

Chapitre VI Le Problème traditionnel

6.1 Au chapitre I, j'ai soutenu que l'on devrait considérer le problème du libre arbitre et du déterminisme comme deux problèmes : le Problème de la Compatibilité, et le Problème Traditionnel. Les chapitres III et IV ont été consacrés au Problème de la Compatibilité, et à ses ramifications. C'est-à-dire que dans ces deux chapitres, nous avons tenté de répondre à la question de savoir si le libre arbitre et le déterminisme sont compatibles. Nous avons conclu qu'ils sont en fait incompatibles. Cette conclusion a pour conséquence que le Problème Traditionnel existe, une proposition que tous ceux qui ont réfléchi avec le plus d'intensité sur le libre arbitre, du dix-septième siècle jusqu'à présent, ont mis de grands efforts à récuser. (Je ne suis pas heureux de m'engager dans une guerre dont les forces ennemies sont commandées par Hobbes, Hume et Mill. Mais je dois suivre l'argument là où il conduit.) Dans le présent chapitre, je ferai ce que je peux avec le Problème Traditionnel. Je ne sais pas comment le résoudre. Autrement dit, je ne sais pas comment montrer que nous avons le libre arbitre. Il y a, je crois, un bon argument pour l'existence du libre arbitre, et aucun bon argument contre lui, à moins qu'il n'y ait un bon argument pour le déterminisme. Je crois, en outre, qu'il n'y a pas de bons arguments pour le déterminisme et qu'il y a des arguments plutôt bons contre lui. Les deux phrases précédentes résument ce que je crois être la « situation dialectique » dans laquelle se trouve le philosophe qui se confronte au Problème Traditionnel. Le corps de ce chapitre sera une défense de la thèse que cette description de la situation dialectique d'un tel philosophe est une description exacte.

6.2 Revenons à la question du déterminisme. Demandons-nous quelles raisons il y a de penser que le déterminisme est vrai, et quelles raisons il y a de penser qu'il est faux. Le déterminisme, rappelons-nous, est la thèse que le passé et les lois de la nature déterminent ensemble, parmi les futurs possibles ou consistants de manière interne, un seul d'entre eux à être le futur, le futur actuel. (Rappelons-nous que le déterminisme doit être soigneusement distingué de la thèse de la causalité universelle, la thèse que tout événement – ou état de choses, ou comme vous voudrez – a une cause. Même si je me trompe en croyant que la causalité universelle n'implique pas le déterminisme¹, que cette implication soit valable est certainement une thèse philosophique substantielle. Et je doute que quiconque penserait que la causalité universelle est incompatible avec le libre arbitre, s'il ne pensait pas que la causalité universelle implique le déterminisme.) Il sera opportun de présenter notre enquête comme un commentaire de deux assertions : 'La science montre que le déterminisme est vrai' et 'la raison montre que le déterminisme est vrai'.

(i) *La science montre que le déterminisme est vrai.* Cela ne semble simplement pas être le cas. Si la science montre quelque chose, c'est tout à fait le contraire, car l'interprétation standard de la mécanique quantique est indéterministe. Cet énoncé brut demande un certain nombre de qualifications et de commentaires.

(a) On entend parfois dire que l'indétermination quantique ne s'applique qu'aux événements inobservables, sous-microscopiques et non aux événements observables de la vie ordinaire. C'est tout-à-fait faux : la mécanique quantique est une théorie *physique* – par opposition à une théorie biologique ou géologique – et on l'applique ainsi de manière légitime au comportement d'un éléphant ou d'un glacier comme à celui d'un neutrino ou d'un positron. Il est vrai que l'indétermination prédite par la théorie dans le comportement d'un éléphant ou d'un glacier est plus petite, selon plusieurs ordres de grandeur, que quoi que ce soit qui puisse être mesuré par les instruments les plus merveilleux dont notre science fiction

¹. Voir la section 1.3

puisse rêver. Néanmoins, elle est là.

Plus important encore est le fait que le domaine du sous-microscopique et le domaine de l'observable ne sont pas isolés causalement l'un de l'autre. C'est en un sens, un truisme, puisque le comportement de l'observable survient, on le présume, sur le comportement du sous-microscopique². Mais il y a un sens plus fort et plus intéressant dans lequel ces deux domaines ne sont pas isolés l'un de l'autre : des événements sous-microscopiques *isolés* peuvent « déclencher » des événements observables. Un moyen de produire effectivement un tel déclenchement est un compteur Geiger. Le passage d'une seule particule alpha ou d'un électron ou l'apparition d'une radiation électromagnétique de haute énergie dans un tube de gaz peut ioniser suffisamment le gaz dans le voisinage immédiat du passage de l'intrus pour produire un effet qui – après les amplifications adaptées, que le compteur Geiger est conçu pour produire – est enregistré comme un clic audible ; et toute particule ou éclat aura été produit par le changement du noyau ou des électrons orbitaux d'un seul atome. Ces faits peuvent, je crois, être résumés justement en disant qu'un événement sous-microscopique est la cause du clic produit par le compteur Geiger. Etant donnée cette caractéristique des compteurs Geiger, il serait assez facile de construire des instruments dont le comportement soit indéterminé (pourvu, bien sûr, que la théorie quantique soit effectivement correcte). Nous pourrions, par exemple construire une bombe à retardement incorporant une source radioactive, un compteur Geiger, et un mécanisme de mise à feu conçu pour prendre les « sorties » du compteur Geiger – les clics ou autres – comme « entrées ». Une telle bombe pourrait être conçue pour exploser si le compteur cliquait cinq fois, par exemple, pendant un intervalle de dix secondes. Supposons que nous devions construire une telle bombe et l'armer ; supposons que l'obligeant Dieu des philosophes nous en donne un millier de doubles parfaits – les doubles les plus précis que la toute-puissance puisse former, cette duplication étant, bien entendu, comprise comme incluant le dispositif du mécanisme d'armement. Qu'arriverait-il ? Ce qui n'arriverait *pas*, si ce n'est avec la chance la plus infime, c'est l'explosion simultanée de toutes les bombes. Ou du moins, cela n'arriverait pas sans que la base de la théorie physique en cours ne soit une illusion épouvantable. Dire cela ce n'est pas fixer des limites à la toute-puissance, à moins que l'on n'estime lui fixer des limites en disant que même un être tout-puissant ne pourrait pas faire un aéroplane fonctionnel entièrement fait de fromage et de crème fouettée. Car, tout comme le fromage et la crème fouettée sont des matériaux intrinsèquement inaptes aux besoins de l'ingénierie aéronautique, *tout* matériel physiquement possible est intrinsèquement inapte à la tâche de la construction d'une bombe à retardement qui est à la fois déterministe et construite à la manière spécifiée plus haut dans ce paragraphe. La matière actuelle, la matière qui obéit aux règles de la mécanique quantique, est intrinsèquement incapable de porter avec elle les dispositions parfaitement déterminées de son comportement futur que requiert le déterminisme strict.

Il semble donc que la vérité de la mécanique quantique soit incompatible non seulement avec le déterminisme en général, mais également avec la doctrine plus modeste que les états passés de l'univers, observables ou non, déterminent une succession unique d'états futurs *observables*.

(b) Dans la physique classique, l'état momentané d'un système physique ou, en tout cas, d'une masse ponctuelle) est représenté par un ensemble de vecteurs du système – l'ensemble contenant position et les vecteurs dynamiques du système. En physique quantique, cependant, l'état momentané d'un système est représenté par une *fonction*, la *fonction- Ψ* du système. Tout comme les lois de la mécanique classique déterminent la manière dont un système

². Même si, comme nous allons le voir, cette survenance (*supervention*) est purement statistique, si la mécanique quantique est correcte.

classique changera d'état avec le passage du temps (en réponse aux forces qui lui sont imprimées), de même les lois de la mécanique quantique déterminent la façon dont un système quantique changera d'état avec le passage du temps. J'utilise le mot 'déterminer' tout à fait délibérément : étant donné l'état quantique d'un système clos (comme le monde physique dans son entier) à un moment donné, un seul « futur » - une séquence d'états quantiques – est possible pour ce système³. Il semble alors que la mécanique quantique soit une théorie déterministe.

Il y a un sens à décrire les choses ainsi. Mais si nous les décrivons de cette manière, nous devons alors distinguer la question de savoir si la mécanique quantique est une théorie déterministe au sens du paragraphe précédent et la question de savoir si la mécanique quantique implique le déterminisme au sens du chapitre III. Bien que la mécanique quantique puisse être une théorie déterministe, supposer que sa vérité implique le déterminisme ne me semble pas du tout plausible. Ce serait une supposition plausible si *toutes* les propriétés d'un système physique à un moment étaient déterminées par l'état quantique du système à ce moment. Plutôt, il n'y a en général qu'une corrélation *statistique* entre le fait qu'un objet est dans certain état quantique – que cet objet soit un observable ou un objet sous-microscopique - et le fait qu'il possède une propriété observable ou mesurable. Considérons le démon omniscient de Laplace. Même s'il lui était possible de connaître l'état quantique du monde entier au moment présent, le maximum qui puisse suivre du fait que la théorie quantique est déterministe est ceci : le démon connaîtrait l'état quantique du monde à tous les moments futurs. Mais de la supposition qu'il connaîtra l'état quantique du monde, par exemple, en 2000, on ne peut pas inférer qu'il saura si je serai en vie ou mort cette année là ni où je serai, ni quels seront les problèmes philosophiques qui m'intrigueront. Au plus, me semble-t-il, le démon serait-il capable d'assigner des probabilités à plusieurs hypothèses alternatives sur ma condition cette année-là. Quant à savoir si une hypothèse raisonnablement spécifique recevrait une assignation de probabilité plus élevée de façon significative que les hypothèses tout aussi spécifiques avec lesquelles elle est en compétition, je n'en ai pas la moindre idée. Le démon ne pourrait certainement pas faire de prédictions très détaillées sur les bombes à retardement indéterminées que nous avons considérées plus haut ; il ne serait capable de nous dire que des choses comme : « la probabilité que la bombe explose dans l'heure prochaine est de 0,7362. Donc, même le démon dont Laplace pensait qu'il n'aurait pas d'usage de la probabilité trouverait ce concept indispensable dans un monde dont la mécanique quantique donne une description correcte.

Autrement dit, il trouverait ce concept indispensable s'il s'intéressait aux propriétés autres que celles qui appartiennent à un système en vertu de son état quantique. Il est simplement possible que quelqu'un maintienne que, en un sens, seules les propriétés quantiques sont « réelles » : qu'une image correcte de la réalité ultime ne donnerait de place que pour les propriétés quantiques et doit exclure toutes les propriétés comme *être rouge*, *être en mouvement*, ou *penser à la mécanique quantique*. Un philosophe qui aurait adopté cette position pourrait dire de manière consistante que la mécanique quantique est une théorie correcte de la réalité physique et que le déterminisme est vrai. Mais une telle position extrême risque de ne pas gagner beaucoup d'adeptes, et je ne la discuterai pas davantage.

(c) De nombreux philosophes et savants ont du mal à accepter que la mécanique quantique soit plus qu'un bouche-trou, une théorie qui marche bien et qui est certainement meilleure que rien, mais qui *doit* d'une façon ou d'une autre être supplantée par une théorie déterministe.

³. C'est en tout cas aussi vraisemblable que l'est dans le cas de la mécanique classique la thèse sur les états classiques. Cf. Richard Montague, « Deterministic Theories », in Richmond H. Thomason, ed., *Formal Philosophy, op. cit.*, p. 332-6

(C'était par exemple, la pensée d'Einstein.) Ce remplacement pourrait aboutir de deux façons.

D'abord, la mécanique quantique pourrait être supplantée par une théorie incompatible avec elle. Sans doute, ce devrait être une théorie nouvelle et révolutionnaire, dont les postulats seraient « dès le départ » différents de ceux de la mécanique quantique. Nul ne peut dire, bien entendu, que cela ne peut pas arriver. Mais il vaut la peine de noter qu'aucune théorie de ce genre n'a fait son apparition ; la mécanique quantique n'a pas de rival sérieux⁴.

Deuxièmement, la mécanique quantique pourrait être supplantée par une théorie compatible avec elle mais plus compréhensive. Le traitement, par Einstein, de la théorie statistique du mouvement brownien donne un exemple du type de possibilité auquel je pense. Dans un article classique, Einstein a montré que les lois statistiques décrivant le trajet bizarre (la « marche de l'ivrogne ») d'une particule de poussière en suspension dans un fluide (le mouvement brownien) pourrait être dérivée de certaines thèses plausibles concernant l'agitation de la particule par le mouvement thermique des molécules dont le fluide était composé ; il a montré cela, qui plus est, de façon tout à fait indépendante de la question de savoir si ces mouvements thermiques étaient eux-mêmes déterministes. En ce sens, Einstein a montré comment une théorie statistique pourrait être enveloppée dans une théorie déterministe (la théorie essentiellement Newtonienne de la mécanique du mouvement thermique). Des cas de ce genre pourraient conduire à se demander si quelque chose du même ordre ne pourrait pas être le cas avec la mécanique quantique. Ne pourrait-il pas y avoir un micro-microcosme déterministe sous-jacent au comportement apparemment indéterministe du microcosme, d'une manière analogue à la manière dont le mouvement moléculaire, que l'on peut présumer déterministe pour les besoins de l'exemple, est sous-jacent au mouvement brownien apparemment indéterministe ? Ou, pour poser la question de manière plus formelle, ne serait-il pas possible d'envelopper la mécanique quantique dans une théorie déterministe⁵, à la façon dont Einstein a enveloppé la théorie statistique du mouvement brownien dans la mécanique classique ? (Les théories déterministes dans lesquelles la mécanique quantique pourrait être enveloppée sont souvent appelées des théories des « variables cachées », les variables « cachées » étant les variables de la théorie qui représente les traits déterministes sous-jacents d'un système physique qui ne sont pas représentés par la mécanique quantique.)

Les problèmes soulevés par cette question sont très difficiles. Comme je les comprends – et je les comprends difficilement – la réponse semble être « Probablement pas ». On a largement cru autrefois que Von Neumann avait montré que la structure mathématique de la théorie quantique – à la différence de la structure mathématique du mouvement brownien – est telle qu'elle interdit son enveloppement par une théorie déterministe. Autrement dit, on croyait qu'il avait montré que toute théorie traitant du même sujet que la mécanique quantique, et qui pourrait être enveloppée par une théorie déterministe, devait faire des prédictions différentes de celles de la mécanique quantique sur les observables, des prédictions dont on sait empiriquement qu'elles sont fausses. La réception de l'argument de von Neumann est une histoire tortueuse, que je ne chercherai pas à résumer⁶. L'argument n'a jamais été accepté par Einstein ni par de Broglie, et a souvent été reçu avec plus de rhétorique qu'il n'est habituel pour une démonstration mathématique. Le passage suivant de deux disciples de de Broglie représente une réaction extrême mais loin d'être unique :

... le grand mathématicien von Neumann ... a même produit un théorème établissant qu'il serait impossible de

⁴. Cf. N. R. Hanson, « Quantum Mechanics, Philosophical Implications of », in Paul Edwards, ed., *The Encyclopedia of Philosophy*, New York and London, MacMillan, 1967, vol. 7, p. 45.

⁵. C'est-à-dire, dans une théorie selon laquelle les variables observables comme la position et la vitesse sont déterminées. Nous avons déjà vu que la mécanique quantique elle-même est en un sens une théorie déterministe, bien qu'il ne donne pas, en général, de prédictions déterministe des valeurs des variables observables

⁶. Voir Max Jammer, *The Philosophy of Quantum Mechanics*, New York, Wiley, 1974.

représenter un déterminisme caché qui reproduirait les lois statistiques de la mécanique quantique. Ce théorème étrange énonce que la théorie expliquera non seulement tous les phénomènes observables, mais aussi, en vertu de sa logique interne, qu'elle exclut toute autre explication. Pourtant, de Broglie en a donné une analyse critique, et montré que le théorème de von Neumann était fondé sur une hypothèse *ad hoc* qui a rendu le théorème entier sans valeur. Nous ne pouvons toujours pas exclure la possibilité d'un déterminisme caché.⁷

(Je doute que beaucoup des gens compétents pour juger de ces questions admettront sans réserve ce que ces auteurs prétendent sur la base de l'analyse réalisée par leur *maître*.)

La pensée courante sur ces matières est difficile à cerner. Je *crois* qu'elle revient à quelque chose comme ceci : (i) il y a des *lacunes* mathématiques dans l'argument de von Neumann ; (ii) on a construit d'autres arguments rigoureux pour la même conclusion⁸ ; (iii) même le plus rigoureux de ces arguments dépend nécessairement de certaines hypothèses sur la mesure et la nature des « variables cachées » dont l'existence possible fait l'objet de recherches. Les hypothèses qui ont été faites sont attractives et apparemment évidentes ; mais des hypothèses attractives et apparemment évidentes ont dû être rejetées en plusieurs occasions notables dans la physique du XXe siècle. Et même s'il est possible d'envelopper la théorie quantique dans une théorie déterministe, nul n'a réalisé cela d'une manière telle que les experts en théorie quantique la jugent à la fois logiquement satisfaisante et dotée d'un « intérêt physique ». Par conséquent, semble-t-il, ce qu'il y a de plus raisonnable à croire à l'heure actuelle pour un non-expert c'est qu'il est peu vraisemblable que la mécanique quantique sera enveloppée dans une théorie déterministe, et qu'il pourrait bien être impossible que cela se produise.

Ces commentaires et qualifications ayant été faits et dûment notés, il paraît encore raisonnable de penser que la science montre que le déterminisme est faux, et il semble tout à fait impossible de dire que la science montre que le déterminisme est vrai. Mais certains philosophes critiquent cette conclusion. J'en ai entendu argumenter plus ou moins comme ceci :

La mécanique quantique est très satisfaisante. Mais nul ne suggère sérieusement que des mécanismes de « déclenchement » comme ceux de vos bombes à retardement fassent partie des antécédents de l'action humaine. Et si aucun mécanisme de ce genre ne fait partie de nous, alors le degré d'indétermination que la mécanique quantique assigne aux pensées, aux paroles, aux actions humaines est si petite qu'elle n'a pas de signification expérimentale. Un être humain pourrait donc être un système tout aussi déterministe qu'un ordinateur digital, même si la mécanique quantique est correcte et essentiellement indéterministe. Et un ordinateur (à moins que ses concepteurs n'y aient délibérément intégré un mécanisme indéterministe) est certainement le genre de choses qui est déterministe – pour tout ce qui est pratiquement pertinent – même si, pour des raisons théoriques, il peut être sujet à un degré d'indétermination trop petit pour être mesuré. Assurément, les grands succès des physiologistes du cerveau dans la compréhension de son fonctionnement (notre « hardware ») et les succès aussi importants des psychologues dans la compréhension de nos fonctions cognitives (notre « software ») conduisent à la seule conclusion possible : un être humain, comme un ordinateur, est, pour tout ce qui est pratiquement pertinent, une chose dont le comportement futur est entièrement déterminé par ses états passés et les « entrées » actuelles.

J'accepte, et tiens pour importante, la distinction entre détermination stricte et détermination « pour tout ce qui est pratiquement pertinent ». Et je ne doute pas que, si la

⁷. J. Angrade e Silva et G. Lochak, *Quanta, grains et champs*, Paris, Hachette, 1969

⁸. Par Jauch et Piron, par exemple. Pour une présentation raisonnablement accessible de leur argument, voyez les chap. 5, 6 et 7 de Josef M. Jauch, *Foundations of Quantum Mechanics* Reading (Mass.), Addison-Wesley Educational Publishers, 1968

détermination stricte est, comme je l'ai soutenu, incompatible avec le libre arbitre et la responsabilité morale, alors la détermination en ce sens très légèrement plus faible est aussi incompatible avec les deux. Il serait, en fait, assez facile d'adapter les arguments du chapitre III pour montrer que le « quasi-déterminisme » est incompatible avec le libre arbitre, même si les arguments adaptés seraient comparativement négligés. Mais pour ce qui est de l'argument des « grands succès », je ne peux qu'y voir un coup de bluff. J'ai eu beaucoup de difficultés à voir pourquoi tant de philosophes semblent penser que les résultats de l'étude empirique des êtres humains donnent du crédit à l'hypothèse que le comportement humain est déterminé. (Par la suite, j'utiliserai 'déterminé' pour signifier 'toujours déterminé, pour ce qui est pratiquement pertinent, par des stimuli passés et présents'.) Je ne veux pas dire que je crois que les investigations empiriques ont montré que le comportement humain n'est pas déterminé. Ma difficulté est tout simplement celle-ci : l'organisme humain et le comportement humain sont des réalités d'une telle complexité (en comparaison avec ce qu'il y a à connaître), qu'il est difficile de voir pourquoi quelqu'un pourrait penser que ce que nous savons rende raisonnable une croyance que le comportement humain est déterminé. Je peux seulement conclure que ces philosophes sont convaincus, sur une base a priori, ou peut-être sur aucune base, que le comportement humain est déterminé, et, en raison de cette conviction, sont prédisposés à tenir à peu près tout pour un élément de preuve de cette croyance. (S'il en est ainsi, cela ne devrait pas constituer une surprise. Des cas semblables sont légion. De nombreux théistes sont prédisposés à surestimer la valeur de preuve d'un dessein apparent dans la nature, et de nombreux athées la valeur de preuve de la douleur et de la souffrance.)

Regardons les choses de cette façon. Il y a plusieurs thèses qui, si elles se révèlent vraies, ne pourraient conduire une personne instruite qu'à prétendre que le monde a été, épistémiquement parlant, très injustement arrangé. L'astrologie est un bon exemple de cette thèse. Si l'astrologie se révélait vraie, alors le monde devrait être un genre de monde très différent de ce que la science contraint ceux qui lui prêtent attention à penser qu'il est. (Je ne parle pas de la possibilité qu'il y ait un grain de vérité dans l'astrologie ; j'envisage le cas où il serait avéré qu'il n'y a que de la vérité dans l'astrologie.) La vérité ultime de l'astrologie est une possibilité peut-être consistante avec la totalité de mon expérience sensible à ce jour ; néanmoins, si l'astrologie se révélait vraie, je pourrais seulement dire que le monde était une *tricherie épistémique*, un monde qui a été faussé, comme une roulette truquée ou un accessoire de magicien, afin de réaliser un divorce absolument déloyal de l'apparence et de la réalité. Mais supposons qu'il soit avéré que le comportement des êtres humains n'est pas déterminé. Supposons, pour rendre cette question un petit peu plus abstraite, qu'il soit avéré que les résistances électriques à travers certaines synapses cérébrales changent à l'occasion de manière indéterminée et que ces changements soient d'une grandeur suffisante et suffisamment fréquente pour rendre fausse l'hypothèse que le comportement humain est déterminé⁹. Quelles théories physiologiques devraient-elles être écartées ? Quelles expériences bien attestées devraient-elles être regardées comme ayant produit des résultats égarants ? Quelles parties de nos manuels sur la structure et le fonctionnement du cerveau devraient-elles être réécrites ? La réponse, me semble-t-il, est que la découverte que j'ai imaginée ne révélerait aucun défaut dans nos théories et nos manuels, simplement parce qu'aucun scientifique n'a jamais prétendu être capable de prédire le cours actuel des événements dans un cerveau vivant, tout comme aucun météorologiste n'a jamais prétendu être capable de prédire le temps de la semaine prochaine dans tous les détails.

Considérons une analogie. Supposons que nous découvrions d'une manière ou d'une autre que des démons ont influencé le comportement humain pendant des siècles, en introduisant de menus changements dans les cerveaux humains – en introduisant, en fait, des changements dans

⁹. Cf. C.D. Broad, *The Mind and its Place in Nature*, London, Routledge and Kegan Paul, 1925, p. 113.

les résistances électriques entre certaines synapses. Je crois qu'il devrait être assez clair qu'aucune de nos recherches dans le fonctionnement du cerveau ne rend cette supposition le moins du monde improbable. Si nous l'apprenions, nous saurions quelque chose de nouveau sur les sources du comportement humain, mais aucune de nos informations péniblement acquises sur le fonctionnement du cerveau ne devrait être écartée ou modifiée en quelque façon. Nul n'aurait le droit de s'exclamer : « Mais cela va contre ce que nous croyions avoir découvert sur le cerveau ! ». Il est sans doute vrai que nous sommes parfaitement justifiés à assigner une probabilité subjective négligeable à l'hypothèse que des démons ont affecté nos cerveaux de la façon que j'ai imaginée. Mais cette justification ne repose pas sur nos découvertes *empiriques* sur le cerveau. (La question « Sur quoi repose-t-elle, alors ? » est une question que je laisserai aux théoriciens de la connaissance.) Si nous devions apprendre qu'il y a de nombreuses races intelligentes dans l'univers, dont exactement la moitié est sous l'influence de démons manipulateurs, toute notre énorme connaissance de la structure et du fonctionnement du cerveau ne serait d'aucune aide pour décider si notre propre race est sous une telle influence, car les effets proches des manipulations démoniaques seraient perdues dans la nature touffue incomparablement vaste du cerveau.

Assurément le cas est le même si nous nous interrogeons non pas à propos de l'influence démoniaque, mais à propos du changement indéterminé. Supposons que Dieu nous dise que les cerveaux de la moitié des races dans l'univers sont déterminés, et que les cerveaux de l'autre moitié sont sujets à de menus changements indéterminés du genre que j'ai imaginé. Comment pourrions-nous découvrir à quelle race nous appartenons ? Il semble qu'il n'y ait pas de raison *a priori* de penser qu'il y aurait une différence entre les cerveaux des deux sortes de races que l'on pourrait détecter en pratique. Il semble donc tout à fait déraisonnable de dire que nous avons une raison empirique de dire que nous sommes déterminés.

Je ne dis pas que nous n'avons aucune raison de dire que nous sommes déterminés. Je dis que toute raison que nous pourrions avoir doit être au moins en partie une raison *a priori* ou conceptuelle. Par exemple, quelqu'un pourrait dire que nous sommes déterminés parce que, si nous ne l'étions pas, notre comportement serait une série de mouvements décousus, et notre parole un pur babillage. J'ai dit au chapitre IV pourquoi je pensais que cette idée était erronée. Ma présente remarque est qu'elle n'est fondée sur aucune recherche empirique concernant notre manière de fonctionner, mais plutôt sur l'identification d'une séquence indéterminée avec une séquence aléatoire, identification qui repose, sans doute, sur une base purement conceptuelle. Cet argument n'est ni meilleur ni pire aujourd'hui qu'il ne l'était, par exemple, au dix-septième siècle : les réussites des sciences empiriques sont tout simplement sans pertinence pour la question de savoir si c'est un argument contraignant.

Je ne dis pas que nous n'avons jamais de preuve empirique en faveur de la thèse qu'une chose était déterminée « à toutes fins pratiques ». Si je n'avais jamais vu de montre et en heurtait une du pied un jour en traversant une lande, j'ose dire que je pourrais assez facilement découvrir que son comportement était déterminé par ses états antérieurs, et par les influences externes auxquelles elle était soumise à tel moment. Je pense que les résultats d'une telle enquête seraient tout aussi fiables que peuvent l'être des résultats qui exigent qu'on les énonce au moyen de hautes abstractions comme 'détermination' et 'état', et qu'on pourrait qualifier ces recherches d'*empiriques* de manière parfaitement appropriée. Mais nous ne sommes pas des montres. Déterminés ou non, nous sommes trop complexes pour que nos mécanismes internes puissent être décortiqués comme ceux d'une montre.

Je ne dis pas que nous ne pourrions pas avoir de preuve empirique pour la thèse qu'un homme était déterminé, « à toutes fins pratiques ». Peut-être sera-t-il possible dans le futur de construire des modèles informatiques de cerveaux particuliers qui nous rendront capables de prédire comment une personne particulière réagira à divers stimuli. Si ces stimuli étaient

d'une complexité comparable à celle des stimuli que nous rencontrons dans la vie ordinaire, et si les prédictions étaient toujours, ou presque toujours, correctes, alors nous aurions bien une excellente raison de croire que les êtres humains sont déterminés. Et, bien sûr, il y a presque certainement d'autres sortes de preuves, des sortes de preuves que je ne peux pas imaginer, et qui rendraient très probables l'hypothèse que nous sommes déterminés. Mais aucune preuve de ce genre n'a encore été présentée, et, même si les êtres humains sont déterminés, il semble qu'il faille attendre très longtemps avant qu'on en présente une.

Récapitulons : il semble n'y avoir que deux manières dont les données scientifiques pourraient nous convaincre que nous sommes déterminés ; d'abord, nous pourrions le croire, comme l'a fait Laplace, sur la base de nos théories physiques les plus générales (qui s'appliquent à tous les systèmes physiques et donc à nous) ; ensuite, nous pourrions le croire sur la base de l'étude empirique de l'homme. Mais nos théories physiques les plus générales ne sont plus déterministes. Et l'étude empirique de l'homme a encore beaucoup à faire avant de pouvoir nous dire quoi que ce soit quant à notre éventuelle détermination.

(ii) *La raison montre que le déterminisme est vrai.* Il y a des philosophes qui pensent que le déterminisme est une vérité de raison. C'est-à-dire qu'il y a des philosophes qui prétendent voir, par une sorte d'intuition conceptuelle ou ontologique, que les lois de la nature *doivent* être suffisamment « serrées » pour que, à tout moment, il n'y ait qu'un seul futur possible. Je ne partage pas cette intuition et je doute qu'elle existe vraiment. Bien plus, je ne connais pas d'argument très plausible qui procède à partir de prémisses connues *a priori* pour conclure que le monde est entièrement déterminé¹⁰.

Il y a un principe à partir duquel le déterminisme peut être déduit, cependant, et sur lequel j'ai mieux à dire que simplement « je n'y crois pas ». C'est le fameux Principe de Raison Suffisante :

(PRS) Pour tout état de choses qui se réalise, il y a une raison suffisante pour qu'il se réalise.

(Dans ce principe, 'état de choses' doit être compris au sens de la section 5.6) On peut soutenir de manière plausible que ce principe implique le déterminisme. Supposez, en effet, que le déterminisme soit faux. Alors il y a un temps après lequel deux futurs sont possibles. Mais l'un des deux doit finir par se réaliser ; et il est difficile de voir quelle « raison suffisante » il y aurait pour la réalisation de *ce* futur particulier. Ce raisonnement est cependant loin d'être sans faille. On pourrait suggérer, par exemple, que le futur actuel est devenu actuel pour aucune raison que l'on puisse trouver dans le monde naturel, mais plutôt parce que Dieu a choisi qu'il se réaliserait, le choix de Dieu étant dans ce cas la raison suffisante que réclame PRS. Il est donc clair que nous aurions besoin de prémisses supplémentaires pour déduire le déterminisme de PRS. Quoi qu'il en soit, PRS doit être rejeté, car il a une conséquence absurde : l'effondrement de toutes les distinctions modales.

Pour le montrer, nous devons jeter un rapide coup d'œil sur le concept de raison suffisante. Je ne sais comment donner une explication satisfaisante de ce concept – si un concept est bien attaché aux mots 'raison suffisante' – mais je crois que je peux voir certains traits que doit avoir tout concept cohérent pouvant raisonnablement être exprimé par ces mots. D'abord, si x est une raison suffisante pour y , alors x doit impliquer y . C'est-à-dire qu'il doit être impossible que x se réalise sans que y se réalise. Car s'il était *possible* que x se réalise et que y manque à se réaliser, comment la réalisation de x pourrait-elle être une raison *suffisante* pour celle de

¹⁰. Ce n'est pas dire qu'il n'y a pas d'arguments plausibles *a priori* pour la conclusion que tout événement a une cause. Il y en a peut-être.

y ?¹¹ Ensuite, aucun état de choses *contingent* ne peut être sa propre raison suffisante. Cela semble être une caractéristique essentielle du concept de raison suffisante. (J'introduis la qualification 'contingent' pour satisfaire ceux qui pensent qu'un état de choses nécessaire est sa propre raison suffisante. Qu'il en soit ainsi ou non n'aura pas d'influence sur notre argument.)

Nous pouvons maintenant montrer que PRS conduit à l'effondrement de toutes les distinctions modales. Soit P la conjonction, en une seule proposition, de toutes les propositions vraies de manière contingente¹². (Dans ce qui suit, j'utiliserai 'P' indifféremment pour dénoter la conjonction de toutes les propositions vraies de manière contingente et pour l'état de choses qui consiste en la vérité de ces propositions. Dans le langage de la section 5.6, ce vaste état de choses serait dénoté en préfixant 'le fait qu'il soit vrai que' à une phrase exprimant P. Cet état de choses est ce que Plantinga appelle 'le monde actuel'. Ce que j'appelle 'le monde actuel' dans la section 3.6 est la possibilité que P soit vraie.) Il est évident que P elle-même est une proposition contingente, car une proposition nécessaire ne peut pas avoir une seule proposition contingente comme conjoint. Or, selon PRS, il y a un état de choses S qui est une raison suffisante pour P. S doit être contingent ou nécessaire. Mais il ne peut être ni l'un ni l'autre. Il ne peut pas être nécessaire car, s'il était nécessaire alors P (qui, d'après notre premier principe, est impliqué par S) serait nécessaire. Il ne peut pas être contingent, car s'il était contingent, il serait un des conjoints de P ; et s'il était un conjoint de P, il serait impliqué par P ; et s'il était impliqué par P, il impliquerait P et serait impliqué par P ; auquel cas il serait P (par le critère d'identité des états de choses donné dans la section 5.6¹³) ; et si S *était* P, alors un état de choses contingent serait sa propre raison suffisante, contrairement à notre second principe. Puisque S ne peut être ni nécessaire ni contingent, il ne peut pas exister, et PRS est faux.

Ce résultat s'ensuit si nous admettons qu'il y a quelque chose comme P – à savoir la conjonction de toutes les propositions vraies de manière contingente. Et il y a une telle chose dès lors qu'il y a ne serait-ce qu'une proposition vraie de manière contingente. Donc, si PRS est vrai, il n'y a d'autres vérités que les vérités nécessaires : aucune distinction ne doit être faite entre vérité et nécessité¹⁴. De grands philosophes l'ont cru – Spinoza par exemple. Mais je doute que beaucoup de philosophes contemporains puissent même considérer sérieusement cette hypothèse. Je sais que je ne le pourrais pas. (Notamment parce que, étant donnée la règle d'inférence appelée (α) dans la section 3.10, cela impliquerait que personne n'a le libre

¹¹. Leibniz a cru, bien entendu, que le fait que ce monde soit le meilleur des mondes possibles était une raison suffisante pour qu'il soit actuel, malgré le fait que *ce monde est le meilleur des mondes* n'implique pas que *ce monde est actuel* (car si cette implication était valide, alors les autres mondes ne seraient pas possibles ; c'est-à-dire qu'il ne pourraient pas être actuels).

¹². Quelqu'un pourrait s'inquiéter de savoir si cette définition imprédicative conduit au paradoxe de l'auto-référence. Ce n'est pas le cas. Le modèle suivant le montre. Identifiez une proposition à l'ensemble des mondes dans lesquelles elle est vraie, comme nous l'avons fait à la section 3.3. La conjonction de deux propositions sera alors leur intersection et un conjoint d'une proposition sera toute proposition dont la première est un sous-ensemble. Ainsi, toute proposition est un conjoint d'elle-même. Dans ce modèle, 'P' dénote l'ensemble dont le seul membre est le monde actuel (pourvu qu'il y ait au moins deux mondes possibles : s'il n'y a qu'un monde, toutes les propositions sont nécessaires, et 'P' n'a pas de dénotation), un ensemble qui peut être pris à part sans paradoxe. Comparez cette identification de P dans notre modèle avec les identifications faites dans la remarque entre parenthèses qui suit dans le texte.

¹³. Ce critère d'identité n'est pas réellement nécessaire pour l'argument. Si nous le rejetons, nous devrions remplacer notre second principe par le principe qu'aucun état de choses contingent ne peut avoir pour raison suffisante un état de choses auquel il est logiquement équivalent. Et cela semble assez plausible.

¹⁴. Il suit de ce résultat que PRS implique bien le déterminisme, pour la simple raison que le déterminisme est une thèse qui, conçue très abstraitement, affirme seulement que certaines propositions vraies sont impliquées par certaines autres propositions ; mais si la vérité et la vérité nécessaire coïncident, alors tout ensemble de propositions vraies est impliqué par n'importe quoi.

arbitre). Nous devons donc rejeter PRS et avec lui la seule tentative plausible de montrer que le déterminisme est une vérité de raison.

Puisque, comme nous l'avons vu, ni la science empirique, ni la raison pure ne montrent que le déterminisme est vrai, ou même ne nous donnent de bonnes raisons de penser que le déterminisme est vrai, nous pouvons seulement conclure que la croyance au déterminisme est – au moins pour le temps présent – totalement injustifiée.

6.3 Quelles raisons y a-t-il de penser que nous avons ou que nous n'avons pas le libre arbitre ?

Je ne crois pas qu'il y ait la moindre façon de simplement découvrir si nous avons le libre arbitre, dans le sens où nous pouvons découvrir s'il y a de la vie sur Jupiter. Certains philosophes croient qu'il est possible de découvrir si nous avons le libre arbitre par *introspection*. Mais cela semble tout bonnement faux, de manière évidente, puisque, s'ils avaient raison, nous pourrions découvrir par introspection si nous avons été équipés de dispositifs martiens comme ceux que nous avons imaginés dans notre discussion de l'Argument du Cas Paradigmatique, au chapitre IV. Et, bien sûr, c'est ce que nous ne pouvons pas faire.

Il est certainement vrai que la plupart d'entre nous sommes parfaitement certains que nous avons le libre arbitre. Mais il n'y a aucune raison de penser que notre parfaite certitude sur cette question dérive d'un « accès direct », en quelque sorte, que nous aurions aux prémices de l'action, à la façon dont certains philosophes croient que nous avons un accès direct à nos propres états mentaux. Elle dérive plutôt, selon moi, de notre connaissance, le plus souvent implicite et inarticulée, que l'on ne peut pas délibérer sans croire à son propre libre arbitre. Il me semble raisonnable de supposer que pour toute espèce d'êtres rationnels, et pour tout type d'activité mêlée de manière inextricable à pratiquement tous les aspects de leurs vies – comme la délibération concernant une action future est inextricablement mêlée à presque tous les aspects de *nos* vies – s'il y a une proposition dont l'acceptation inconsciente est un présupposé de cette activité, alors l'assentiment théorique, réfléchi, à cette proposition sera commun, s'il n'est pas universel, à tous les membres de cette espèce. Les philosophes qui appartiennent à cette espèce peuvent, sans contredire ma thèse, poser parfois des questions sur la manière de justifier ces propositions, présupposées par presque tous les aspects de leur vie ; certains peuvent même, sans contredire ma thèse, rejeter *verbalement* ces propositions¹⁵.

C'est parce que la proposition que nous avons le libre arbitre est liée de manière inséparable à notre vie délibérative, selon moi, que la plupart d'entre nous sommes certains d'avoir le libre arbitre. Mais si c'est bien l'explication correcte de notre certitude, alors cette certitude est sans valeur de preuve. Notre certitude sur nos propres états mentaux conscients peut venir du fait que chacun de nous occupe, nécessairement, une situation privilégiée dans une discussion sur le point de savoir s'il a le libre arbitre. Si chacun de nous a, implanté dans son cerveau, un dispositif martien comme celui que nous avons imaginé, alors ce sont les Martiens, et non pas nous, qui sont experts sur la question de savoir si nous avons le libre arbitre.

Y a-t-il une autre manière que l'introspection pour découvrir si nous avons le libre arbitre ? Eh bien, si le déterminisme est vrai, nous pourrions découvrir que nous n'avons pas le libre arbitre, en découvrant que nous sommes déterminés. Mais les difficultés qui s'opposent à ce que l'on découvre que nous sommes déterminés ont déjà été discutées. Qui plus est, même si le déterminisme est faux, il semble que ce soit un vœu pieux de supposer que nous découvrirons que nous avons le libre arbitre en découvrant que le déterminisme est faux.

¹⁵. Cf. la discussion du Baron d'Holbach, section 5.2.

D'abord, il est peu vraisemblable que nous découvririons que nous sommes indéterminés (au sens de la section 6.2), même en supposant que nous sommes indéterminés : nous sommes tout simplement trop compliqués pour qu'une telle découverte soit vraisemblable. Ensuite, même si nous découvriions que nous étions indéterminés, cette découverte ne montrerait pas que nous avons le libre arbitre. J'ai soutenu en plusieurs endroits de ce livre que le fait d'être indéterminé était une condition nécessaire pour avoir le libre arbitre, et n'était pas une condition suffisante pour ne pas avoir le libre arbitre. Mais je n'ai jamais dit que le fait d'être indéterminé était une condition suffisante pour avoir le libre arbitre. Je ne l'ai pas dit, parce que c'est faux : la proposition qu'il est indéterminé si je lèverai ou non ma main à un certain moment n'implique pas la proposition qu'il est en mon pouvoir de lever ma main à ce moment-là. Supposez, par exemple, qu'il soit physiquement possible que je lève ma main dans une minute. Supposez aussi que je suis lié à une bombe dont l'explosion est temporellement indéterminée du type de celle que nous avons considéré dans la section précédente. Supposez que la bombe va en fait exploser dans 30 secondes. Alors, bien qu'il y ait des futurs consistants à la fois avec le présent état du monde et avec les lois de la nature dans lesquels je lève ma main dans une minute, je n'ai pourtant pas le choix de lever ma main dans une minute, puisque je serai en morceaux dans une minute et que je n'ai pas le choix sur cela¹⁶.

Donc, il semblerait qu'il n'y a pas d'espoir de découvrir si nous avons le libre arbitre en découvrant si le déterminisme est vrai. Et il n'y a pas d'autre moyen (outre l'introspection) dont je puisse penser qu'il apparaisse seulement pertinent pour la question de savoir si nous avons le libre arbitre. Notre position est-elle alors sans espoir ? Je crois que non. Revenons un instant aux conséquences du rejet de la thèse du libre arbitre, un sujet que nous avons discuté longuement au chapitre V. Nous avons vu dans ce chapitre que le rejet de la thèse du libre arbitre impliquait qu'il n'y a rien de tel que la responsabilité morale. Et c'était là une conclusion peu surprenante : on aurait du mal à trouver quelqu'un qui ait supposé que nous pouvions être tenus pour responsables de ce que nous faisons, si nous n'avons pas le choix sur ce que nous faisons. Mais si la réalité de la responsabilité morale implique l'existence du libre arbitre alors je suggère que nous avons une raison parfaitement bonne, en fait si bonne qu'on n'en voit pas de meilleure, de croire au libre arbitre. Car assurément nous ne pouvons pas douter de la responsabilité morale ?

Il y a peut-être des gens qui, non seulement en doutent, mais rejettent la thèse qu'ils sont moralement responsables. (Je pense aux psychopathes. Je dis 'peut-être' parce que j'ai du mal à envisager clairement ce qui se passe dans l'esprit d'un homme qui viole et assassine une petite fille et ne ressent ensuite aucun remords). Mais il y a peu de gens, si même il y en a, qui réagiront à un acte blessant gratuit qui leur serait infligé de manière délibérée par un être humain de la même manière qu'ils réagiraient si la même blessure leur était infligée par un coup de foudre ou par une branche brisée par le vent. Quand une *personne* nous blesse – au moins si nous croyons qu'elle savait ce qu'elle faisait, et qu'elle pouvait s'en empêcher – nous avons une manière humaine caractéristique de réagir : nous blâmons, nous grondons, nous haïssons, nous réfléchissons sur la futilité de la haine, nous planifions une vengeance, nous nous rappelons que le désir de vengeance est un désir d'usurper une prérogative divine. Ce que nous ferons, parmi ces possibilités, sera sans doute en partie fonction de notre constitution et de notre éducation. Mais que nous ferons l'une de ces choses découle du fait que nous sommes humains, si ce n'est pas à partir du seul fait d'être rationnel. Et réagir d'une de ces manières c'est montrer plus sûrement que ne pourrait le faire le discours le plus élevé que nous croyons en la responsabilité morale.

¹⁶. Pour un autre exemple, voyez la discussion de la relation entre possibilité physique et pouvoir d'agir, à la section 1.4.

J'ai écouté des philosophes qui rejettent l'existence de la responsabilité morale. Je ne peux pas les prendre au sérieux. Je connais un philosophe qui a écrit un article où il rejette la réalité de la responsabilité morale. Et pourtant ce même philosophe, le jour on lui a volé plusieurs de ses livres, a dit : « C'était une mauvaise action ! » Mais personne ne peut dire de manière cohérente qu'un certain acte était une mauvaise action *et* dire que son agent n'était pas moralement responsable quand il l'a accompli : ceux qui ne sont pas moralement responsables de ce qu'ils font peuvent peut-être mériter notre pitié ; ils ne méritent sûrement pas notre condamnation.

Ce que j'ai dit dans les derniers paragraphes à propos de notre croyance dans la responsabilité morale est à certains égards semblables à ce que j'ai dit au chapitre V à propos de la délibération et de notre croyance au libre arbitre. Mais il y a une différence. Le philosophe qui nie l'existence du libre arbitre se contredit continuellement lui-même parce que son comportement non verbal manifeste constamment sa croyance au libre arbitre. Mais je suggère que le philosophe qui rejette la responsabilité morale dit des *paroles* qui contredisent ses théories, des paroles comme « c'était une mauvaise action ». Ce n'est pas seulement que ses actions trahissent ses mots (bien que ce soit aussi vrai, bien entendu), mais que ses mots trahissent ses mots. J'ai suggéré au chapitre V qu'il nous serait impossible de cesser de nous comporter de diverses manières qui manifestent une croyance au libre arbitre. Mais je ne crois pas qu'il soit *impossible* de cesser de parler de diverses manières qui manifestent une croyance en la responsabilité morale. Ce serait simplement très, très difficile. Je vous demande d'essayer d'imaginer à quoi cela ressemblerait de ne jamais faire de jugements comme 'Quelle manière méprisable de se comporter' ou 'On penserait qu'une personne douée des mêmes talents en saurait plus que cela' ou 'Je ne peux jamais penser à ce que j'ai fait sans me sentir mal'. Si vous essayez d'imaginer cela, vous ressentirez peut-être la même chose que moi quand j'essaie de faire l'expérience. Je trouve qu'une difficulté que je vois à renoncer à faire de tels jugements ressemble beaucoup à une difficulté que j'aurais à abandonner des jugements comme 'Cette voiture est dangereuse parce qu'elle a de mauvais freins' et 'Les champignons qui poussent sous cet arbre sont vénéneux'. La difficulté tient à deux faits : de tels jugements sont souvent *corrects* et ils sont extrêmement importants pour avancer dans la vie. Pensez à un comportement que vous avez observé et que vous qualifieriez vraiment de 'parfaitement méprisable'. N'est-ce pas la bonne manière de le qualifier ? N'est-ce pas la bonne *description* ? N'est-ce pas une description aussi « objective » (quel qu'en soit le sens) que 'dangereux' ou que 'vénéneux' ?

De nombreux philosophes, je m'en doute, diront qu'ils s'expriment ainsi pour décrire le comportement des gens, mais qu'en agissant de la sorte, ils n'attribuent pas à ces personnes de responsabilité morale pour leur comportement. Cela me semble erroné. Supposons que quelqu'un ait fait quelque chose qui nous conduise à dire « c'était de sa part une action parfaitement méprisable ». Supposons que nous découvriions ensuite qu'il a fait cela peu après avoir reçu, à son insu ou sans son consentement, une drogue connue pour altérer le comportement humain de manière radicale et imprévisible. Supposons que cette découverte nous ait conduit à décider qu'il n'était pas « responsable de ce qu'il avait fait » au moment où il a réalisé cet acte. Il me semble que nous ne pourrions plus continuer à dire « c'était de sa part une action parfaitement méprisable », même en ajoutant à cette assertion « bien qu'il n'ait pas été responsable de ses actions quand il l'a fait ». Cette addition ne me semble pas être une qualification cohérente de l'assertion originale (à moins que « c'était de sa part une action parfaitement méprisable » soit tenu pour signifier « *Normalement*, ce qu'il a fait serait une action parfaitement méprisable » ; mais il n'en va pas ainsi dans le cas considéré). La raison est simple. Appeler un acte 'méprisable', c'est blâmer son agent pour l'avoir réalisé ; tandis que dire que l'agent n'était pas responsable de ce qu'il a fait quand il a réalisé un acte

c'est l'excuser pour l'avoir réalisé ; et on ne peut pas à la fois blâmer et excuser.

Nous croyons donc tous que les gens sont parfois moralement responsables de ce qu'ils font. Nous croyons tous que la responsabilité existe. Et je crois que, si nous examinons nos convictions honnêtement, sérieusement et soigneusement, nous verrons que nous ne pouvons pas croire que cet assentiment correspond à une contrainte que notre nature et la nature de la vie sociale humaine exercent sur nous, à la manière dont les manifestations dans notre comportement de l'assentiment à la proposition que nous avons le libre arbitre nous sont imposées par la simple impossibilité d'une vie sans délibération. Je crois que nous verrons que nous devons tenir notre croyance en la responsabilité morale pour une croyance justifiée, une croyance qui n'est tout simplement pas accessible à un doute raisonnable. Je voudrais même aller plus loin : selon moi, la proposition que nous sommes souvent moralement responsables pour ce que nous avons fait est quelque chose que nous savons tous être vrai.

Etre convaincu que nous savons quelque chose ne prouve évidemment pas que nous le savons, ni même que ce soit vrai. Mais il *est* vrai que nous sommes moralement responsables, non ? Et nous savons bien que c'est vrai, n'est-ce pas ?

Si nous savons que la responsabilité morale existe, alors nous ne devrions avoir aucun doute sur le point de savoir si nous avons une bonne raison de croire que nous avons le libre arbitre. C'est cela, et uniquement cela qui nous donne, selon moi, une raison de croire que nous avons le libre arbitre. Il peut bien se faire que, de façon paradoxale, nous croyons que nous sommes libres parce que nous n'avons pas le choix à l'égard de ce que nous croyons sur ce point (étant donnée la nécessité, pour celui qui délibère, de la croyance en son propre libre arbitre). Mais ce fait ne peut être pour personne une *raison* de croire au libre arbitre : ce pourrait au plus être une excuse pour qui serait accusé de croire au libre arbitre sans avoir aucune raison en faveur de cette croyance. Si quelqu'un devait *défendre* sa croyance au libre arbitre, il ne pourrait pas répondre que ni lui ni nul autre n'ont le choix sur ce qu'il croit à propos du libre arbitre. En revanche, c'est une défense de la thèse du libre arbitre tout aussi appropriée que celle de n'importe quelle position philosophique que de dire : « sans libre arbitre, nous ne serions jamais moralement responsables de quoi que ce soit ; et nous sommes parfois moralement responsables ».

Je vais maintenant considérer une objection à cet argument (ou, plus exactement, une objection à l'emploi de cet argument par un incompatibiliste) que je trouve, à certains égards, très puissante et dont je pense que de nombreux philosophes la trouveront convaincante :

Si ce que vous dites est juste, alors, parce que nous savons que la responsabilité morale existe, nous savons que nous avons le libre arbitre. Mais, selon vous, la thèse du libre arbitre implique l'indéterminisme. Et, sans doute, vous pensez que vos arguments donnent au philosophe attentif une bonne raison de croire que la thèse du libre arbitre implique l'indéterminisme. Par conséquent, si tous vos arguments sont corrects, notre connaissance alléguée de l'existence de la responsabilité morale, couplée avec certains arguments *a priori*, peuvent constituer une bonne raison de croire que le déterminisme est faux. Mais ce n'est pas la sorte de chose qui peut être une bonne raison de croire à l'indéterminisme. L'indéterminisme est, pour le dire grossièrement, une thèse sur le mouvement des particules de matière dans le vide ; la thèse plus spéciale que *nous* sommes indéterminés (au sens de la section 6.2) est une thèse sur les détails structurels et les menus fonctionnements de notre système nerveux. Seules les investigations scientifiques sont pertinentes pour la vérité ou la fausseté de telles thèses. (Cette assertion est vraie même si vous avez le droit de dire qu'il est peu vraisemblable que la science sera capable de montrer que nous sommes indéterminés). Par conséquent, vos arguments représentent simplement une tentative de plus par un philosophe de régler

par l'intuition intellectuelle et la raison pure une question qui devrait être laissée à la science empirique. Et si elle devait prouver que cette question *ne peut pas* être réglée par la science empirique, en raison peut-être de « l'obscurité de la matière et la brièveté de la vie humaine », alors nous devrions simplement choisir de n'avoir aucune opinion quant à la réponse correcte.

J'ai dit que je trouvais cet argument très puissant, à certains égards. Pourtant, mon jugement tout bien réfléchi est qu'il doit être rejeté. Son point faible peut être mis en évidence en le comparant à un argument souvent utilisé dans la discussion du scepticisme. Voici une version de cet argument, c'est le sceptique qui parle :

Vous dites que la plupart des propositions que nous acceptons sans réfléchir comme vraies *sont* vraies. Vous dites, qui plus est, que nous le savons, ou, du moins, que nous avons de bonnes raisons de le croire. Mais nous pouvons déduire de ce que vous dites que nous avons de bonnes raisons de croire qu'il n'existe pas de Trompeur Universel Cartésien – aucun être qui nous trompe sur à peu près tout. (Car s'il y avait un tel être, alors presque toute proposition que nous pensons savoir être vraie serait fausse). Or, la validité évidente de la déduction entre parenthèses donne clairement au philosophe attentif une bonne raison d'accepter la proposition conditionnelle correspondante :

Si la plupart des propositions que nous pensons savoir être vraies *sont* vraies, alors, il n'existe pas de Trompeur Universel

Mais si nous avons de bonnes raisons d'accepter cette conditionnelle et son antécédent, nous avons de bonnes raisons d'accepter son conséquent ; c'est-à-dire que nous avons de bonnes raisons d'accepter la proposition qu'il n'y a pas de Trompeur Universel. La thèse qu'il n'y a pas de Trompeur Universel est, pour le dire grossièrement, une thèse sur les caractéristiques d'une partie du monde qui est inaccessible à toute investigation humaine possible. Votre affirmation que nous avons une bonne raison de croire que la plupart de nos prétentions de connaissances sont correctes est juste un autre cas d'un philosophe prétendant avoir une bonne raison de croire quelque chose que personne ne peut avoir une bonne raison de croire.

L'anti-sceptique qui fait face à cet argument a, je crois, deux types de réponses à sa disposition.

Tout d'abord, il peut simplement répliquer que la déduction présentée par le sceptique – associée à sa propre thèse sur la validité en général de nos prétentions à la connaissance – *montre* qu'il a une bonne raison de croire qu'il n'y a pas de Trompeur Universel. Ce peut être un résultat surprenant, concèdera-t-il, mais l'argument paraît difficile à contrer. Quant à la prétention du sceptique que la thèse de la non-existence d'un Trompeur Universel n'est pas le genre de thèse que les êtres humains *pourraient* avoir une bonne raison d'accepter (continue l'anti-sceptique), la déduction du sceptique montre que cette prétention plausible au début est finalement fausse.

En second lieu, l'anti-sceptique peut souligner que l'argument dépend essentiellement de la règle d'inférence

Rp

$R(p \supset q)$

Donc, Rq

(où 'R' est mis pour 'on a une bonne raison de croire que') et conteste la validité de cette règle.

Personnellement je suis favorable à la première réponse. Mais ce n'est pas un livre d'épistémologie, et je n'en dirai pas plus sur la question de la réponse appropriée à l'argument du sceptique. Mon but, en présentant l'argument du sceptique, est de la comparer à l'argument sur le déterminisme, qui est notre présent sujet. Ces deux arguments sont très semblables. Ils sont si semblables, en fait, que le philosophe qui rejette la conclusion de l'argument sur le déterminisme a deux types de réponses disponibles, exactement parallèles aux deux types de réponses disponibles pour l'anti-sceptique.

Tout d'abord, il peut simplement répondre que l'incompatibilité de la responsabilité morale et du déterminisme - associée à sa thèse que nous savons, et avons donc de bonnes raisons de croire, que la responsabilité morale existe – *montre* qu'il a une bonne raison de croire que le déterminisme est faux. Quant à l'affirmation que le déterminisme n'est pas le type de thèse que nous pourrions avoir de bonnes raisons de tenir pour fausse, à moins que ces raisons ne soient fournies par la science empirique, il peut répondre que les arguments qu'il accepte pour l'incompatibilité du déterminisme et de la responsabilité morale montre que cette affirmation plausible au début est finalement fausse.

En second lieu, il peut souligner que l'argument du déterminisme dépend essentiellement de la règle d'inférence exposée plus haut, et contester la validité de cette règle.

Je suis favorable à la première réponse, mais je ne défendrai pas la thèse que cette réponse est la meilleure des deux. Je soutiendrai plutôt que tout philosophe qui rejette l'argument sceptique que nous avons examiné devrait aussi rejeter l'argument que nous considérons maintenant, et pour la même raison, quelle que soit cette raison. Tout philosophe qui rejette l'argument du sceptique pour la raison que 'on a une bonne raison de croire cela' ne se comporte pas comme un opérateur de nécessité devrait, bien sûr, rejeter l'argument du déterminisme, puisque cet argument dépend de la même hypothèse sur la logique des bonnes raisons. Et tout philosophe qui souhaite dire que, puisqu'il a une bonne raison de croire que nos prétentions à connaître sont généralement correctes, il a donc une bonne raison de croire qu'il n'y a pas de Trompeur Universel, ne devrait avoir aucune réticence à dire que puisqu'il a une bonne raison de croire que la responsabilité morale existe, il a donc une bonne raison de croire que le déterminisme est faux. Bien plus, tout philosophe qui accepte l'argument du déterminisme sera contraint de trouver une manière de rejeter de manière cohérente l'argument du sceptique. Supposons que ce philosophe accepte la règle d'inférence ' $Rp, R(p \supset q) \mid - Rq$ ', qui est commune aux deux arguments ; comment répondra-t-il alors au critique qui considère ainsi sa manœuvre :

Vous dites que le fait que vous avez de bonnes raisons de croire que la responsabilité morale n'est pas le genre de choses qui pourrait nous donner de bonnes raisons de croire que le déterminisme est faux. Mais vous dites *aussi* que le fait que nous ayons de bonnes raisons de croire que nos prétentions à la connaissance sont généralement correctes *est* le genre de choses qui peut nous donner de bonnes raisons de croire qu'il n'y a pas de Trompeur Universel. Mais n'est-il pas vrai que vous répondez différemment, de manière arbitraire, à des problèmes philosophiques qui sont essentiellement les mêmes ? Si la réflexion sur la connaissance humaine peut nous donner une bonne raison d'accepter une thèse sur ce qui est au-delà des limites de l'observation possible, pourquoi la réflexion sur la responsabilité morale ne nous donnerait-elle pas une bonne raison d'accepter une thèse sur le mouvement des particules de matière dans le vide ?

Je ne peux pas voir de réponse très convaincante à cela. J'en conclus que l'incompatibiliste qui croit que l'existence de la responsabilité morale est une bonne raison d'accepter la thèse du libre arbitre ne doit pas être troublé par l'accusation que ses vues l'engagent à la thèse que

l'existence de la responsabilité morale est une bonne raison de croire à l'indéterminisme. Ou, en tout cas, il ne doit pas être plus troublé par cette accusation qu'il ne l'est pas les arguments « Trompeur Universel » en faveur du scepticisme.

6.4 Nous avons vu qu'il y a toute raison de penser que le déterminisme est faux, et aucune bonne raison de penser qu'est vraie la thèse plus faible que notre comportement est déterminé à la façon dont le comportement des ordinateurs est déterminé. Nous avons vu qu'il y avait au moins une excellente raison de croire que nous avons le libre arbitre : c'est que nous sommes au moins parfois moralement responsables. Nous avons vu que le libre arbitre était incompatible avec le déterminisme. Nous sommes donc en mesure d'avancer, peut-être de manière seulement provisoire, une réponse à la question « Les êtres humains ont-ils le libre arbitre, ou sont-ils déterminés ? » - Les êtres humains ont le libre arbitre et ils ne sont pas déterminés.

Cette réponse peut être difficile à accepter pour certains. (Je ne dis pas, bien sûr, que c'est la réponse correcte de manière si évidente que l'on devrait l'accepter sans argument. Mais j'ai donné des arguments : ce livre.) Je pense à ceux qui trouvent cette réponse si peu plausible à première vue, si intrinsèquement incroyable, qu'ils répondront aux arguments que j'ai donnés soit en les ignorant, soit en cherchant quelque chose qui ne va pas avec eux. (Et, bien sûr, un philosophe qui cherche ce qui ne va pas avec un argument trouvera généralement un aspect de celui-ci qu'il considèrera comme une erreur. S'il ne trouve rien d'autre, il peut toujours souligner le fait que les prémisses de l'argument implique la conclusion et accuser l'argument de commettre une pétition de principe).

Ce livre ne s'adresse pas à ce genre de philosophes. Il s'adresse à ceux qui ont gardé un esprit ouvert sur la question du libre arbitre et du déterminisme et qui aimeraient voir quels sont les arguments. Je ne fais pas partie de ces gens qui considèrent que « garder un esprit ouvert » soit, *per se*, une vertu. Garder un esprit ouvert sur les théories raciales des Nazis ou sur l'astrologie ne serait pas une marque de vertu, intellectuelle ou morale. Mais, selon moi, le cours du débat sur le libre arbitre depuis le dix-septième siècle a été suffisamment source d'embarras, et la physique et la psychologie du vingtième siècle, tout comme les recherches sur le système nerveux humain, ont été suffisamment complexes, pour que garder un esprit ouvert sur le libre arbitre et le déterminisme soit consistant avec les réquisits de la vertu. Pourquoi, en ce cas, certains philosophes se sentent-ils contraint de réagir à la thèse de l'avant-dernier paragraphe en qualifiant ses partisans de « superstitieux », de « croyants en la magie » ou de « physiologistes en fauteuil » ?

Il n'y a peut-être pas de réponse générale à cette question. Mais je suis convaincu que dans un grand nombre de cas la réponse est que les gens qui considèrent ma thèse centrale comme simplement incroyable sont des victimes du scientisme. Le scientisme, dans mon usage du mot, est une sorte de respect indu pour la science – c'est-à-dire pour les sciences physiques et biologiques – et un mépris corrélatif pour tous les autres domaines où s'exerce l'entreprise intellectuelle humaine. Il n'est guère besoin de souligner que le scientisme est la première idéologie de notre temps. Il n'est guère besoin de souligner que les illusions engendrées par le scientisme sont si répandues et si insidieuses qu'il est pratiquement impossible d'obtenir que celui qui y est sujet envisage la possibilité que ce ne soient que des illusions¹⁷. (J'espère que la dénégation suivante n'est pas nécessaire : si je désapprouve le scientisme, je ne

¹⁷. J'utilise le mot *illusion* au sens de Freud (*Die Zukunft einer Illusion*, ch. 6) : la croyance de l'alchimiste que le plomb pouvait être transformé en or et la croyance de Christophe Colomb qu'un bateau du XV^e siècle pourrait atteindre une terre en naviguant vers l'ouest de l'Europe étaient des illusions, bien qu'elles fussent vraies. Pour un exemple d'esprit subtil pris dans les mailles du scientisme, lisez l'autobiographie de Carnap dans P.A. Schilpp, ed., *The Philosophy of Rudolf Carnap*, La Salle (Ill.), Open Court, 1963.

désapprouve pas pour autant la science. Nier que des honneurs divins soient dus à César n'est pas mettre en cause son commandement militaire.)

Mon propos n'est pas d'entrer dans une discussion étendue du scientisme, ni d'en offrir une analyse. Je souhaite seulement attirer l'attention sur deux effets du scientisme sur la pensée philosophique contemporaine.

Le premier de ces effets pourrait être décrit comme une révolte contre le holisme. Le raisonnement suivant est assez courant :

Un être humain est, en dernière analyse, un système d'atomes, une chose matérielle énormément complexe dont les propriétés surviennent sur les propriétés des atomes individuels qui en sont les parties, et sur les arrangements de ces atomes individuels. Mais le type d'indéterminisme que vous attribuez aux êtres humains ne pourrait pas survenir sur un système d'atomes. Dire que si toutes les parties du système sont déterminées le système entier est déterminé, ce n'est pas commettre le sophisme de la composition, car la détermination est clairement le genre de propriété qui se transfère des parties au tout. Il est vrai que le cas est compliqué par le fait que les atomes individuels *sont* indéterminés sous quelques aspects en raison de considérations quantiques. Mais la mécanique quantique n'a rien à voir avec le libre arbitre, et vous devez donc postuler une autre source d'indétermination, qui n'a aucune base dans le comportement des atomes et donc qui ne peut avoir aucune place dans une vision scientifique du monde.

Je ne sais vraiment pas quoi dire sur ce type d'argument, sinon qu'il doit y avoir quelque chose qui ne va pas avec lui. Il procède à partir de certaines prémisses (comme l'anti-holisme et la thèse que l'indétermination quantique n'a rien à voir avec le libre arbitre) et la plausibilité de ces prémisses doit être examinée. Je suis enclin à penser que les deux prémisses mentionnées sont plutôt plausibles. Je ne crois pas qu'il soit très vraisemblable que nous ayons des propriétés qui ne surviennent pas en *quelque* sens sur les propriétés des atomes dont nous sommes composés ; je ne suis certainement pas très tenté de penser que la mécanique quantique ait quoi que ce soit à voir avec le choix libre¹⁸. Mais, même si c'était là les deux seules prémisses de l'argument, et même si cet argument était parfaitement clair dans sa terminologie, et logiquement valide de façon indiscutable, je ne le trouverais pourtant pas convaincant. Ses prémisses peuvent être plausibles, mais je les considère comme bien moins plausibles que la proposition que nous sommes parfois moralement responsables pour nos actions et que la proposition que la règle (β) est valide. Et si l'argument de ce livre est correct, alors il suit de ces deux prémisses que quelque chose ne va pas dans celui que nous sommes en train de considérer.

J'aimerais savoir ce qui ne va pas avec cet argument, mais je l'ignore. Je n'ai jamais prétendu comprendre « comment marche le libre arbitre ». Si je le savais je vous le dirais, mais je l'ignore. Les questions soulevées par l'argument précédent sont des questions profondes auxquelles je n'ai pas de réponse, tout comme je n'ai pas de réponses aux questions dont j'ai avoué qu'elles me troublaient à la Section 4.4. Je n'ai pas d'affection pour les mystères en philosophie. Mais il n'est pas bon de prétendre qu'il n'y a pas de mystères s'ils existent de façon évidente. Bien plus, je préfère les petits mystères aux grands mystères. Une manière de considérer ce livre est d'y voir une tentative de présenter une position sur le libre

¹⁸. J'espère que le fait que j'ai discuté de la mécanique quantique assez longuement n'a pas donné l'impression opposée. Mon but dans cette discussion était simplement de montrer clairement que s'il a pu être raisonnable de croire que la science supportait le déterminisme (la thèse générale, pas la thèse que les êtres humains sont déterminés 'à tous égards pratiques'), ce n'est plus le cas. Par conséquent une raison que nos prédécesseurs pouvaient avoir de croire au déterminisme n'est pas une raison que nous avons.

arbitre qui engage ses adhérents à de plus petits mystères que ceux auxquels engagent toutes les positions rivales. Quiconque rejette les thèses centrales et les arguments de ce livre – j’inclus sous ce titre les thèses du présent chapitre, y compris la thèse que nous ne sommes pas déterminés – doit, s’il a une théorie élaborée du libre arbitre, du déterminisme et de la responsabilité morale, accepter au moins l’une des trois propositions suivantes :

La responsabilité morale n’existe pas

La responsabilité morale existe même si personne n’a jamais le choix de rien

Il y a des propositions vraies p et q telles que qui n’a pas le choix quant au fait que p soit vraie, ni quant au fait que si p est vraie, q l’est aussi, a pourtant le choix quant au fait que q soit vraie

Si l’une de ces propositions était vraie, cela serait, d’après moi, un très grand mystère en vérité, un mystère beaucoup plus grand que celui qui serait attaché à la vérité de

Nous avons des propriétés qui ne surviennent pas sur les propriétés des atomes dont nous sommes constitués

ou

Le libre arbitre implique une indétermination quantique

Ceux qui sont sous l’influence du scientisme ne le verront sans doute pas sous cet angle. Selon eux, aucun mystère ne saurait être plus grand que la vérité d’une proposition en conflit avec « la vision scientifique du monde ». Et la première de ces deux propositions au moins est en conflit avec cette vision du monde, puisqu’elle s’oppose à l’anti-holisme qui est une partie essentielle de la « vision scientifique du monde »¹⁹.

Ici, ces philosophes que je décris comme victime du scientisme et moi avons, je crois, atteint le fond ultime de la discussion. Nous n’avons rien de plus à nous dire ; ou, du moins, quoique nous puissions nous interpeler l’un l’autre, nous n’avons plus d’arguments.

Le second effet du scientisme sur certains philosophes est de produire chez eux une crainte profonde d’être laissés de côté par le progrès de la science²⁰. Je crois que de nombreux philosophes verraient dans les mots suivant une expression adéquate de leur pensée. « Supposons que j’accepte l’argument selon lequel, puisque la responsabilité morale existe, le déterminisme est faux, et supposons ensuite que les physiiciens décident que le déterminisme est vrai, après tout. Ou supposons que les neurophysiologistes, ou les psychologues cognitifs ou qui que ce soit nous donnent une raison vraiment bonne de croire que nous sommes des systèmes déterministes, dans un sens au moins aussi fort que le sens dans lequel les ordinateurs sont des systèmes déterministes. Je paraîtrais plutôt stupide, non ? On devrait rire de moi comme on rit de Hegel, qui avait déduit a priori l’existence d’exactly six planètes. Mon argument et moi-même devrions être mis dans le même sac que Kant et la géométrie euclidienne ou Paley et l’Argument du Dessein. Cela ne me va pas ».

Je ne suis peut-être pas charitable ; les mots que j’ai attribués, ou presque attribués, à de « nombreux philosophes », pourraient sans doute être remplacés par des formules plus honorables. Mais je crois que les idées exprimées dans cette citation imaginaire ont plus

¹⁹. C’est bien entendu un terme tendancieux. J’aimerais me décrire comme soutenant une « conception scientifique du monde », si le terme n’avait pas été associée avec une certaine métaphysique (qui n’a pas de connexion avec les résultats de la science) que je rejette.

²⁰. On trouve parfois une autre crainte mêlée à celle d’être laissé de côté : la crainte d’être jeté à la poubelle. C’est la crainte que le progrès de la science – ou peut-être un changement d’attitudes sociales attribuable au progrès de la science – fera que ses conceptions ne seront pas tant réfutées qu’irréremédiablement démodées.

d'influence que l'on a bien voulu le reconnaître en général, et je ne considère pas que cette influence soit salutaire. Alors qu'il serait stupide, pour des philosophes, ou pour quiconque, y compris les scientifiques, de se mettre à prédire l'évolution à venir des sciences, je ne vois pas de raison pour que des philosophes se retiennent de soutenir des positions qui impliquent que la science ne sera jamais capable de faire ceci ou cela. De telles théories seraient empiriquement réfutables, bien sûr, mais je ne vois pas de raison de considérer que c'est une mauvaise chose. Si une théorie philosophique est empiriquement réfutable et se trouve empiriquement réfutée, alors il y aura une théorie philosophique en moins, et ce ne serait pas une mauvaise chose.

Certains philosophes peuvent dire qu'une théorie qui se dénomme elle-même philosophique ne devrait pas être empiriquement réfutable, puisque toute théorie empiriquement réfutable est « automatiquement » non pas de la philosophie mais de la science. Selon cette conception, Kant et Paley – je passe sur la blague à propos de Hegel²¹ – quand ils discutaient sur la géométrie et la finalité dans la nature, étaient en train de faire de la science (de la très mauvaise science, en l'occurrence) plutôt que de la philosophie. Selon cette conception, je suis moi-même coupable de faire passer de la science pour de la philosophie. Cette thèse méta-philosophique me semble n'avoir guère de raison en sa faveur. Les théories et les arguments de ce livre sont philosophiques de manière paradigmatique, quoi qu'ils puissent être d'autre. Si quelqu'un veut dire qu'ils sont aussi « scientifiques », cela m'est indifférent. Bien sûr, je ne voudrais pas que quiconque dise que mes théories et mes arguments sont de la *mauvaise* science. Sauf s'il est possible pour la mauvaise science d'être de la bonne philosophie, ce qui me semble hautement improbable. Je pense que ce que j'ai produit est de la bonne philosophie, et ce que j'aimerais voir de la part de mes adversaires, ce sont des *arguments* en faveur de la conclusion qu'il s'agit de mauvaise philosophie.

Malgré ce que j'ai dit, il y a une question que me poseront très vraisemblablement les philosophes qui pensent que j'ai outrepassé les limites de la philosophie et dont je crois qu'il s'agit d'une bonne question, qui m'oblige à répondre. On pourrait la formuler ainsi :

Très bien. Vous admettez que votre théorie est en principe réfutable empiriquement. Si la physique devait un jour devenir plus déterministe, ou si les sciences empiriques de la nature nous donnaient de vraiment bonnes raisons de croire qu'un être humain est un système déterministe, alors (vous le concédez) votre rejet du déterminisme serait rendu intenable par la science. Eh bien, supposons que cela *soit arrivé* malgré votre prédiction que cela n'arrivera pas. Que diriez-vous alors ?

Je ne suis pas tout à fait sûr de ce que je dirais, mais je crois que je dirais que (β) était finalement invalide. (Je vais donner les raisons de mon hésitation sur ce point.) Cette réponse n'est pas purement *ad hoc* ; je n'ai pas simplement sélectionné une des réactions possibles à la validation scientifique du déterminisme pour embrasser ensuite arbitrairement cette réaction. Considérez la liste de propositions suivante :

- (1) Nous sommes parfois moralement responsables des conséquences de nos actes ;
- (2) La validité de (β) implique que notre possession du libre arbitre implique l'indéterminisme (argument de la conséquence 3^e version)
- (3) Si (1) est vraie, alors nous avons le libre arbitre
- (3a) Nous avons le libre arbitre ;
- (4) (β) est valide ;

²¹. Voyez Bertrand Beaumont, « Hegel and the Seven Planets », *Mind* (1954).

(4a) Notre possession du libre arbitre implique l'indéterminisme ;

(5) L'indéterminisme est vrai

Dans ce livre j'ai argumenté en faveur de chacune de ces propositions, ou du moins, je les ai soutenues. Pourtant, je ne les tiens pas pour toutes aussi bien établies les unes que les autres. Bien qu'elles soient toutes très vraisemblables, selon moi, au sens où je regarde leur conjonction comme très vraisemblable, et *a fortiori* chacune d'elle comme très vraisemblable, certaines me semblent pour ainsi dire *préférables* à d'autres. (Disons qu'une personne tient *p* pour préférable à *q* si (i) elle croit que la vérité de la conjonction de *p* et *q* est beaucoup plus vraisemblable que sa fausseté et (ii) elle croit que si l'une ou l'autre est finalement fautive, alors *p* et vrai et *q* est fautive.) J'ai indiqué mes préférences entre ces propositions par leur numérotation : je tiens (1) pour préférable à (2), (2) pour préférable à (3), etc. (Mais ne considère ni (3) ni (3a) comme préférable à l'autre, et ni (4) ni (4a) comme préférable à l'autre.) Pourquoi ai-je cet ensemble de préférences ? Eh bien, certaines contraintes purement formelles jouent un rôle. Par exemple, puisque (1) et (3) impliquent (3a), et puisque je le sais, il serait bizarre que je tiens les deux pour préférables à (3a). Mais, en gros, c'est juste l'ordre de préférences que je me trouve avoir. Je ne sais pas comment le défendre, même si je considérerais soigneusement tout argument qui voudrait établir que je devrais ordonner mes préférences différemment. Je ne crois pas, bien entendu, que mon ordre de préférences soit entièrement arbitraire. Je crois que mon idée est qu'il se fonde sur plusieurs intuitions conceptuelles et que si vous ne partagez pas mes préférences, vous ne partagez pas mes intuitions.

Cela ne m'embarasse pas de l'admettre. Les positions de n'importe quel philosophe, aussi argumentées soient-elles, doivent en dernière instance se fonder sur des propositions qui lui *semblent* simplement vraies. Il n'y a pas d'échappatoire ; c'est une simple conséquence du fait que tout argument à un nombre d'étapes fini. Bien plus, tout philosophe doit penser que ceux de ses collègues dont les « prémisses ultimes » sont incompatibles avec les siennes ou bien n'ont pas considéré ces prémisses avec suffisamment de soin, ou bien manquent d'intuition sur la manière d'être des choses. Que peut-il penser d'autre ? S'il croit que les prémisses ultimes de ceux qui sont en désaccord avec lui sont « tout aussi bonnes » que les siennes, alors il devra admettre qu'il n'a pas de raison de s'en tenir aux siennes, et devra, pour le moins, faire retraite vers une position où il n'accepte rien de plus fort que la disjonction de ses premières prémisses et de toutes les propositions qui sont « tout aussi bonnes ».

Un point méta-philosophique de plus. Etant donné qu'un philosophe doit avoir un ensemble de « prémisses ultimes », il semblerait qu'il n'y a pas de raison de supposer qu'il doive considérer qu'aucune n'est préférable aux autres. Il n'y a pas de raison pour qu'un philosophe ne puisse pas dire quelque chose comme « les propositions A et B constituent à présent le fondement de ma philosophie ; je n'ai pas d'argument en faveur de l'une ou de l'autre, mais je crois que je peux tout simplement *voir* qu'elles sont vraies. Pourtant, si l'une d'elle est fautive, alors c'est B ».

Etant donné que la liste ci-dessus représente l'ordre de mes préférences, il m'est facile de dire ce que je ferais si la science devait me donner une raison indubitable de croire au déterminisme : je devrais rejeter (5). Mais (5) est impliqué par (3a) et (4a), donc je devrais rejeter (3a) ou (4a). Puisque je préfère (3a) à (4a), je devrais donc rejeter (4a). Mais (4a) est impliquée par (2) et (4), donc je devrais rejeter (2) ou (4). Puisque je préfère (2) à (4), je devrais donc rejeter (β). Et cela semblerait devoir clore le débat. (Dans ce raisonnement, je présuppose bien sûr qu'une « raison indubitable de croire au déterminisme » ne demanderait pas de moi que j'accepte sans preuve une proposition à laquelle je préférerais (β)).

J'ai défendu (β) sur une base entièrement *a priori*. Mais je ne serais pas trop surpris si je

découvrais que cette proposition, qui me semble à présent être une vérité de raison, avait été réfutée par le progrès de la science. De telles réfutations se sont produites souvent. Et il ne suit pas du fait qu'elles se sont produites qu'il y ait quoi que ce soit d'erroné dans le fait d'accepter sur une base *a priori* un principe qui se révèle plus tard être empiriquement réfutable. On doit simplement se rendre compte que des convictions *a priori* sont aussi corrigibles que d'autres.

J'ai exprimé plus haut une certaine hésitation à dire que je devrais rejeter (β) si la science devait me forcer à accepter le déterminisme. J'hésite parce que cette supposition contrefactuelle est décrite de manière très abstraite. Je souhaite laisser ouverte au moins la pure possibilité qu'il y ait une façon dont la science pourrait me forcer à accepter le déterminisme, qui me forcerait aussi à modifier l'ordre de mes préférences, exprimé dans la liste précédente. Je ne considère pas que ce soit très vraisemblable mais qui peut être certain du futur ?

6.5 Dans la section finale de ce livre, je vais résumer son argument central.

La thèse du libre arbitre est la thèse que, très souvent, quand nous avons à choisir entre des lignes de conduite incompatibles, chacune est telle qu'il est en notre pouvoir de la choisir. Ou, dit de manière plus idiomatique, la thèse du libre arbitre est la thèse que nous avons très souvent le choix sur ce que nous allons faire.

Le déterminisme est la thèse que, étant donnés le passé et les lois de la nature, il n'y a qu'un futur possible.

La thèse du libre arbitre et le déterminisme sont incompatibles. Autrement dit, l'*incompatibilisme* est vrai. On peut le voir sur la base de plusieurs arguments détaillés, qui sont tous des élaborations de l'argument simple suivant (l'Argument de la Conséquence) :

Si le déterminisme est vrai, alors nos actes sont les conséquences des lois de la nature et d'événements dans le passé éloigné. Mais ni ce qui s'est passé avant notre naissance, ni ce que sont les lois de la nature ne dépendent de nous. Par conséquent, les conséquences de ces choses (y compris nos actes présents) ne dépendent pas de nous.

L'un des ces trois arguments détaillés – le Troisième Argument du chapitre III – est un instrument particulièrement utile dans la discussion du compatibilisme parce qu'il isole, comme une prémisse séparée, la thèse la plus douteuse que doit accepter l'incompatibiliste. Cette thèse est la validité de la règle d'inférence suivante :

(β) p et nul n'a, ni n'a jamais eu, le choix quant à p

Si p alors q , et nul n'a ni n'a jamais eu le choix quant au fait que si p alors q

Donc, q et nul n'a, ni n'a jamais eu, le choix quant à q

La proposition que cette règle est valide, quand on la considère attentivement, dispose l'esprit à l'assentiment ; elle semble être une vérité de raison. Mais, malgré tout, elle peut être fausse. Pourtant, elle semble plus vraisemblable que certaines prémisses du compatibiliste. En particulier, elle semble plus vraisemblable que toute analyse conditionnelle de la capacité qui est suffisamment complexe pour pouvoir « rendre compte » (*handle*) des contre-exemples bien connus aux analyses conditionnelles plus simples. Il en est ainsi tout simplement parce que de telles analyses sont trop complexes pour être saisies comme des tous et ne peuvent donc même pas *sembler* être intuitivement évidentes.

Nous devrions donc être incompatibiliste, à moins qu'il n'y ait une objection plausible à l'incompatibilisme, quelque chose, qui plus est, qui le soit *davantage* que le rejet de la règle d'inférence (β) apparemment évidente d'elle-même. Et il y a bien une objection à

l'incompatibilisme. L'incompatibiliste doit ou bien rejeter la thèse du libre arbitre, ou bien accepter la thèse qu'un agent à la souvent le choix quant au fait que l'un de ses états internes est suivi d'un acte externe (l'état étant la cause de l'acte), même si ni l'état interne ni rien d'autre ne *détermine* que l'acte se produise²². Je trouve chacune de ces alternatives sujette à objection, mais la seconde me semble bien *moins* sujette à objection que le rejet de (β). Je suggère donc d'accepter l'incompatibilisme.

Si l'incompatibilisme est vrai, alors ou bien le déterminisme ou bien la thèse du libre arbitre est fausse. Nier la thèse du libre arbitre c'est nier l'existence de la responsabilité morale, ce qui serait absurde. En outre, il semble n'y avoir aucune bonne raison d'accepter le déterminisme (qui, on doit le rappeler, n'est *pas* la même thèse que le Principe de Causalité Universelle). Par conséquent, nous devrions rejeter le déterminisme.

Cette conclusion est, au moins en principe, ouverte à la réfutation scientifique, puisqu'il est concevable que la science nous présente un jour des raisons contraignantes de croire au déterminisme. Alors, et seulement alors, devrions-nous selon moi devenir compatibilistes, car, dans le cas imaginé, la science aura *par hypothèse* montré que quelque chose que j'ai soutenu par des arguments est faux, et les plus faibles des mes arguments sont ceux en faveur de l'incompatibilisme.

²². Ce dilemme, à strictement parler, ne se pose qu'à l'incompatibiliste qui accepte notre second modèle de l'action. Mais un dilemme semblable se posera à tout incompatibiliste qui soutient une théorie, quelle qu'elle soit, très explicite de l'action.