortalité de l'âme ou à un thème spirituel voisin ». Mon thème spirituel voisin est, si l'on veut, la mortalité de l'âme. Parmi les conférenciers qui m'avaient précédé ceux qui étaient des philosophes cherchaient, je pense, à déterminer si, oui ou non, nous étions immortels ; tel n'est pas mon propos, mais plutôt de montrer quelle chance nous avons de ne pas l'être. J'avancerai que l'immortalité, ou une condition ignorant la mort, serait dépourvue de sens ; qu'ainsi, d'une certaine façon, la mort donne du sens à la vie. Cela ne veut pas dire que nous devrions ne pas craindre la mort (quelle que soit l'efficacité qu'on puisse, par ailleurs, accorder à cette injonction). Effectivement, il y a plusieurs manières de concevoir que, dans le même temps, la mort donne son sens à la vie et qu'elle soit, abstraction faite du reste, quelque chose qui doit être redouté. Par exemple, certains existentialistes semblent avoir affirmé que la mort était ce qui donnait du sens à la vie, à supposer que quelque chose le pût, précisément parce que c'était la peur de la mort qui donnait du sens à la vie ; je ne les suivrai pas sur ce terrain. Je veux plutôt explorer l'idée que, à considérer les réalités du désir et du bonheur humains, et

1. A Berkeley, Université de Californie, sous le patronage de la fondation Agnès et Constantin Foerster. Je remercie le comité exécutif de m'avoir invité à prononcer la conférence de l'année 1972.

d'une discussion d'expériences de pensée sur la division de l'esprit dans sa base physique (scission d'un cerveau par exemple), envisage de manière critique un examen analytique de ce qu'est une personne. Il examine deux théories de la personne : la théorie de l'Ego et celle du Faisceau, théorie qui nie qu'il existe un supposé personnel et réduit le moi à un agrégat. Parfit voit dans la théorie du non Soi du Bouddha un échantillon de théorie du faisceau. Parfit soutient que le cas des cerveaux divisés favorise la théorie du faisceau : qu'il y ait un cerveau divisé n'implique pas que deux personnes occupent mon corps. L'unification ne suppose pas l'existence d'un moi mais une théorie psychologique que défend Parfit. À l'opposé R. Chisholm pose le problème métaphysique de la liberté humaine à l'intérieur d'une théorie qui ne se confond ni avec celle de l'Ego, ni avec celle du Faisceau, et qui reprend certaines intuitions aristotéliennes sur la base d'une analyse de l'action responsable. Chisholm reprend le contexte classique de discussion, celui du déterminisme et de l'indéterminisme. Le concept qu'il analyse de la manière la plus minutieuse est celui de la cause. Jérôme Dokic aborde, dans un texte inédit, la question de la perception, en lien avec la métaphysique de la manière suivante : est-ce que l'analyse de la perception peut donner des arguments en faveur de telle ou telle thèse métaphysique ? Il soutient la thèse de la neutralité : cette analyse est neutre quant aux grandes options métaphysiques. Cette neutralité ne signifie pas que l'analyse cognitive de la perception soit parfaitement sans intérêt pour la métaphysique. Si cette analyse ne commande pas le choix de telle ou telle option métaphysique, elle peut dissuader d'en retenir certaines et en ce sens elle peut avoir un rôle critique et évidemment de clarification.

DEREK PARFIT

LES ESPRITS DIVISÉS ET LA NATURE DES PERSONNES*

Ce sont les cas de cerveaux divisés qui m'ont conduit à la philosophie. Ce que nous en savons repose sur le résultat de plusieurs tests psychologiques, décrits par Donald McKay¹. Ces tests exploitent deux faits. Nous contrôlons chacun de nos bras, et voyons ce qu'il y a dans chaque moitié de notre champ visuel, avec un seul de nos hémisphères cérébraux. Après que les hémisphères d'une personne ont été déconnectés, les psychologues peuvent lui présenter deux questions écrites dans chaque moitié de son champ visuel, et recevoir deux réponses différentes écrites avec l'une ou l'autre main.

Voici une version imaginaire simplifiée de ce que ces tests mettent en évidence. Une personne dont le cerveau a été divisé regarde fixement le centre d'un grand écran, dont la moitié gauche est rouge et la moitié droite bleue. Si les mots « Combien de couleurs voyez-vous ? » apparaissent en lettres plus foncées à gauche et à droite, la personne écrira « Une seule » avec une main et la même chose avec l'autre. Si à présent les mots « Quelle est

* Le texte original est intitulé « Divided Minds and the Nature of Persons » et fait partie du recueil édité par S. Blakemore et S. Greenfield, *Mindwaves*, Oxford, Oxford UP, 1987, p. 19-26.

1. Cf. la contribution de MacKay au volume édité par Blackemore et Greenfield (chap. 1).

l'unique couleur que vous voyez ? » apparaissent, la personne écrira « Bleu » avec une main et « Rouge » avec l'autre.

Si c'est ainsi que la personne répond, je conclurais qu'elle a deux sensations visuelles – qu'elle voit, comme elle le prétend, à la fois du rouge et du bleu. Mais en voyant une couleur, elle n'a pas conscience de voir l'autre. Elle a deux courants de conscience, dans chacun desquels elle ne voit qu'une seule couleur. Elle voit du rouge dans un courant et, au même moment, du bleu dans l'autre. Plus généralement, elle pourrait avoir au même moment deux séries de pensées et de sensations, et en ayant l'une ne pas avoir conscience d'avoir l'autre.

Cette conclusion a été contestée. Certains ont nié qu'il y ait là deux courants de conscience, au motif que l'hémisphère sous-dominant est une partie du cerveau et que son fonctionnement n'implique aucune conscience. S'ils ont raison, les cas de cerveaux divisés perdent beaucoup de leur intérêt. Je crois qu'ils ont tort, principalement parce que si l'hémisphère dominant d'une personne est détruit, celle-ci est capable de réagir comme l'hémisphère sous-dominant dans les cas de cerveaux divisés, et nous ne la considérons comme un automate dépourvu de conscience. Certes, l'hémisphère sous-dominant est bien moins développé à plusieurs égards, il possède typiquement les capacités linguistiques d'un enfant de trois ans. Mais les enfants de trois ans sont conscients, ce qui confirme la thèse selon laquelle, dans les cas de cerveaux divisés, il y a deux courants de conscience.

Selon une autre thèse, de tels cas impliquent deux personnes partageant le même corps. Comme le Professeur McKay, je pense qu'il faut rejeter cette thèse. Ma justification est toutefois différente. Le Professeur McKay nie qu'il y ait deux personnes impliquées parce qu'il croit qu'il n'y en a qu'une. Pour ma part, je crois qu'en un sens, il n'y en a aucune.

La Théorie de l'Ego et la Théorie du Faisceau

Pour expliquer le sens en question, je dois me détourner un instant des cas de cerveaux divisés. Deux théories s'opposent sur ce qu'est une personne, et ce que constitue l'existence continue d'une personne à travers le temps. Selon la Théorie de l'Ego, l'existence continue d'une personne ne peut être expliquée que comme l'existence continue d'un *Ego* ou *sujet d'expérience* particulier. Les partisans de la Théorie de l'Ego prétendent que la réponse à la question de savoir ce qui unifie la conscience d'une personne à un moment donné – ce qui fait que, par exemple, je peux en ce moment à la fois voir ce que je suis en train de taper à l'écran et entendre le vent à travers ma fenêtre – est que ces deux expériences sont vécues par moi, cette personne, à ce moment. De même, ce qui explique l'unité de la vie entière d'une personne est le fait que la même personne, ou sujet d'expérience, a toutes les expériences dans cette vie. Dans sa forme la plus connue, la *Conception cartésienne*, chaque personne est une entité persistante purement mentale – une âme ou une substance spirituelle.

La conception rivale est la *Théorie du Faisceau*. Comme la plupart des styles artistiques – gothique, baroque, rococo, etc. –, cette théorie doit son nom à ceux qui la critiquent. Mais ce nom n'est pas si mauvais. Selon la Théorie du Faisceau, ni l'unité de la conscience à un moment donné, ni l'unité d'une vie entière ne peut être expliquée en faisant référence à une personne. Il faut plutôt parler de longues séries d'états et d'événements mentaux différents – des pensées, des sensations, et ainsi de suite –, chaque série constituant ce que nous appelons une vie. Chaque série est unifiée par différents types de relations causales, telles que les relations qui lient des expériences et les souvenirs ultérieurs les concernant. Chaque série est donc comme un faisceau ou un fagot attaché par une ficelle.

En un sens, les partisans de la Théorie du Faisceau nient l'existence des personnes. Une dénégation catégorique serait certes absurde. Comme Reid l'écrivait au XVIII^e siècle en guise de protes-

tation, « Je ne suis pas une pensée, je ne suis pas une action, je ne suis pas un sentiment ; je suis une entité qui pense, agit, et ressent ». Je ne suis pas une série d'événements, mais une personne. Les défenseurs de la Théorie du Faisceau l'admettent, mais prétendent qu'il s'agit là d'un fait de notre grammaire, ou de notre langue. Il y a des personnes et des sujets en ce sens linguistique. Si les personnes sont conçues d'une manière qui transcende le sens linguistique – comme des entités existant séparément, distinctes de nos cerveaux et de nos corps, et des différents types d'états et d'événements mentaux – la Théorie du Faisceau nie leur existence.

Le premier défenseur de la Théorie du Faisceau est Bouddha, qui a enseigné l'*anatta*, ou la *Conception du Non-Soi*. Les bouddhistes admettent que le soi ou la personne a une « existence nominale », au sens où elle n'est qu'une combinaison d'autres éléments. Seul ce qui existe par lui-même, comme un élément séparé, possède ce que les bouddhistes appellent une « existence réelle ». Voici quelques citations de textes bouddhistes :

Au début de leur conversation, le roi demande poliment au moins son nom, et reçoit la réponse suivante : « Monsieur, mon nom est "Nagasena" ; c'est ainsi que mes compagnons dans la vie religieuse s'adressent à moi. Bien que mes parents m'aient donné ce nom... celui-ci n'est qu'une simple appellation, une forme de discours, une description, un usage conventionnel. "Nagasena" n'est qu'un nom, car vous ne trouverez nulle personne ici ».

Tu crois qu'un être sensible existe, Ô Mara ? Tu es leurré par une fausse conception. Ce faisceau d'éléments est vide de Soi, il ne contient aucun être sensible. De même qu'un ensemble de morceaux de bois reçoit le nom de chariot, nous donnons aux éléments le nom d'êtres imaginaires.

Bouddha a parlé ainsi : « Ô Frère, les actions existent, de même que leurs conséquences, mais la personne qui agit n'existe pas. Il n'y a personne pour rejeter cet ensemble d'éléments, et personne pour

assumer un nouvel ensemble. Il n'y a pas d'Individu, ce n'est qu'un nom conventionnel donné à un ensemble d'éléments »¹.

Les déclarations de Bouddha ressemblent de manière frappante à celles de plusieurs auteurs occidentaux. Puisque ceux-ci ne connaissent rien à Bouddha, cette ressemblance suggère que de telles déclarations ne font pas seulement partie d'une tradition culturelle, liée à une période historique. Elles pourraient bien être vraies, comme je le pense.

Ce que nous croyons être

Étant donné les progrès de la psychologie et de la neurophysiologie, la Théorie du Faisceau peut paraître aujourd'hui trivialement vraie. Il peut sembler banal de nier l'existence d'Ego existant séparément, distincts de nos cerveaux, corps, états et événements mentaux de différents types. Mais là n'est pas l'unique enjeu. Il se peut que nous soyons convaincus que la Théorie de l'Ego est fautive, voire même dénuée de sens. Pourtant, la plupart d'entre nous ont aussi certaines croyances, dont nous ne sommes pas toujours conscients, au sujet de ce qui constitue notre existence continue à travers le temps. Or ces croyances ne pourraient être justifiées que par une théorie comme celle de l'Ego. Par conséquent, la plupart d'entre nous avons des croyances fausses sur ce qu'est une personne, et sur nous-mêmes.

Ces croyances se manifestent le plus clairement quand nous considérons certaines situations imaginaires, souvent empruntées à la science-fiction. La *télétransportation* est une situation de ce genre. Supposons que vous entriez dans une cabine dans laquelle, quand vous appuyez sur un bouton, un scanner enregistre l'état de toutes les cellules de votre cerveau et de votre corps, en détruisant ces derniers par la même occasion. Cette information est ensuite transmise à la vitesse de la lumière à quelque autre planète, sur

1. Pour les sources de ces citations et d'autres citations similaires, cf. mon *Reasons and Persons*, Oxford, Oxford UP, 1984, p. 502-503 et 532.

laquelle un duplicateur produit une copie organique parfaite de vous-même. Puisque le cerveau de votre Réplique est exactement similaire au vôtre, il aura l'impression de se souvenir d'avoir vécu votre vie jusqu'au moment où vous avez appuyé sur le bouton; sa personnalité sera exactement comme la vôtre, et il sera sous chaque autre aspect en continuité psychologique avec vous. Cette continuité psychologique n'aura pas la cause qu'elle a normalement, à savoir l'existence continue de votre cerveau, puisque la chaîne causale aura transité par la transmission radio de votre « empreinte ».

Plusieurs auteurs prétendent que si vous choisissez d'être télétransportés, parce que vous pensez que c'est le mode de transport le plus rapide, vous faites une terrible erreur. Ce n'est pas un mode de transport, mais une façon de mourir. Ils admettent qu'une telle mort pourrait ne pas être tout à fait aussi terrible que la mort ordinaire. Vous pouvez vous consoler à la perspective qu'après votre mort, votre Réplique pourra finir le livre que vous êtes en train d'écrire, jouer le rôle de parent pour vos enfants, et ainsi de suite. Mais ils insistent sur le fait que la Réplique n'est pas vous. Elle est seulement quelqu'un d'autre qui vous ressemble en tous points. C'est pourquoi cette perspective est presque aussi terrible que la mort ordinaire.

Imaginons à présent toute une série de cas dans lesquels une proportion croissante de cellules de votre cerveau et de votre corps sont remplacées, en une seule opération, par des répliques exactes. Vers le début de la série, seules un ou deux pour cent de cellules sont remplacées. Au milieu de la série, 40 ou 60 pour cent, et à la fin de la série, 98 ou 99 pour cent. À cet extrême se place la télétransportation pure, où toutes vos cellules sont « remplacées ».

Quand vous imaginez qu'une proportion de vos cellules va être remplacée par des répliques exactes, il est naturel d'avoir les croyances suivantes. Premièrement, il doit y avoir une réponse aux questions « Vais-je survivre? Serai-je la personne résultant de ce processus? ». Ou bien vous allez survivre, ou bien vous êtes sur le point de mourir. Deuxièmement, la réponse à ces questions doit être

ou bien simplement « oui » ou bien simplement « non ». Ou bien la personne qui se réveillera sera vous, ou bien elle sera quelqu'un d'autre. Il ne saurait y avoir de troisième réponse, comme si la personne qui se réveillera pouvait être à moitié vous. Vous pouvez vous imaginer vous-même ultérieurement à demi conscient. Mais si la personne résultante est totalement consciente, elle ne peut pas être à moitié vous. Pour résumer ces croyances, à la question « Serai-je la personne qui résultera de ce processus? », il *doit* y avoir une réponse, de l'ordre du tout ou rien.

Il y a de bonnes raisons de penser que dans le cas de la télétransportation, votre Réplique n'est pas vous. Selon une variante un peu différente de ce cas, votre Réplique est créée alors que vous êtes toujours en vie, de telle sorte que vous pouvez lui parler. Cela semble montrer que si 100 pour cent de vos cellules sont remplacées, le résultat est seulement une Réplique de vous. À l'autre extrémité de ma série, où seul un pour cent de cellules sont remplacées, la personne résultante *est* clairement vous. Il apparaît donc que dans les cas intermédiaires, la personne résultante doit être ou bien vous ou bien une simple Réplique. Il nous semble que l'un ou l'autre doit être vrai, et que la différence est capitale.

Pourquoi nous ne sommes pas ce que nous croyons

Si ces croyances étaient correctes, il devrait y avoir un pourcentage critique, quelque part dans la série des cas, en deçà duquel la personne résultante est vous, et au-delà duquel elle n'est que votre Réplique. Par exemple, peut-être allez-vous vous réveiller si la proportion de cellules remplacées est de 49 pour cent, alors que quelques cellules supplémentaires également remplacées auraient fait toute la différence, avec pour conséquence quelqu'un d'autre se réveillant à votre place.

Nos croyances naturelles impliquent qu'il doit y avoir un pourcentage critique de ce genre. Mais cette conclusion est fort peu plausible. Comment quelques cellules pourraient-elles faire une telle différence? De plus, si ce pourcentage critique existe,

personne ne pourra jamais découvrir d'où il vient. Puisque, dans tous les cas, la personne résultante croira qu'elle est vous, jamais aucune donnée probante ne pourra déterminer le moment où, dans la série des cas, elle cessera soudainement d'être vous.

Selon la Théorie du Faisceau, nous devrions renoncer à nos croyances naturelles. Puisque vous, en tant que personne, n'avez pas d'existence séparée, nous sommes en mesure de savoir exactement ce qui se passe dans les cas pertinents sans répondre à la question de savoir ce qui vous arrive. De plus, dans les cas vers le milieu de la série, la question de savoir si la personne résultante est vous ou seulement quelqu'un d'autre vous ressemblant en tous points est vide. Il ne s'agit pas là de deux options différentes, dont une seulement doit être vraie. Il s'agit seulement de deux descriptions différentes des mêmes événements. Si 50 pour cent de vos cellules sont remplacées par des doubles parfaits, nous pouvons appeler « vous » la personne résultante, ou nous pouvons l'appeler seulement « votre Réplique ». Il s'agit là d'un choix purement terminologique.

Comme Bouddha le disait, la Théorie du Faisceau est difficile à croire. Il est difficile d'accepter que la question de savoir si l'on est sur le point de mourir, ou si au contraire l'on va vivre encore plusieurs années, pourrait être vide.

Ce que l'on nous demande d'accepter peut être clarifié au moyen d'une analogie. Supposons qu'un certain club existe depuis quelque temps, et qu'il tienne régulièrement des réunions. Puis les réunions cessent. Quelques années plus tard, plusieurs personnes forment un club portant le même nom, et avec le même règlement. Nous pouvons nous demander : « Ces personnes ont-elles fait revivre le même club, ou ont-elles seulement inauguré un autre club exactement similaire au précédent ? ». Étant donnés certains autres détails, la question est à nouveau vide. Nous sommes en mesure de savoir exactement ce qui s'est passé sans y répondre. Supposons que quelqu'un dise « Mais il doit y avoir une réponse. Le club formé ultérieurement doit être ou non le même club que l'original ». La

personne montrerait ainsi qu'elle n'a pas compris la nature des clubs.

De même, quiconque a le moindre doute sur les cas que j'ai imaginés ne comprend pas la nature des personnes. Dans chaque cas, nous savons que la personne résultante est exactement comme vous tant sous l'aspect psychologique que sous l'aspect physique, et qu'elle a une certaine proportion des cellules qui étaient dans votre cerveau et dans votre corps – 90 pour cent, 10 pour cent ou, dans le cas de la télétransportation, zéro pour cent. En sachant cela, nous savons tout ce qu'il y a à savoir. Comment donner un sens à la question de savoir ce qui vous arriverait à moins de supposer que vous êtes un Ego existant séparément, distinct d'un cerveau et d'un corps, et d'états et d'événements mentaux de différents types ? Si de tels Ego n'existent pas, la question est sans objet.

Non seulement la Théorie du Faisceau est difficile à accepter, mais elle peut aussi affecter nos émotions. Comme Bouddha le disait, elle peut diminuer nos préoccupations sur notre avenir. Un tel effet peut être induit en re-décrivant ce changement de perspective. Supposons que vous soyez sur le point d'être détruit, mais que vous aurez plus tard une Réplique sur Mars. Vous pensez naturellement que cette perspective est aussi sombre qu'une mort ordinaire, puisque votre Réplique ne sera pas vous. Selon la Théorie du Faisceau, le fait que votre Réplique ne sera pas vous consiste seulement dans le fait que bien que votre Réplique soit en parfaite continuité psychologique avec vous, la cause de cette continuité n'est pas normale. Mais lorsque vous objectez à la télétransportation, vous n'objectez pas seulement au caractère anormal de cette cause. Votre objection est que cette cause ne *vous* conduira pas sur Mars. Vous craignez que la cause anormale échoue à produire un fait additionnel de la plus haute importance, différent du fait que votre Réplique est en continuité psychologique avec vous. Vous ne voulez pas seulement qu'il y ait une continuité psychologique entre vous et quelque personne future. Vous voulez *être* cette personne future. Selon la Théorie du Faisceau, un tel fait additionnel n'existe pas. Ce que vous craignez n'arrivera pas et, dans le cas imaginé,

n'arrive *jamais*. Vous voulez que la personne sur Mars soit vous en un sens particulièrement intime, mais aucune personne future ne sera jamais vous en ce sens. Il s'ensuit que du point de vue de vos croyances naturelles, même la survie ordinaire est presque aussi terrible que la télétransportation. *La survie ordinaire est à peu près aussi terrible que le fait d'être détruit et d'avoir une Réplique.*

Comment les cas de cerveaux divisés favorisent la Théorie du Faisceau

La vérité de la Théorie du Faisceau me semble être, dans le sens le plus large, une conclusion aussi bien scientifique que philosophique. Je peux imaginer le type de données qui auraient pu justifier la croyance qu'il existe des Ego existant séparément, et que l'existence continue de ces Ego explique la continuité de notre vie mentale. Mais en fait, peu de choses militent en faveur de la Théorie de l'Ego, alors qu'une pléthore de données conforte la Théorie rivale du Faisceau.

Ces données sont fournies entre autres par les cas de cerveaux divisés. Selon la Théorie de l'Ego, l'explication de ce qui unifie nos expériences à un moment donné repose sur la thèse selon laquelle toutes ces expériences sont vécues par la même personne. Les partisans de la Théorie du Faisceau rejettent cette explication. Le désaccord qui s'ensuit est difficile à régler dans les cas ordinaires. Mais considérons le cas simplifié de cerveau divisé que j'ai décrit. Nous montrons au patient que j'ai imaginé un panneau dont la moitié gauche est bleue et la moitié droite rouge. Dans l'un de ses courants de conscience, la personne est consciente de ne voir que du bleu, alors qu'au même moment, dans l'autre courant de conscience, elle est consciente de ne voir que du rouge. Chacune de ces deux expériences visuelles se combine avec d'autres expériences, comme celle d'être conscient de bouger une main ou l'autre. Qu'est-ce qui unifie les expériences que la personne a à un moment donné dans chacun de ses deux courants de conscience? Qu'est-ce qui unifie sa conscience de ne voir que du rouge et sa conscience de bouger une

main? La réponse ne peut pas être que ces expériences sont vécues par la même personne. Cette réponse ne peut pas expliquer l'unité de chaque courant de conscience de la personne, car elle néglige l'absence d'unité entre ces courants. La personne a présentement toutes les expériences résidant dans ses deux courants de conscience. Si c'était ce qui unifie ces expériences, les deux courants n'en feraient qu'un.

Les cas de cerveaux divisés n'impliquent pas, comme je l'ai affirmé, deux personnes partageant un seul corps. Puisqu'il n'y a qu'une personne en jeu, qui a les deux courants de conscience, l'explication du partisan de la Théorie de l'Ego devrait prendre la forme suivante. Il serait obligé de distinguer entre des personnes et des sujets d'expérience, et prétendre que les cas de cerveaux divisés impliquent *deux* sujets d'expérience. Ce qui unifie les expériences dans l'un des deux courants de conscience de la personne devrait être le fait que ces expériences sont toutes vécues par le même sujet d'expérience. Ce qui unifie les expériences dans l'autre courant de conscience de la personne devrait être le fait qu'elles sont vécues par un autre sujet d'expérience. Quand l'explication prend cette forme, elle devient beaucoup moins crédible. Tant que l'on pouvait supposer que le terme «sujet d'expérience», ou «Ego», veut simplement dire la même chose que «personne», il était facile de croire en l'existence de sujets d'expérience. Mais s'il peut y avoir des sujets d'expérience qui ne sont pas des personnes, et si dans la vie d'un patient dont le cerveau est divisé, il y a au même moment deux sujets d'expérience différents – deux Ego différents – pourquoi devrions-nous croire à l'existence de telles entités? Ces remarques ne constituent pas une réfutation. Mais elles me semblent fournir un argument solide contre la Théorie de l'Ego.

En tant que partisan de la Théorie du Faisceau, je crois que ces deux Ego sont des rouages qui tournent à vide. Il y a une autre explication de l'unité de la conscience, tant dans les cas ordinaires que dans ceux de cerveaux divisés. C'est simplement un fait que les gens ordinaires sont, à un moment donné, conscients d'avoir plusieurs expériences différentes. Cette conscience de plusieurs

expériences différentes peut être utilement comparée avec la conscience, dans la mémoire à court terme, de plusieurs expériences différentes. De même qu'il peut y avoir un seul souvenir d'avoir vécu juste auparavant plusieurs expériences, comme celle d'avoir entendu une cloche sonner trois fois, il peut y avoir un seul état de conscience à la fois d'entendre la cloche sonner une quatrième fois et de voir, au même moment, des corbeaux s'envolant du clocher.

Cette explication, contrairement à celle du partisan de la Théorie de l'Ego, peut facilement être étendue aux cas de cerveaux divisés. Dans ces derniers cas, il n'y pas un mais deux états de conscience de plusieurs expériences différentes à un moment donné. Dans l'exemple que j'ai décrit, il y a un état de conscience à la fois de ne voir que du bleu et de bouger une main, et il y a un autre état de conscience à la fois de ne voir que du rouge et de bouger l'autre main. En affirmant qu'il y a deux états de conscience de ce type, nous ne postulons pas d'entités étranges, comme deux Ego existant séparément et qui seraient distincts de la personne en question. Cette explication fait appel à deux états mentaux auxquels il faudrait de toute façon faire référence dans une description complète de l'exemple.

J'ai essayé de montrer comment les cas de cerveaux divisés pouvaient fournir un argument en faveur d'une certaine conception de la nature des personnes. Il me faut mentionner un autre argument de ce type, fourni par une extension imaginaire de ces cas, dont Wiggins a été le premier à fournir une discussion détaillée¹.

Dans ce nouveau scénario, le cerveau d'une personne est divisé, et les deux moitiés sont transplantées dans deux corps différents. Les deux personnes résultantes vivent leur vie chacune de leur côté. Un tel scénario montre que l'identité personnelle n'est pas ce qui compte. Si j'étais sur le point de me diviser, je devrais conclure qu'aucune des personnes résultantes ne sera moi. J'aurai cessé

1. À la fin de son ouvrage *Identity and Spatio-temporal Continuity*, Oxford: Blackwell, 1967.

d'exister. Mais cette façon de cesser d'exister est à peu près aussi réjouissante – ou aussi terrible – que la survie ordinaire.

Certains traits du scénario imaginaire de Wiggins resteront probablement techniquement impossibles. Mais le scénario ne peut pas être écarté, car son trait le plus frappant, la division d'un courant de conscience en des courants séparés, s'est déjà réalisé. C'est le deuxième aspect sous lequel les cas réels de cerveaux divisés ont une importance théorique capitale. Ils mettent au défi certaines des hypothèses les plus profondes que nous entretenons sur nous-mêmes¹.

Traduction Jérôme DOKIC

1. Je discute ces hypothèses plus avant dans la partie 3 de mon ouvrage *Reasons and Persons*.

R. Swinburne, *Is there a God?*, ch. 5 (traduction Paul Clavier)

Les propriétés et les événements peuvent être physiques ou mentaux. Par événement physique, j'entendrai un événement pour lequel aucune personne n'est forcément mieux placé que les autres pour savoir qu'il s'est produit. Les événements physiques sont extérieurs; il n'y a pas d'accès privilégié à ces événements. Ainsi, le fait que mon bureau soit rectangulaire est un événement physique puisque, même si j'étais la seule personne à l'avoir observé, n'importe qui pourrait le vérifier aussi bien que moi. Au nombre des événements physiques, il y a les événements cérébraux. Que tel neurone se décharge dans le cerveau à tel moment, c'est quelque chose qui pourrait être observé indifféremment par de nombreux observateurs distincts : la décharge est donc un événement physique. Au contraire, les événements mentaux sont ceux qu'une seule et unique personne explore par une voie qui lui est propre — en les éprouvant directement. La personne a un accès privilégié à ses événements mentaux, un moyen d'exploration qu'elle ne peut partager avec personne d'autre.

Il est clair — c'est même plus clair que tout le reste — qu'il existe réellement des événements mentaux, comme notre expérience nous l'apprend. Ils comprennent les zones de couleur dans le champ visuel, les peines, les émotions, les estimations, les pensées et les impressions. Ils comprennent également les buts que nous essayons d'atteindre au moyen de notre corps ou autrement (comme je l'ai dit au chapitre 2). Considérons les événements suivants : le fait que j'aie eu du chagrin hier midi, ou une image rouge dans mon champ visuel, ou que je pense au déjeuner, ou que j'aie l'intention d'aller à Londres. Si quelqu'un d'autre que moi avait un moyen de les explorer, je pourrais moi aussi les explorer par le même moyen. Un autre que moi peut faire des découvertes sur mon chagrin et mes pensées par un examen de mon comportement, et peut-être également de mon cerveau. Mais moi aussi, je pourrais étudier mon comportement : je pourrais visionner un film de moi; je pourrais étudier mon cerveau — via un système de miroirs et de microscopes — aussi bien que n'importe qui d'autre. Et pourtant, j'ai certainement un moyen de connaître ces chagrins, ces pensées et le reste, un moyen qui échappe même au meilleur observateur de mon comportement ou de mon cerveau : c'est de les éprouver directement. C'est pourquoi ils doivent être distincts des événements cérébraux, ou de tout autre événement corporel. Un neurophysiologiste ne peut pas observer la qualité de la couleur dans mon champ visuel, ou l'âcreté de l'odeur du roast-beef que je respire. Un martien qui viendrait sur Terre, et capturerait un être humain, et inspecterait son cerveau pourrait

découvrir tout ce qui se passe dans ce cerveau mais pourrait continuer à se demander "Est-ce que cet humain ressent vraiment quelque chose quand j'écrase son orteil?". Qu'il y ait des chagrins et des souvenirs, des pensées et des intentions, c'est un fait plus profond, situé au-delà de la production des événements cérébraux. De tels événements doivent être distingués des comportements spécifiques auxquels, ils donnent naissance. Les gens ont des sensations qu'ils n'expriment pas — des chagrins qu'ils cachent ou des sensations en rêve qu'ils ne racontent à personne — et d'ailleurs, si une sensation donne lieu à un comportement, le sujet est conscient de la sensation comme d'un événement distinct du comportement auquel elle donne lieu.

Je voudrais insister sur cette définition du mental comme ce à quoi le sujet a un accès privilégié. Il y a beaucoup de propriétés que nous attribuons aux gens, que nous pourrions quelquefois appeler "mentales" alors qu'elles ne sont pas mentales en ce sens, et ne sont en fait que des propriétés du comportement extérieur. Quand nous disons que quelqu'un est généreux ou irritable, ou qu'il est une source d'information utile, il se peut que nous nous contentions de dire quelque chose sur son comportement en public, et rien sur la vie de pensée et d'impression sous-jacente à de tels comportements. Naturellement, on peut toujours décrire l'irritabilité comme une propriété mentale, mais ce n'est pas une propriété mentale au sens où je l'ai définie. Mon objectif est d'établir qu'il y a des états mentaux en un sens précis, qui sont distincts des événements cérébraux. En établissant ce point, je ne nie pas un seul instant que la plupart de mes événements mentaux soient causés par mes événements cérébraux. Un événement dans mon cerveau (lui-même causé par un événement dans ma dent) a provoqué mon mal de dents; et un autre événement dans mon cerveau (lui-même causé par le mouvement de la branche derrière la fenêtre) a fait que je pense que la branche a bougé. Mais l'essentiel, c'est que, tout comme l'inflammation du pétrole diffère de l'explosion qui s'ensuit, l'événement cérébral diffère du chagrin (ou de n'importe quel sentiment) qu'il provoque. Et, bien entendu, la causalité marche également dans l'autre sens : tel projet que j'ai formé entraîne (sans intention de ma part) les événements cérébraux qui à leur tour provoquent le mouvement de mes membres (que j'avais l'intention de faire).

Les humains, contrairement aux choses inanimées, sont doués de propriétés mentales; ils ont une vie mentale. Mais il y a plus dans l'humain que le simple fait d'avoir une vie mentale connectée à un corps. Cette vie mentale elle-même, comme je vais le soutenir, est l'état d'une substance immatérielle, l'âme, qui est connectée au corps. Que les humains soient constitués de deux substances en connexion — un corps et une âme — c'est l'opinion connue sous le nom de

dualisme des substances. L'alternative est l'affirmation que les humains ne sont que des corps (je serais donc la même chose que ce qu'on peut appeler rapidement mon corps). Dans ce cas, mes propriétés mentales, comme avoir du chagrin ou un souvenir, seront des propriétés de mon corps. Appelons cette opinion *monisme substantialiste* — l'idée qu'il n'existe que des substances d'un seul type, les substances matérielles. Si le monisme était correct, il n'y aurait rien d'autre dans l'histoire du monde que la succession des événements impliquant des substances matérielles, leur origine et leur destruction, leurs propriétés et leurs relations (physiques ou mentales). Mais, comme je voudrais l'indiquer maintenant, si vous aviez connaissance de tout cela, vous ignorerez toujours l'une des choses les plus importantes de toutes — à savoir si vous, et tout autre humain, continuez ou non à mener une vie consciente.

Permettez-moi d'illustrer ceci par une histoire de transplantations cérébrales. Le cerveau est constitué de deux hémisphères et d'une tige cérébrale. On a la preuve qu'un être humain peut survivre et se comporter comme un être conscient même si une grande partie d'un de ses hémisphères est détruite. Alors imaginez que mon cerveau (les hémisphères et la tige cérébrale) soit séparé en deux parties, et que chaque demi-cerveau soit extrait de mon crâne pour être transplanté dans le crâne vide d'un corps dont on vient de retirer le cerveau; et qu'on ajoute à chaque demi-cerveau toutes les parties (par exemple plus de tige cérébrale) prélevées sur un autre cerveau (par exemple celui de mon vrai jumeau) qui sont nécessaires pour que la transplantation prenne et pour qu'il y ait deux personnes vivantes, vivant des expériences conscientes. Bien sûr, je suis tout à fait conscient qu'une opération d'une telle précision n'est pas pour le moment réalisable, et ne le sera peut-être jamais pour des savants purement humains, avec des ressources purement humaines; mais je ne vois pas qu'on rencontre dans le cours d'une telle opération de difficulté théorique insurpassable. (En réalité c'est un euphémisme - je m'attends tout à fait à ce que cela soit réalisé un jour). Nous sommes donc autorisés à poser la question suivante : si cette opération était réalisée et que nous ayons deux personnes vivantes, vivant toutes deux des expériences conscientes, laquelle serait moi ? Probablement toutes les deux auraient, dans une certaine mesure, un comportement semblable au mien, affirmeraient être moi, se souvenir qu'elles ont fait ce que j'ai fait; puisque le comportement et la parole dépendent, pour une large part, des états cérébraux, et qu'il y a des recouvrements tout à fait considérables entre "l'information" transitant par les deux hémisphères et qui donne naissance au comportement et à la parole. Pourtant, aucune des deux personnes ne serait moi. Si elles étaient toutes deux identiques à moi, elles seraient la même personne (si a est identique à b , et b identique à c , alors a est identique à

c), or elles ne le sont pas. Elles ont maintenant des expériences distinctes et mènent des vies distinctes. Il reste trois autres possibilités : la personne qui a mon hémisphère droit est moi, la personne qui a mon hémisphère gauche est moi, ou aucune des deux n'est moi. Laquelle est la vraie ? Aucune certitude possible. Il s'ensuit que la simple connaissance de ce qui arrive aux cerveaux ou aux corps ou à tout ce qui est matériel ne nous dit pas ce qui arrive aux personnes.

On peut être tenté de dire que la question de savoir laquelle de ces trois possibilités est la bonne est une question arbitraire de définition. On doit résister à cette tentation. C'est un fait crucial qui est ici en jeu — on peut le montrer en imaginant que j'aie été capturé par un chirurgien fou qui s'apprête à réaliser sur moi l'opération de dédoublement cérébral. Il m'annonce (j'ai toutes les raisons de le croire) que la personne qui va être formée à partir de mon hémisphère cérébral gauche va mener une vie agréable, et que la personne qui va être formée à partir de mon hémisphère cérébral droit va être assujettie à une vie de torture. La question de savoir si ma vie future sera une vie heureuse ou très douloureuse, ou tout simplement si je survivrai à cette opération, est clairement une question de fait. (Seule une personne sous l'emprise d'un dogmatisme philosophique très dur songerait à le nier). Pourtant, alors que j'attends la transplantation et que je sais exactement ce qui va arriver à mon cerveau, je n'ai aucune possibilité de connaître la réponse à la question : qu'est-ce qui va m'arriver ? Peut-être qu'aucune des deux personnes futures ne sera moi — il se peut que la section de la tige cérébrale détruise la personne originale une fois pour toutes et que, bien qu'en réparant la tige débranchée on crée deux nouvelles personnes, aucune des deux ne soit moi. Peut-être serai-je la personne à l'hémisphère gauche, ou peut-être sera-ce la personne à l'hémisphère droit qui sera moi. Même si une des personnes après la transplantation ressemble, davantage que l'autre, au moi d'avant, par son caractère et ses prétendus souvenirs, il se peut qu'elle ne soit pas moi. Peut-être survivrai-je à l'opération mais changerai-je de caractère et perdrai-je beaucoup de mes souvenirs du fait de cette opération, en conséquence de quoi l'autre personne, après la transplantation, ressemblera plus dans son comportement public au moi d'avant que je ne le ferai.

Réfléchir sur cette expérience de pensée montre que, quelle que soit la quantité de connaissances dont on dispose sur ce qui est arrivé à mon cerveau ou à tous mes autres constituants matériels — on pourrait savoir exactement ce qui est arrivé à chacun de ses atomes — nous ne savons pas pour autant forcément ce qui m'est arrivé. Il s'ensuit qu'il doit y avoir en moi plus que la matière dont mon corps et mon cerveau sont faits, une part immatérielle plus essentielle encore, dont la continuation dans l'existence fait que le cerveau (et donc le corps) auquel je suis

connecté est mon cerveau (et mon corps), et à cette chose je donne le nom traditionnel d' "âme". Je suis mon âme plus tout ce que vous voudrez de cerveau (et de corps) auquel elle est connectée. En temps normal, mon âme marche de pair avec mon cerveau, mais dans des circonstances inhabituelles (comme quand mon cerveau est dédoublé) on ne sait pas où elle va.

Prenons un exemple légèrement différent. Je meurs d'une hémorragie cérébrale que les médecins d'aujourd'hui ne peuvent pas soigner, mais mes proches prennent mon corps et le mettent directement dans un super congélateur en Californie. Peu après se produit un tremblement de terre, avec comme conséquence que mon cerveau congelé éclate en de nombreux morceaux, dont on perd une partie. Pourtant, cinquante ans plus tard, comme les techniques médicales ont progressé, mes descendants décident de prendre les morceaux de mon corps éparpillé, le réchauffent, le raccommode, en remplaçant les pièces manquantes par d'autres prélevées ailleurs. Le corps devient le corps d'une personne vivante dont le comportement ressemble assez au mien et qui a l'air de se rappeler un bon bout de ma vie passée. Suis-je revenu à la vie, ou non ? Peut-être bien que oui, peut-être bien que non. On retrouve ici une question de vérité : ai-je oui ou non survécu à l'hémorragie comme je l'espérais ? Mais c'est une vérité dont nous ne pouvons être sûrs, quelle que soit la masse de connaissances que nous avons sur l'histoire de mon cerveau. Ma survie consiste donc dans la continuation de quelque chose d'autre, que j'appelle mon âme, qui est liée à mon corps précédent: et je pourrai survivre dans ce nouveau corps si et seulement si cette âme est connectée avec lui. Notez que cette vérité supplémentaire ne concerne pas le genre de vie mentale qui est connectée au cerveau. Ce n'est pas une vérité concernant les propriétés mentales, les pensées, les impressions et les intentions qu'a la personne rappelée à la vie. Mais plutôt, cette vérité supplémentaire, cette vérité sur la question de savoir si j'ai survécu, concerne la question QUI — c'est-à-dire quelle est la substance dans laquelle ces propriétés sont instanciées. Étant donné que la simple connaissance de ce qui est arrivé à toutes les substances matérielles ne me le dit pas, ce doit être une vérité concernant une substance immatérielle. Aussi longtemps que je continue à avoir des pensées, des impressions et des intentions, c'est que j'ai survécu à l'opération — quoi qu'il arrive à chacun de mes constituants matériels. Mon âme est donc mon constituant essentiel— c'est sa survie qui fait ma survie; pensées, impressions etc. m'appartiennent du fait qu'elles appartiennent à mon âme. L'âme est le constituant essentiel de la personne humaine.