

## *Introduction*

### LES SCIENCES COGNITIVES : UN TOUR D'HORIZON

*Daniel Andler*

Ce chapitre a pour objectif de fournir au nouveau-venu quelques repères dans le paysage complexe des sciences cognitives : repères dans le temps, repères dans l'espace des disciplines et programmes de recherche actifs aujourd'hui.

Remarquons qu'il n'est pas strictement nécessaire de pouvoir se repérer dans ce paysage, lorsqu'on débute, ni même du reste lorsqu'on est plus avancé : la spécialisation est le lot de tout chercheur. Pensons aux mathématiques, ou bien aux sciences sociales, ou encore aux sciences de la vie : la grande majorité des étudiants, et bien des chercheurs, sont largement ignorants de l'ensemble du champ dont relèvent leurs études ou leurs recherches. Mais d'abord cette réalité n'est pas nécessairement une bonne chose : avoir une petite idée de ce qui s'étend au-delà de son village peut se révéler précieux le jour où l'on veut ou l'on doit s'en éloigner, ou même si l'on envisage seulement de se promener ou encore de commercer avec les peuplades voisines. Ensuite, et surtout, les sciences cognitives ont deux caractéristiques qui rendent un repérage plus important que dans les autres domaines mentionnés : elles sont intrinsèquement interdisciplinaires, et elles sont jeunes.

L'*interdisciplinarité* fait que beaucoup de notions et de résultats ne prennent tout leur sens que dans le cadre de rapports étroits avec des notions et des résultats relevant d'autres secteurs. Pour prendre un exemple, quand on parle des calculs (ou computations) qu'effectue le cerveau, de modèle computationnel de telle ou telle capacité, ou encore de codage neural, il est utile et peut-être même indispensable, à un moment ou un autre, de savoir ce que calcul (ou computation) et codage veulent dire, dans le contexte où ces notions ont été développées, à savoir la logique et la théorie de l'information. De même, lorsqu'on parle, en psychologie du développement, du caractère inné d'une capacité, ou de la fonction d'un mécanisme, ou d'un dysfonctionnement, il peut être bon de savoir comment les notions de fonction, d'innéité ou d'adaptation sont comprises par la biologie et en particulier par la biologie évolutive. Le lecteur trouvera beaucoup d'autres exemples au fil des chapitres, expliqués de manière moins allusive, et qui montreront en outre que les théories et les modèles proposés dans un secteur de la recherche sont soumis à des contraintes issues d'autres secteurs, un peu comme une théorie mécanique du moteur à explosion doit être compatible avec le second théorème de la thermodynamique, ou comme un théorème d'optique géométrique doit s'inscrire dans le cadre de l'optique physique (mais c'est généralement plus compliqué que ne le suggèrent ces exemples !).

La *jeunesse* des sciences cognitives, seconde particularité, fait qu'elles sont plus instables que les autres champs scientifiques, en tout cas ceux qui relèvent des sciences de la nature, alors même qu'elles aspirent au statut, et à la stabilité, de ces sciences. Aucune science vivante n'est parfaitement stable, ni parfaitement consensuelle : les révisions, mineures le

plus souvent, exceptionnellement plus radicales, et les controverses sont le sort de toute discipline. Mais les branches mûres possèdent un noyau pratiquement invariant, et les controverses se limitent sauf exception à certaines questions relativement circonscrites. L'instabilité et la controverse sont en proportion sensiblement plus importantes dans les sciences cognitives. Ce qui ne s'explique pas seulement par leur jeunesse (comme on le comprend sans peine) mais sans doute aussi par l'extraordinaire difficulté de leur entreprise. Ce n'est pas un hasard si l'objet composite esprit/cerveau est le dernier à être abordé de front par la science. On peut penser qu'il constitue le défi le plus difficile qu'elle ait jamais relevé. Toujours est-il que les turbulences du passé laissent des traces plus visibles dans les sciences cognitives actuelles que, mettons, dans la physique, la chimie ou la géologie, et que les controverses y sont plus vives. Il y a pourtant de l'ordre dans ce désordre apparent, et il importe de faire voir concrètement à qui aborde le domaine qu'il n'est pas un simple forum livré aux querelles d'opinion : il est pour l'essentiel, par-delà les incertitudes, une entreprise progressiste et cumulative ; l'hésitation porte surtout sur la meilleure manière d'organiser une énorme moisson de résultats solides, provenant de cadres théoriques nécessairement différents car relatifs à des niveaux de description et des ordres de phénomènes distincts et qui sont néanmoins en attente d'intégration.

Enfin, l'émergence et le déploiement des sciences cognitives sont une belle aventure, pleine d'enseignements, qu'on peut vouloir explorer pour le simple plaisir intellectuel... Notre objet n'est cependant pas celui de l'historien des sciences, qui se place, en bonne méthodologie, *en amont* et cherche à comprendre comment on est passé d'un état de la science à l'état suivant, et ainsi de suite jusqu'à l'état présent, sans supposer que toute la trajectoire soit guidée par son point terminal : l'historien se garde de ce que Raymond Aron appelait l'illusion rétrospective de fatalité. Nous partirons au contraire de l'état présent, notre objectif étant de répondre à la question que se posent naturellement les étudiants d'aujourd'hui : pourquoi voit-on aujourd'hui les choses comme nous les voyons, et comme ne pouvaient les voir nos prédécesseurs ? Parler du projet des sciences cognitives dans les termes actuels est une chose, parler des obstacles et des manques qui ont gêné sa réalisation par le passé en est une autre, qu'il faut comprendre comme un artifice d'exposition : nos prédécesseurs n'avaient pas ce projet en tête, ils étaient occupés, comme nous le sommes aujourd'hui, par leur programme de recherche, formulé dans les termes qui faisaient sens à leur époque<sup>1</sup>.

---

1. On touche là à un problème central de la méthodologie historique : comprendre le passé à la lumière du présent, c'est pratiquer une conception dite « whiggish » ou rétrospective de l'histoire, conception abondamment critiquée par les historiens eux-mêmes, mais qui, chassée par la porte rentre régulièrement par la fenêtre. Nous ne pouvons entrer ici dans le débat, que je voulais signaler aussi pour indiquer que les sciences de la nature ne sont pas les seules à avoir des exigences de rigueur. Un point important est que l'histoire whiggish ne consiste pas nécessairement à assigner au passé le « but » de conduire au présent : son biais principal (peut-être inévitable) est d'interpréter l'état passé des connaissances dans les termes d'aujourd'hui.

# I. UN PROJET ANCIEN QUI S'EST LONGTEMPS HEURTE A DES OBSTACLES CONCEPTUELS

## 1. *Le projet*

L'esprit humain est l'objet d'une fascination très ancienne. En témoignent toute l'histoire de la psychologie, et toute celle de la philosophie, histoires qui sont du reste longtemps restées inextricablement liées<sup>2</sup> – et même tout bonnement confondues, si l'on adopte une perspective qu'on peut appeler *psychologisme*, selon laquelle l'esprit, comme produit de processus mentaux, n'a pas d'autre réalité que ces processus eux-mêmes. Point de vue que ne partagent pas ceux (nombre de philosophes notamment) qui estiment qu'on peut et qu'on doit s'intéresser aux produits (pensées, concepts, langage, images, décisions, etc.) sans nécessairement savoir quels mécanismes les produisent chez les êtres humains. On peut en tout cas reconnaître que, d'une part, nous avons appris beaucoup de choses sur les pensées, les concepts, etc., avant d'avoir la moindre idée solide sur les mécanismes qui nous permettent d'entretenir des pensées, d'appliquer des concepts, etc. ; et que, d'autre part, cette connaissance a de bonnes chances de nous servir dans notre quête des mécanismes en question. À quoi il faut ajouter qu'inversement, toute information sur les mécanismes doit être prise en compte dans nos conceptions des produits : pensées, concepts, etc. La manière la plus concise et la moins partisane de caractériser le projet des sciences cognitives est précisément celui-là : rendre compte à la fois des capacités mentales et des mécanismes de l'organisme humain (voire d'autres organismes) en vertu desquels les humains (et parfois d'autres organismes) sont dotés de ces capacités et mécanismes.

Ce projet est celui d'une science coordonnée de l'esprit et du cerveau, puisque les mécanismes recherchés sont localisés dans le cerveau, du moins dans une très large mesure<sup>3</sup> – le sens précis de cette « coordination », et, partant, la place respective de l'esprit et du cerveau, restant largement ouverts au débat. On peut aussi parler d'une science *naturelle* de l'esprit, par opposition à une science abstraite de ses produits. Or, si l'ambition de développer une telle science était présente depuis des siècles, au moins à partir de la philosophie mécaniste cartésienne (XVII<sup>e</sup> siècle), au XVIII<sup>e</sup> chez les philosophes empiristes, au XIX<sup>e</sup> chez les premiers neurologues, les moyens de le faire ne sont apparus qu'au premier tiers du XX<sup>e</sup> siècle. Nous verrons dans un instant ce que sont ces moyens, mais il faut dire quelques mots de ce qui existait avant qu'ils ne soient découverts.

## 2. *L'état des lieux avant l'émergence des sciences cognitives*

---

2. La psychologie comme discipline académique ne s'est séparée institutionnellement de la philosophie qu'à la charnière des XIX<sup>e</sup> et XX<sup>e</sup> siècles.

3. Je dirai « cerveau » pour simplifier, étant entendu que c'est tout le système nerveux central qui est impliqué dans la « production » des phénomènes mentaux, sans oublier le système nerveux entérique, dont le rôle est si important qu'on l'appelle parfois le second cerveau.

Puisque l'histoire de la psychologie est liée à celle de la philosophie, et que cette dernière remonte, dans la tradition occidentale, à plus de deux millénaires ; puisque la psychologie possède son Olympe (d'Aristote (384-322 av. J.C.), Wolff (1679-1754), Bonnet (1720-1793) ou Kant (1724-1804) <sup>4</sup> à Asch (1907-1996) en passant, chronologiquement, par Fechner (1801-1887), Bain (1818-1903), Helmholtz (1821-1894), Wundt (1832-1920), James (1842-1910), Pavlov (1849-1936), Freud (1856-1939), Binet (1857-1911), Titchener (1867-1927), Claparède (1873-1940), Jung (1875-1961), Watson (1878-1958), Wallon (1879-1962), Köhler (1887-1967), Lewin (1890-1947), Vygotski (1896-1934), Piaget (1896-1980), Skinner (1904-1990)<sup>5</sup>), et que la philosophie n'est pas en reste ; pour toutes ces raisons il semble peu vraisemblable que ces brillants esprits, animés généralement par l'esprit scientifique, n'aient pas anticipé les découvertes qui vont suivre ; et plus généralement, qu'il n'y ait aucune continuité importante entre la préhistoire et l'histoire des sciences cognitives. De fait, comme toute révolution, la révolution cognitive doit beaucoup à ce qui l'a précédée, et innove moins que ses hérauts ont parfois voulu faire croire. Pour nous, néanmoins, qui ne sommes pas chargés de distribuer les bons points ni de retracer avec scrupule la voie des idées, l'important est bien de prendre la mesure du saut accompli. Une chose semble assez claire : avec le temps, on accède à des données empiriques nouvelles, qui excluent certaines spéculations et mettent sur la voie de nouvelles théories plausibles. En ce sens, si ingénieux qu'aient été nos prédécesseurs, ils souffraient de ce qui apparaît aujourd'hui comme un déficit empirique.

Qu'avaient donc construit, au fil des siècles, ces grands esprits, et qu'on peut considérer rétrospectivement comme le point de départ des sciences cognitives ? En schématisant beaucoup, on peut distinguer quatre grands corpus, d'ampleur inégale.

#### a. Psychologie non empirique

C'est le plus important de ces corpus, en volume, et le plus hétérogène sur le plan des résultats et des méthodes. Il comprend deux volets. Le premier consiste en une psychologie philosophique, anecdotique, historique, littéraire, appuyée sur l'observation quotidienne, l'introspection, la littérature, les témoignages des chroniqueurs, biographes, essayistes, moralistes, journalistes, voyageurs, médecins, militaires..., bref, de tous ceux qui observent

---

4. Kant est certes l'un des grands philosophes, mais il est aussi l'auteur de ce qu'on peut considérer comme le premier manuel de psychologie au sens moderne du terme.

5. Chaque historien de la psychologie a sa propre liste de grands, « pères » de ceci ou de cela – celle-ci ne prétend pas constituer mon panthéon personnel, il ne s'agit que de fournir quelques repères. Remarquons au passage l'absence de femmes, qui serait totale même au second rang si l'on excluait la psychanalyse ; la situation a évidemment changé à partir de la seconde moitié du XX<sup>e</sup> siècle. Ma liste est donc en partie arbitraire, mais pas complètement. En particulier, elle fournit implicitement une périodisation : tous les auteurs énumérés après Asch sont nés au XIX<sup>e</sup> siècle, à quelques années près dans le cas de Skinner et d'Asch lui-même : la psychologie scientifique, au sens où nous l'entendons aujourd'hui, est née au XVIII<sup>e</sup> siècle (que représentent les trois auteurs suivant le nom d'Aristote), mais a pris véritablement son essor au XIX<sup>e</sup>. Le XX<sup>e</sup> se divise, grossièrement, en deux moitiés : la première est dominée (sans s'y ramener entièrement) par le béhaviorisme (Watson, Skinner...) aux États-Unis, seul pays où il joue un rôle essentiel, la psychologie de la Gestalt (Köhler, Lewin...) en Allemagne et la psychanalyse (Freud, Jung...) un peu partout ; la seconde, selon une lecture qui ne fait d'ailleurs pas l'unanimité, par les sciences cognitives, dont les fondateurs, sans exception, sont nés au cours de ce siècle ; l'un des plus importants, le Canadien Donald Hebb (1904-1985), a joué un rôle fondamental dans la transition du béhaviorisme aux sciences cognitives.

les êtres humains cheminer au travers des épisodes, infimes ou décisifs, de leur existence, et cherchent à en tirer des enseignements de portée plus ou moins générale. On peut penser que ce corpus n'a aucun rôle à jouer dans la psychologie scientifique, mais ce serait une erreur : il constitue un réservoir de concepts et d'hypothèses ; et il contient les manifestations de ce qui est probablement une capacité essentielle de notre appareil cognitif, la « psychologie naïve » ou « théorie de l'esprit » (cf. chap. XIII « Émotions », et chap. IV « Le développement »). Il joue un peu<sup>6</sup> le rôle que le savoir des naturalistes a joué pour la biologie naissante, et continue du reste de jouer sous des formes plus scientifiques. À côté de ce premier volet, de nature descriptive, la philosophie a développé un considérable volet normatif, portant sur la structure logique des pensées et sur leur bon enchaînement : la majeure partie de la tradition rationaliste occidentale est consacrée à cette tâche. Plus rigoureux, systématique et unifié que le premier, ce second volet n'est cependant pas scientifique, dans la mesure où il repose sur le jugement des philosophes, sans contribution de résultats empiriques. Sa jonction avec les sciences empiriques sera précisément l'un des objectifs des sciences cognitives.

#### b. Médecine et psychiatrie

Centrales déjà chez Aristote, relayées au XIX<sup>e</sup> siècle par la neuro-anatomie et la neurophysiologie (Flourens, Gall, Broca, Wernicke...), médecine et psychiatrie ont érigé une connaissance empirique et médicale du système nerveux, et jeté les bases d'un naturalisme résolu dans l'étude de l'humain, en particulier parce que leur perspective embrasse l'animal, considéré comme proche de l'homme. Ce qu'on appelle aujourd'hui les neurosciences s'enracine dans un ensemble riche, quoique fragmenté de spécialités aux objectifs divergents.

#### c. Psychophysique

La perception, surtout dans ses modalités visuelle et auditive, donne au contraire naissance, à partir du XVIII<sup>e</sup> siècle, à une science au plein sens du terme, qui prend bientôt le nom de psychophysique, et constitue la première branche de la psychologie scientifique, sur laquelle pourra se greffer sans solution de continuité le segment correspondant des sciences cognitives (cf. chap. VI « La perception visuelle »). Une école, celle de la Gestalt, part de la perception, dont elle met au jour des propriétés insoupçonnées, pour aller vers l'intuition intellectuelle (celle de la solution soudaine d'un problème).

#### d. Linguistique

Enfin, le langage donne lieu depuis l'Antiquité, dans les traditions babylonienne, grecque, hindoue, chinoise, hébraïque, arabe, etc., à une enquête systématique, qui se développe au Moyen Âge sous le nom de « grammaire » et prend son essor au XVIII<sup>e</sup> siècle ; elle devient une discipline à part entière, la linguistique, au siècle suivant. En raison même de sa richesse et de sa solidité, elle résistera longtemps (et résiste encore pour une part) à l'incorporation dans les sciences cognitives : le langage semble à beaucoup de chercheurs pouvoir être étudié

---

6. L'analogie est imparfaite, dans la mesure où le savoir des philosophes, moralistes et autres n'avait pas (et n'a toujours pas) atteint le degré de systématisme que les naturalistes avaient su donner, au cours des XVIII<sup>e</sup> et XIX<sup>e</sup> siècles, à la biologie naissante, et qui conduira directement à la théorie de l'évolution de Darwin. Mais le vivant est un sujet apparemment plus accessible que le mental, et l'analogie vaut si l'on tient compte de cette différence.

de manière profonde et détaillée sans référence aux mécanismes qui permettent aux humains de l'employer. Il se présente sous les dehors d'un système abstrait, comparable à celui des nombres ou des objets mathématiques, qui peut (et selon certains) doit être étudié en lui-même. L'articulation avec les sciences cognitives sera proposée par Noam Chomsky, qui est, en partie pour cette raison, considéré comme l'un des pères, peut-être le plus important, de la révolution cognitive (cf. chap. VIII « Le langage »).

À mesure qu'elle se familiarisera avec les recherches présentées au fil des chapitres suivants, la lectrice pourra se faire une idée plus précise de la distance qui les sépare des quatre corpus qui viennent d'être sommairement présentés : distance grande, voire considérable, mais certainement pas infinie. Interrogeons-nous maintenant sur ce qui nous apparaît rétrospectivement comme des obstacles qui ont empêché les philosophes, logiciens, psychologues, linguistes, médecins et biologistes de s'engager ensemble sur le chemin des sciences cognitives.

### 3. Des obstacles

Par « obstacles » il ne faut entendre ici rien d'autre que certaines attitudes intellectuelles assez courantes avant l'émergence des sciences cognitives, dont il est probable qu'elles gênaient l'accès à leurs intuitions de base, d'autant qu'elles sont encore actives aujourd'hui et tendent toujours à s'y opposer. Je ne prétends pas en dresser une liste exhaustive, ni les considérer comme des tares dont il fallait à tout prix, selon une « logique de l'histoire », se débarrasser. Notre lecture, redisons-le, est volontairement rétrospective et à visée pédagogique.

Je distingue les obstacles des manques, et je ne parle ici que des premiers, la section suivante étant consacrée aux seconds. Par « obstacle » j'entends ce qui empêche de voir ce qui manque, d'imaginer que cela puisse exister, et ainsi de se mettre en quête de le trouver. J'en vois trois principaux, qui tous ont trait au statut de l'esprit.

Le premier qui vient... à l'esprit est le dualisme, la conviction que l'esprit appartient à une sphère de réalité différente de la sphère de la matière. Érigée, notamment par Descartes<sup>7</sup>, en dogme philosophique, cette thèse semble exclure la possibilité d'une science *naturelle* de l'esprit, si on entend par là une explication des phénomènes mentaux comme résultant de causes mécaniques ou physico-chimiques. Cela ne rend pas pour autant inconcevable une science rigoureuse de l'esprit, appuyée sur l'expérience : la psychologie scientifique du XVIII<sup>e</sup> siècle, par exemple, a pu se développer dans un cadre dualiste ; et comme on le verra, une partie importante des travaux en sciences cognitives peut s'accommoder d'une certaine forme de dualisme (ce que certains d'ailleurs lui reprochent). Il n'empêche que le dualisme interdit de considérer le cerveau, ou le système nerveux central, comme la source déterminante des phénomènes mentaux : pour le dire (trop) simplement, on ne peut être dualiste et considérer que « le cerveau secrète la pensée comme le foie secrète la bile », comme le proclamait (la formule est célèbre) le médecin et philosophe français Cabanis en 1802.

---

7. La position de Descartes est beaucoup plus complexe que ce qu'on en retient généralement dans ce contexte, mais un dualisme cartésien simplifié suffit ici pour caractériser l'obstacle en question.

Ce dualisme (qu'on appelle parfois métaphysique ou ontologique, du fait qu'il introduit au sein de ce qui existe une scission entre deux types d'entités) constitue donc un premier obstacle : il bloque la voie d'une enquête sur la façon dont le cerveau peut engendrer les phénomènes mentaux, leur conférer l'existence et leur donner leurs caractéristiques spécifiques. C'est une autre forme de dualisme qui constitue le deuxième obstacle : il consiste à considérer que l'esprit, même s'il possède une base matérielle, même s'il est *brain-based*, sous-tendu par le cerveau, ne doit son *contenu* à rien qui réside dans l'individu, mais à ce que lui fournit l'environnement au cours du développement et de l'expérience quotidienne de l'individu. L'esprit serait une tablette de cire, selon la métaphore qui parcourt l'histoire de la philosophie depuis Platon. La cire est matérielle, et sans la tablette il n'y aurait pas d'esprit, pas de processus mentaux. Mais ce qui compte, ce qui fait qu'une pensée est ce qu'elle est et se distingue d'une autre, ce qui fait que les pensées s'enchaînent d'une manière et pas d'une autre, ce qui fait finalement que l'esprit de Manon est le sien et pas celui de Mohammed, tout cela est dû à ce qui s'inscrit sur la tablette de Manon et non (identiquement) sur celle de Mohammed.

À ce point s'ouvrent deux voies. La première est largement compatible avec le projet d'une science coordonnée de l'esprit et du cerveau : les traces sur la cire sont l'empreinte mnésique des sensations, des expériences, et ce qui est rejeté est seulement l'idée que l'esprit possède une nature, une forme déterminée, donnée antérieurement à toute expérience. La querelle entre innéisme (ou dit parfois dans ce contexte rationalisme) et empirisme n'a pas cessé depuis le XVII<sup>e</sup> siècle et se prolonge de nos jours dans les sciences cognitives et en particulier la philosophie de l'esprit. La seconde voie diverge en revanche fortement de cette perspective. Elle a non seulement constitué un obstacle à l'essor des sciences cognitives, mais elle est encore aujourd'hui la source d'une forte prévention à leur égard. L'idée bloquante ici est que c'est la société, la culture, l'histoire, la langue qui déterminent le contenu des pensées, les inscriptions sur la tablette. Elle est bloquante parce qu'elle s'accompagne de l'idée que ces choses échappent aux déterminations naturelles : elles possèdent leur autonomie, comme le prouve notamment (selon ce point de vue) leur variabilité illimitée, même si elles ont pour siège les organismes biologiques, donc matériels, que nous sommes. Qui s'intéresse aux pensées des êtres humains, en tout cas des pensées qui vont au-delà des perceptions et des besoins élémentaires, et s'agissant d'individus normaux dans leur état normal, doit faire complètement abstraction de leur constitution biologique, en particulier cérébrale. Cette position conduit à un dualisme épistémologique qui dresse une frontière infranchissable entre les sciences de la nature (dont les neurosciences) et les sciences de l'esprit (ce qu'on appelle aujourd'hui les sciences de l'homme et de la société).

Les dualismes dont il vient d'être question reposent sur la thèse d'une séparation essentielle entre les deux termes d'esprit et de cerveau : entités sans aucun lien entre elles, ou bien sans lien informatif. Une autre façon de résoudre le problème de leur rapport est d'éliminer le plus problématique des deux, à savoir l'esprit, et de lui substituer un ordre de phénomènes observables, donc directement accessible à la science. Cet *ersatz* de l'esprit est le comportement : après tout, abstraction faite peut-être de quelques bribes fugitives, nos pensées se traduisent par des comportements déterminés. Nous pouvons donc faire abstraction des premières pour ne prendre en compte que les seconds, en laissant de côté la question de

savoir si les comportements expriment simplement les pensées, dont ils constitueraient la trace observable, ou s'ils constituent un ordre de phénomènes plus fondamental, les pensées étant peut-être alors des épiphénomènes, voire des illusions. Mettre de la sorte la fabrique des pensées entre parenthèses est l'acte fondateur du béhaviorisme, qui est souvent présenté comme un troisième obstacle à l'émergence des sciences cognitives, mais qui en est aussi, paradoxalement, l'accoucheur : le comportement est bel et bien l'observable fondamental pour les sciences cognitives, qui peuvent ainsi renoncer (dans une grande mesure) à la source peu fiable et limitée de l'introspection. Il n'empêche que pour émerger, les sciences cognitives ont dû s'affranchir de l'interdit jeté par le béhaviorisme sur la « boîte noire » de l'esprit.

## II. PAR-DELA LES OBSTACLES, LES RESSOURCES QUI MANQUAIENT, OU QUELLES ETAPES LES SCIENCES COGNITIVES ONT DU FRANCHIR POUR PRENDRE LEUR ESSOR

Après ce bref examen de ce qui a longtemps gêné, passons à ce qui a enfin permis. Il y a un certain arbitraire, ici encore, dans le choix des ingrédients dont la combinaison a donné naissance aux sciences cognitives ; il faut surtout souligner que, contrairement aux œufs, au beurre, à la farine, à la levure et au sel qui proviennent de canaux indépendants et ne sont mélangés qu'au dernier moment par le pâtissier, les éléments dont la combinaison a rendu les sciences cognitives possibles ont émergé dans un univers scientifique et philosophique commun, et se sont développés en s'appuyant les uns sur les autres. Cela n'empêche pas de les présenter, *post hoc*, séparément ; les chapitres suivants montreront comment ils interagissent encore aujourd'hui et se modifient en conséquence.

### 1. Un cadre conceptuel adapté

On s'accorde pour dater la naissance des sciences cognitives des années 1950 (1956 étant le plus souvent proposée, en raison de la tenue cette année-là de plusieurs rencontres importantes). C'est au cours de cette période, s'étendant jusqu'à la décennie suivante, qu'ont convergé des idées nées en philosophie, en sciences de l'ingénieur, en logique, en psychologie, en neurophysiologie.

#### a. Dualisme

Les philosophes ont d'abord mené une critique efficace contre le dualisme métaphysique. Ils ont montré qu'on peut surmonter les obstacles invoqués depuis des siècles contre la thèse de l'identité entre états mentaux et états cérébraux. Mais ils ont aussi montré que ce monisme n'obligeait pas à épouser le réductionnisme, c'est-à-dire l'idée qu'une science des états et processus mentaux se ramène nécessairement à une science des états et processus cérébraux, *a fortiori* à l'idée qu'elle se ramène nécessairement à la physique et à la chimie. Or le



réductionnisme était l'objet de critiques sévères, et restreignait considérablement l'espace des théories possibles ; en s'affranchissant du réquisit réductionniste, on cessait de restreindre les théories acceptables à celles qui sont issues, du moins directement, des neurosciences ou des sciences physiques. L'autonomie reconnue aux « sciences spéciales<sup>8</sup> » ouvrait l'horizon, sur le plan philosophique comme sur le plan scientifique.

#### b. Comportement

La psychologie fournissait à point nommé un concept essentiel, celui de comportement, dont il vient d'être question comme obstacle. Cet obstacle s'est retourné en ressource, comme il vient d'être dit : alors que le béhaviorisme, école dominante (surtout aux États-Unis, mais influente ailleurs) dans la psychologie expérimentale de la première moitié du XX<sup>e</sup> siècle, prétendait pouvoir faire abstraction de l'esprit, les fondateurs des sciences cognitives ont compris qu'on pouvait à la fois réhabiliter l'esprit (cessant ainsi de faire violence à tout un ensemble d'intuitions apparemment solides<sup>9</sup>) et faire du comportement la manifestation observable des processus mentaux. Cela permettait de soumettre ceux-ci à un contrôle expérimental beaucoup plus sûr et d'application beaucoup plus générale que l'introspection et la verbalisation. Si la notion de comportement devait être à son tour soumise à examen, si elle cessait d'être l'alpha et l'oméga de la psychologie expérimentale, du moins permettait-elle à un programme de recherche à la fois authentiquement mentaliste et authentiquement scientifique de se mettre en place.

#### c. Information et symbole

- La théorie naissante des télécommunications a engendré une notion inédite, celle d'information, prise en un sens général et abstrait. Mobilisée au même moment (les années 1940) par un autre programme de recherche à la frontière des sciences de l'ingénieur et des mathématiques, l'information a été placée au centre d'une discipline nouvelle, la cybernétique. Directement issues de la cybernétique, les sciences cognitives ont fait de l'information, en ce sens nouveau, leur concept clé (un peu ce que la nouvelle physique de Newton, exposée dans ses *Principia* [1687] fait de la notion générale et abstraite de force). L'information inscrit entre le physique et le mental un niveau intermédiaire, qui émerge du premier par abstraction et opérationnalise le second et le dotant d'une notion congruente au comportement observable, celle de processus mental, vu désormais comme un *comportement interne*.

- La logique a apporté une notion qui permettait d'étendre le schème explicatif de mécanisme à la sphère abstraite de l'information. Un mécanisme, au sens classique à partir de Descartes, est un système de parties solidaires exerçant des forces les unes sur les autres en sorte de produire un résultat sous la forme d'une modification de l'environnement matériel.

---

8. Cette autonomie n'est pas revendiquée seulement pour les sciences cognitives, ce qui pourrait paraître « *ad hoc* » : elle s'applique également à des sciences telles que la géologie, la météorologie, l'océanographie... Leurs objets ont beau n'être que des amas de particules élémentaires, elles ont leurs propres concepts, processus et lois, compatibles avec l'ontologie de la physique fondamentale mais qui n'en sont pas déductibles.

9. Au sein de la psychologie scientifique l'école de la Gestalt, rivale et contemporaine du béhaviorisme, a mis en évidence expérimentalement la réalité de phénomènes mentaux dépourvus d'effets comportementaux, contribuant ainsi à la naissance des sciences cognitives.

Dans le néo-mécanisme issu de la logique à la même époque (fin des années 1930), le rôle est occupé par une machine abstraite agissant sur des symboles, selon une suite d'opérations réglées, en sorte de modifier la situation informationnelle. Cette machine (souvent appelée machine de Turing du nom de celui qui en eut l'idée) peut se concrétiser sous forme de différents systèmes physiques, par exemple sous la forme d'un ordinateur ; mais elle ne se réduit à aucune machine particulière, ce qui permet d'étudier son fonctionnement sans se soucier des interactions physiques entre les parties d'une telle concrétisation<sup>10</sup>.

- Mais que sont ces symboles sur lesquels opère la machine informationnelle ? Où réside l'information en question ? Ici intervient un autre concept clé, celui de représentation. Il s'agit d'une reprise moderne d'une notion développée par les philosophes rationalistes du XVII<sup>e</sup> siècle, principalement Descartes et Locke, celle d'idée : une chose interne à l'esprit qui représente (ou présente) à l'esprit un objet, fait ou propriété de l'environnement (et par extension, éventuellement, un objet, fait ou propriété de l'esprit lui-même, ce qui peut ouvrir la porte aux idées abstraites). Ainsi, la machine informationnelle contient-elle des représentations (constituées de symboles simples ou assemblés) sur lesquelles elle effectue une suite d'opérations au terme desquelles est produit un nouveau symbole, porteur d'une nouvelle information, de la même manière qu'un calculateur effectue sur des nombres, ou sur des symboles algébriques, des opérations mathématiques au terme desquelles est obtenu un résultat numérique ou symbolique.

#### d. Implémentation biologique

Reste la question du cerveau : comment peut-il être compris comme un système de traitement de l'information ? Une première étape, décisive, avait été franchie dès la fin du XIX<sup>e</sup> siècle, quand le grand neuroanatomiste espagnol Santiago Ramón y Cajal établit la « doctrine neuronale » selon laquelle le cortex est un assemblage de cellules individuelles transmettant l'influx nerveux par contact ou induction électrique. La seconde étape est l'idée de modéliser le neurone comme un automate à seuil, capable de s'intégrer dans des réseaux réalisant l'équivalent neural d'un circuit électrique, et exécutant les fonctions logiques de base. Proposée en 1943 par le neurophysiologiste américain Warren McCulloch (1898-1969) et son compatriote le mathématicien Walter Pitts (1923-1969), ce modèle très simple est caractéristique de la cybernétique naissante et ouvre la voie aux neurosciences cognitives au seuil des années 1950.

Voilà donc, très rapidement brossé, le cadre général dans lequel il est apparu possible, assez soudainement, d'enquêter simultanément sur les processus mentaux (l'enchaînement des pensées) et sur les phénomènes cérébraux (les processus physico-chimiques au sein du tissu neural). Le cerveau est vu, à un certain niveau de description, comme un système (hautement complexe) traitant de l'information, c'est-à-dire calculant sur des représentations internes. Dans ce cadre, il apparaît comme une manière de réplique biologique de

---

10. C'est dire d'une manière absconse quelque chose que le plus endormi des lecteurs a compris depuis longtemps : ce que fait son ordinateur lorsqu'il met un terme en italiques peut être compris, prédit, commandé, sans qu'on ait la moindre notion d'électronique. Mais l'arrière grand-père de notre lecteur, à moins d'être un génial visionnaire, ne l'aurait sans doute pas imaginé : c'est la familiarité avec les outils numériques qui nous donne à tous désormais une compréhension quasi instinctive de la notion de traitement de l'information.

l'ordinateur, ce qui ne manque pas de soulever des objections. On y reviendra dans la section IV.2, mais deux remarques peuvent être faites dès à présent :

1) Dire que d'un certain point de vue, X est comme Y, ce n'est pas dire que X et Y sont deux exemplaires issus d'un même moule, identiques sous tous les angles : d'un autre point de vue, ils peuvent être très différents.

2) Loin que l'ordinateur ait d'abord été inventé à des fins utilitaires, puis recruté comme métaphore pour le cerveau ou pour l'esprit, il a résulté de la même percée conceptuelle que le cadre qui vient d'être exposé : l'ordinateur et la conception néo-mécaniste de l'esprit/cerveau sont des jumeaux. Quoi que l'on puisse penser, hier, aujourd'hui ou demain, de l'utilité de penser l'un en fonction de l'autre, on ne peut y voir seulement une analogie au mieux superficielle, au pire absurde voire nuisible. Provisoirement, on peut se contenter de s'appuyer sur l'idée qu'on se fait de certains processus informatiques pour s'aider dans la compréhension de certains phénomènes mentaux et/ou cérébraux, sans s'engager dans l'évaluation du « modèle » que l'ordinateur constituerait pour l'esprit/cerveau.

## *2. Un accès empirique aux phénomènes mentaux*

Encore faut-il pouvoir accéder aux phénomènes mentaux par des canaux acceptables pour la science<sup>11</sup>. Or les phénomènes mentaux ne se déploient ni dans l'espace, ni – du moins l'a-t-on longtemps pensé – dans le temps, si le temps est celui des horloges, non celui du flux de la conscience. Aussi la psychologie est-elle restée, dans une large mesure et jusqu'à relativement récemment, une discipline non expérimentale, dont le statut de science était sujet à débat. Non qu'elle prétendît se passer de données : mais ses données provenaient de sources scientifiquement douteuses : l'introspection, les rapports verbaux, sans parler de l'hypnose ou de la libre association. Si la psychophysique et la Gestalt inventèrent des moyens pour contourner la difficulté, en réduisant la subjectivité des données à un minimum strictement contrôlé, elles ne s'occupaient que de la perception ; d'autres domaines limités donnèrent également naissance à des méthodologies acceptables, par exemple, les tests d'intelligence (Binet), ou la capacité de résoudre certains problèmes (Piaget). Mais c'est le béhaviorisme qui a le plus clairement placé l'impératif de la mesure à la première place, en l'imposant à toute forme de comportement (le comportement étant ce dont, dans cette perspective, la psychologie est l'étude). C'est ainsi que la psychologie s'est dotée d'un ensemble de méthodes rigoureuses pour concevoir des paradigmes expérimentaux valides et en tirer des données solides. Sur ce plan, du reste, le psychologue suisse Jean Piaget (1896-1980) comme son contemporain et rival le soviétique Lev Vygotski (1896-1934), tous deux aux antipodes du béhaviorisme sur le plan théorique, puisque chacun à sa manière résolument mentalistes,

---

11. « Empirique » revêt en français deux acceptions. La plus courante est péjorative : est empirique ce qui ne repose pas sur une théorie ou sur une méthode de recherche rigoureuse, mais sur une suite opportuniste d'essais et d'erreurs. Au sens technique, « empirique » marque la relation à des phénomènes objectivement observables ou constatables. « Expérimental » peut être souvent en pratique utilisé à la place d'« empirique », mais le sens est différent : il caractérise les démarches que l'on peut faire pour produire des données empiriques et en tirer des conséquences quant aux phénomènes observés, sous forme d'explications ou plus généralement de théories.

s'en rapprochent, dans la mesure où ils se reposent sur la capacité des sujets à réaliser certaines tâches bien définies, et non sur leurs impressions introspectives.

Dans sa tentative pour ramener tous les processus mentaux à leurs traces comportementales, sur la base d'une conception uniforme infiniment trop restrictive des capacités humaines (et même animales), le béhaviorisme a échoué. Mais il a légué aux sciences cognitives qui lui succèdent ses hautes exigences et son savoir-faire en matière de méthodologie expérimentale.

De fait, la plupart des accès empiriques rigoureux au mental étaient disponibles, parfois sous forme rudimentaire, au moment où les sciences cognitives ont pris leur essor. Ainsi, la méthode des temps de réaction qui permet, dans les cas favorables, de confirmer ou d'infirmer une hypothèse sur le décours temporel d'une suite d'opérations internes, a été inventée par le psychologue néerlandais F. C. Donders (1818-1889), vers la fin des années... 1860 ! Les premiers instruments<sup>12</sup> permettant de mesurer les performances des sujets en matière de sensation, perception, attention ou mémoire, ont été mis au point par les écoles allemandes de la fin du XIX<sup>e</sup> siècle.

Les sciences cognitives ont cependant elles-mêmes enrichi leur trousse à outils. En apportant, tout d'abord, des perfectionnements considérables aux dispositifs existants, rendus possibles par les formidables progrès technologiques en matière de matériaux, d'électronique et bien entendu d'informatique. L'ordinateur a joué un rôle, distinct mais non moins important, dans le traitement des données et la construction et mise à l'épreuve de modèles, comme dans toutes les sciences ; mais il a également révolutionné l'art de concevoir des paradigmes expérimentaux, en démultipliant les possibilités et en facilitant celles qui existaient déjà. Ensuite, la volonté de « faire parler » les sujets non verbaux, animaux non humains, enfants préverbaux, humains n'ayant pas ou plus accès au langage, a suscité chez les psychologues des efforts héroïques d'imagination. Dès lors qu'ils étaient convaincus que les processus mentaux sont réels bien que (contrairement aux réactions organiques et aux comportements) non directement observables<sup>13</sup>, ils se sont mis en quête de moyens *indirects*, quête qui avait si bien réussi aux (autres) sciences de la nature, de l'astronomie à la chimie, à la biologie, à la géologie... C'est ainsi que le paradigme de la succion non nutritive permet d'étudier les processus mentaux du nourrisson dès les premières heures d'existence ; et que les *eye trackers* permettent de déterminer avec une précision extraordinaire la trajectoire du

---

12. Ces appareils comprenaient souvent des parties en laiton (*brass* en anglais, alliage dont sont faits les « cuivres » de l'orchestre), ce qui a valu à cette orientation de la psychologie, de la part de William James qui ne l'aimait guère, l'appellation de « psychologie des cuivres » (*brass instrument psychology*).

13. Comme l'écrivait Ulric Neisser dans un traité qui donna son nom actuel à la psychologie cognitive, « Pourquoi étudier les processus cognitifs ? Pour la même raison, c'est désormais parfaitement clair, qu'on a d'étudier n'importe quoi : à savoir qu'ils sont là. Notre connaissance du monde ne peut se développer, d'une manière ou d'une autre, qu'à partir des stimuli que nous recevons en entrée ... Les processus cognitifs existent, on ne peut en douter, donc en faire l'étude n'est sûrement pas antiscientifique. » (1967, p. 5). Cette conception réaliste des processus mentaux, qui semble aller de soi aujourd'hui, a été arrachée de haute lutte aux écoles dominantes de l'époque, fortement soutenues par l'antiréalisme (on dit parfois aussi dans ce contexte le positivisme) ambiant en philosophie des sciences.

regard, source d'enseignements d'autant plus précieuse que ces trajectoires échappent largement au contrôle de l'agent.

C'est en fin de compte là que réside la conquête méthodologique des sciences cognitives : s'être affranchies du recours obligé au témoignage des sujets, tant de ceux qui ne peuvent l'apporter que de ceux qui le peuvent. Ces témoignages sont en effet constitutivement non fiables, comme on en a progressivement pris conscience tout au long du XX<sup>e</sup> siècle, ou bien inexistantes – nous n'avons rien à dire sur une foule de processus : comme le disait déjà le philosophe écossais Alexander Bain (1818-1903), même s'agissant de processus conscients, la plus méticuleuse introspection ne nous en fournit jamais que quelques instantanés séparés par des noirs ; quant aux processus « subpersonnels », tels que ceux impliqués dans la perception ou la mémoire, nous ne disposons pas même dans le langage courant des mots pour les décrire.

### *3. Un accès empirique aux phénomènes cérébraux*

Dans le domaine de ce qu'on appelle aujourd'hui les neurosciences, la situation est toute différente.

En premier lieu, elles ne connaissent pas les doutes qui ont assailli la psychologie au cours de son histoire quant à la nature véritable de ses objets. Pas plus que leur discipline-mère, la biologie et sa jumelle la médecine, les neurosciences n'éprouvent le besoin de s'interroger sur l'existence des systèmes et processus qu'elles étudient : ils n'ont de choix qu'entre exister et disparaître du discours scientifique. Bien entendu on est parfois incertain sur l'existence d'une entité particulière : les neurones ont refusé longtemps de sauter aux yeux des neuroanatomistes, et il a fallu batailler pour les faire admettre. Mais une fois admis, personne n'a songé à les reconstruire, à les réinterpréter comme des fictions commodes, des résumés d'expérience. C'est donc sans états d'âme, avec la conviction de l'explorateur, que les neuroscientifiques se sont mis en quête de décrire et d'expliquer l'organe cérébral.

En second lieu, la connaissance empirique du cerveau a procédé, pour simplifier, en deux grandes étapes : la première est largement antérieure à l'émergence des sciences cognitives et en est indépendante ; la seconde résulte au contraire d'un processus de coévolution, au cours duquel les sciences cognitives ont subi une profonde transformation, à mesure qu'elles suscitaient la mise au point d'instruments entièrement inédits induisant à leur tour cette transformation.

La première étape a consisté à constituer une image statique de l'appareil cérébral, à différentes échelles, depuis le neurone et ses composantes (cf. chap. I « De la molécule au neurone »), dont les synapses, jusqu'à l'organisation générale du cerveau et à la cartographie des aires corticales, en passant par les structures intermédiaires telles que les colonnes corticales (cf. chap. II « Du neurone au cerveau »).

Entre la première et la seconde grande étape se situe une phase combinant une découverte fondamentale en neurophysiologie et une percée conceptuelle. Elles sont dues toutes deux au psychologue canadien Donald Hebb (1904-1985), dont l'ouvrage de 1949, *The Organization of Behavior*, a fait plus que tout autre pour favoriser à la fois la transition du béhaviorisme au

paradigme cognitif et l'intégration des neurosciences dans ce paradigme<sup>14</sup>. Il mit en évidence un mécanisme (cf. *infra*, encadré « Plasticité synaptique ») rendant compte, au niveau neuronal, de l'apprentissage (cf. chap. IV « Le développement », § III)<sup>15</sup>. Il proposa également de considérer sa conjecture comme relevant de ce qu'il appela le « système nerveux conceptuel », un niveau de théorisation intermédiaire entre la neurophysiologie exacte et la description psychologique, qui était en train de devenir une analyse des traitements de l'information. Anticipée par l'article de McCulloch et Pitts mentionné plus haut, cette proposition ouvrait la porte au connexionnisme dit « PDP » pour « *parallel distributed processing* » des années 1980 et aux modélisations des neurosciences théoriques contemporaines. Elle autorisait Hebb à inférer de données encore très fragmentaires une hypothèse plausible sur le fonctionnement de la synapse ; ce n'est qu'à partir de la fin des années 1960 que les mécanismes cellulaires, moléculaires et biochimiques ont pu être déchiffrés lorsque le rôle des neurotransmetteurs (découverts plus tôt dans le système nerveux périphérique) dans le cerveau fut établi.

Comprendre comment le cerveau se modifie durablement au contact de l'expérience est l'une des façons de passer de la statique à la dynamique. Le tempo est ici relativement lent, c'est celui de l'acquisition de nouvelles capacités. Un demi-siècle plus tard, c'est-à-dire, soulignons-le, bien après que les sciences cognitives se sont installées dans le paysage scientifique et ont adopté le cadre théorique présenté plus haut, un tempo beaucoup plus rapide devient accessible à l'étude expérimentale : c'est celui des processus cérébraux accompagnant la réalisation de tâches mentales. L'imagerie fonctionnelle cérébrale, appuyée principalement sur la technologie nouvelle de la résonance magnétique (IRMf, en anglais fMRI), qui permet de détecter de manière assez précise les aires cérébrales activées spécifiquement lors du déroulement d'opérations mentales d'un certain type, a pris une place de choix (certains diraient la première place) dans la trousse à outil des sciences cognitives, et sont complétées par d'autres techniques d'imagerie aux qualités complémentaires. Si leur importance n'est pas douteuse, elles ne sont ni la première ni l'unique source de données empiriques pour les sciences cognitives – il faut le préciser pour éviter un sérieux malentendu non seulement sur l'histoire, mais sur la structure actuelle du domaine : un ange envoyé du Ciel eût-il mis un scanner IRM 11 Tesla dernier cri entre les mains des chercheurs en 1945, voire 1960, ils n'auraient su qu'en faire. L'outil a été développé une fois jetées les bases du programme de recherche, et il l'a été grâce aux résultats obtenus pendant ses premières phases.

---

14. Je ne saurais trop recommander la lecture de son ouvrage tardif, *An Essay on Mind* (1980).

15. Bien entendu la proposition de Hebb n'est pas le dernier mot de la science sur cette question : les connaissances ont considérablement évolué depuis, sans pour autant reléguer l'idée fondamentale de Hebb aux poubelles de l'histoire.

### III. LA QUESTION DES DISCIPLINES : MONO-, PLURI-, INTERDISCIPLINARITE DANS LES SCIENCES COGNITIVES

Jusqu'à présent il a été essentiellement question de psychologie, et dans une moindre mesure de neurosciences, les autres disciplines étant présentées comme des auxiliaires. Cela reflète l'objectif des sciences cognitives, qui est de rendre compte des phénomènes mentaux : en ce sens elles se situent dans le droit fil des phases antérieures de la psychologie. Mais on a aussi perçu qu'elles sont nées d'un effort pour surmonter les limites et impasses de la psychologie en s'appuyant sur des courants de pensée et des disciplines qui lui étaient extérieures. Ce qui est certain c'est qu'elles ne se réduisent pas à la psychologie, un sujet de controverse étant de savoir si celle-ci en constitue cependant le cœur, ou le fil directeur : certains lui dénie ce rôle, ou estiment que cela n'a pas vraiment de sens de poser la question en ces termes. Par ailleurs, qui consulterait des manuels ou des programmes d'étude en psychologie, des encyclopédies ou des histoires de la psychologie moderne, constaterait que la psychologie poursuit son propre chemin, sur lequel les sciences cognitives constituent un moment, un épisode, un avatar, et non du tout un aboutissement. Nous devons donc nous arrêter un moment sur la structure disciplinaire des sciences cognitives, et leurs rapports aux disciplines constituées.

#### *1. Le rôle des disciplines constituées*

On débat depuis des lustres, en philosophie des sciences, sur la nature des disciplines : sont-elles largement fixées par la nature des choses, par l'organisation du monde en ordres de réalité relativement stables et distincts, ou bien ne font-elles que refléter une organisation du travail académique, imposée par des contraintes sociologiques et historiques, plus ou moins arbitraire, variable et sans signification théorique ? Les sciences cognitives sont concernées par cette question générale : différentes orientations en leur sein sont liées à des réponses divergentes, et leur évolution, dans les années et décennies à venir, constitueront un élément important du dossier. Je ne peux que signaler ici la véritable ampleur de la question, et me bornerai à un état des lieux aussi consensuel que possible.

Les disciplines constituées qui nous concernent se divisent en trois groupes. Le premier comprend les constituants fondamentaux des sciences cognitives. On pourrait être tenté de simplifier la formulation en disant simplement que chacune de ces disciplines est *une* science cognitive et qu'ensemble elles constituent *les* sciences cognitives, mais ce serait non pas légèrement trompeur, ce serait complètement inexact, nous allons vite comprendre pourquoi. Ce premier groupe comprend la psychologie, la linguistique, les neurosciences, l'intelligence artificielle (IA) et la robotique, la philosophie, l'anthropologie. Le deuxième groupe comprend les disciplines qui fournissent aux sciences cognitives des outils indispensables : mathématiques, physique, informatique, biologie (en pratique seules certaines branches de ces disciplines jouent un rôle). Le troisième groupe est un ensemble aux contours mal définis, qui tendent à s'étendre avec le temps ; il s'agit des disciplines qui sont affectées par les sciences

cognitives, au sens où elles tentent d'intégrer certains de leurs résultats et de leurs méthodes. Ce sont, à des degrés divers, toutes les sciences de l'homme et de la société, les sciences juridiques, les disciplines médicales et sportives, les sciences de l'éducation. Certaines (en premier lieu l'économie et la médecine) sont non seulement consommatrices, mais également contributrices, en sorte que même la division en trois groupes est à prendre avec précaution (cf. chap. XV « Les sciences cognitives appliquées », où des exemples sont développés.)

#### a. Recherches mono-disciplinaires

Ce n'est pas parce que les sciences cognitives sont intrinsèquement interdisciplinaires que tout programme de recherche en leur sein l'est nécessairement : de fait, beaucoup, et même probablement le plus grand nombre de travaux relevant incontestablement des sciences cognitives sont mono-disciplinaires : ils se rangent sans conteste dans le cadre d'une des disciplines constituées du premier groupe : cela se vérifie sur le plan institutionnel (lieux de communication et de publication) et sur celui des méthodes. S'ils ne sont pas *seulement* des recherches de psychologie, de linguistique, de neurosciences, de philosophie, etc., c'est parce qu'ils s'insèrent dans une dynamique impliquant des concepts, questionnements, résultats d'autres disciplines, tout cela se plaçant dans un cadre, un contexte qui n'est pas d'abord celui de la psychologie et de son histoire, de la linguistique et de ses traditions, etc., mais celui des sciences cognitives.

Cette double appartenance, dissymétrique comme elles le sont toutes, est une caractéristique de grande importance pour comprendre l'organisation et l'évolution du domaine, et pour s'y orienter efficacement, notamment au cours des études et des premières recherches. Elle explique pourquoi il est recommandé de n'aborder les sciences cognitives qu'armé d'une solide formation dans l'une des disciplines constituantes, et pourquoi à l'issue de la thèse le jeune docteur est nécessairement affilié à l'une d'elle, qui n'est d'ailleurs pas toujours sa discipline initiale.

Les exemples sont si nombreux et variés qu'il ne serait pas très éclairant d'en proposer quelques-uns ici : le présent manuel en est plein. C'est, pour reprendre l'expression des linguistes, le « cas non marqué » : les recherches portent le plus souvent sur un phénomène relativement circonscrit, étudié sous un angle particulier, à un niveau donné, selon un paradigme expérimental, ou dans une tradition disciplinaire existante. Le « cas marqué » est celui, beaucoup moins fréquent, des recherches réellement interdisciplinaires, qui ne se laissent pas décomposer en composantes mono-disciplinaires.

#### b. La partie cognitive des disciplines constituées

Les disciplines participantes conservent donc leur identité même lorsqu'elles s'impliquent dans les programmes de recherche des sciences cognitives. De plus, elles ne s'y impliquent que pour une partie : il s'en faut que tout chercheur en psychologie se considère comme concerné par ces programmes, et il en va de même, à des degrés croissants, lorsqu'on passe aux neurosciences, à l'IA et à la robotique, à la linguistique, à l'anthropologie, à la philosophie.

Chacune de ces disciplines comporte donc à côté d'une partie « cognitive » une partie qui ne l'est pas. Comment se distinguent-elles, et qu'est-ce qui vaut, pour une branche ou un programme de recherche X d'une discipline constituée, appartenance aux sciences



cognitives ? Il n'y a pas consensus sur cette question, mais une majorité de chercheurs s'accordent pour penser que deux conditions doivent être satisfaites :

1) Au sein de X, il doit être admis que les phénomènes étudiés se situent pour une part à un niveau informationnel conceptuellement distinct du niveau de la description physico-chimio-biologique.

2) X accepte le principe d'interactions avec d'autres branches des sciences cognitives, qui peuvent compléter les descriptions développées au sein de X, et peuvent leur imposer des contraintes (par exemple, un processus informationnel postulé par X doit pouvoir être physiquement réalisé en un temps raisonnable par le genre de système dont la branche Y a des raisons de penser qu'il réalise chez l'humain ou telle autre espèce le processus en question). En pratique, X interagit effectivement avec certains programmes en sciences cognitives, dans des configurations qui peuvent varier d'une période à l'autre et d'un laboratoire à l'autre.

Bien entendu, il n'existe pas de procédure formelle d'adhésion aux sciences cognitives. Du reste, cette appartenance est affaire de degré, et peut évoluer. De plus, on le verra, les sciences cognitives abritent des écoles de pensée très diverses et parfois antagonistes : tel type de recherche peut apparaître à une école, mais pas à une autre comme relevant des sciences cognitives. C'est finalement l'historique complexe des filiations théoriques et des interactions avec certains programmes phares qui détermine la place qu'occupe à un moment donné telle ou telle approche au sein de l'une des disciplines constituées.

### c. La partie non cognitive des disciplines constituées

Au sein de chacune des disciplines constituantes des sciences cognitives, il existe, on l'a dit, des branches non cognitives. Les raisons de ne pas participer étant très variables, et les grandes disciplines étant toutes très hétérogènes, ces branches n'ont généralement rien en commun, sinon de ne pas se reconnaître dans le programme et les hypothèses des sciences cognitives. Il y a deux grandes raisons qu'un chercheur peut avoir de s'en démarquer. Il peut considérer, d'abord, que l'objet auquel on s'intéresse n'a pas de rapport ou du moins aucun rapport théoriquement significatif avec les sciences cognitives, dont il n'y a donc rien à attendre (la réciproque pouvant être fautive : pour prendre un exemple, la biochimie de la synapse ne doit rien aux sciences cognitives mais leur apporte une information qui peut se révéler importante). Il peut aussi être en désaccord avec les orientations théoriques des sciences cognitives ou avec la manière dont elles abordent le sujet dont on est spécialiste. Il existe des zones grises, lorsqu'il n'est pas vraiment clair que les sujets d'étude sont distincts, ou que les approches sont incompatibles. Il existe aussi des cas dans lesquels des chercheurs estiment que leurs travaux relèvent des sciences cognitives en raison même de leur objet, alors que leurs collègues des sciences cognitives jugent qu'ils n'en ont pas saisi l'esprit ni la méthode.

## 2. *Par où passent les frontières ? Les sciences cognitives dans le paysage scientifique*

S'il est utile d'être conscient de ces complications, il n'est généralement pas nécessaire de s'y attarder, sauf dans des moments de crise, collective ou individuelle, lors d'un changement de stratégie de recherche ou de plan de carrière. Concluons donc simplement par quelques

indications sur la place des sciences cognitives dans ses différentes disciplines constituantes, ce qu'elles en incluent et ce qu'elles en laissent de côté. La situation diffère beaucoup de l'une à l'autre. Je profite de ce tour d'horizon pour donner quelques aperçus sur le domaine dans son ensemble, en mentionnant notamment certaines provinces que le présent volume, qui ne peut prétendre à l'exhaustivité, laissera de côté. J'examinerai le cas des six disciplines qui sont généralement présentées comme formant l'hexagone fondateur des sciences cognitives, en les regroupant deux à deux selon leur type de connexité avec les sciences cognitives, et en procédant par ordre de complexité croissante : en commençant par deux disciplines dont les rapports avec les sciences cognitives sont relativement simples et directs et en terminant par deux disciplines dont ces rapports soulèvent des questions relativement compliquées.

#### a. Psychologie et intelligence artificielle

Une première configuration est occupée par la psychologie et par l'IA (à laquelle on peut ici adjoindre la robotique). L'objet de ces deux disciplines coïncide largement avec celui des sciences cognitives, et elles ne comportent pas de parties qui en seraient en droit séparées par une frontière théorique étanche. La psychologie a pour objet les phénomènes et processus mentaux, dont le comportement constitue à la fois une fonction et une trace essentielle. Les sciences cognitives sont rigoureusement alignées sur cet objectif, et ne se limitent pas, comme cela a été parfois dit au début, y compris par certains chercheurs du domaine, aux processus d'acquisition des connaissances, de mémoire, de raisonnement, de résolution de problèmes. L'IA quant à elle avait pour vocation officielle de doter des systèmes artificiels de capacités de traitement de l'information requérant chez l'humain de l'intelligence, mais il était en réalité clair dans l'esprit des fondateurs, et il est rapidement apparu aux autres, que ce projet ne pouvait être dissocié de l'étude de l'intelligence humaine, ce qui dans le contexte ne signifie rien d'autre que la cognition humaine : comme l'écrivait dès les années 1970 un des principaux contributeurs à l'IA première manière, Roger Schank, l'étude de l'intelligence artificielle n'est rien d'autre que celle de l'intelligence humaine.

Mais à s'en tenir à cette présentation, on se demande comment la psychologie et l'IA, chacune pour ce qui la concerne, peuvent éviter de se fondre purement et simplement dans les sciences cognitives. La réponse varie dans les deux cas.

La psychologie est, on l'a dit, largement antérieure aux sciences cognitives, et poursuit sa voie sans nécessairement envisager une fusion à terme avec elles : bon nombre de ses branches ne se sentent pas particulièrement concernées – pour elles, la psychologie cognitive n'est qu'une branche parmi d'autres, celle précisément qui a partie liée avec les sciences cognitives. Pendant la phase initiale des sciences cognitives, la psychologie comportait des branches, telles que la psychologie clinique et la psychanalyse, la psychologie du travail, la psychologie des organisations, la psychologie différentielle, la psychologie sociale, qui ne soupçonnaient même pas qu'il puisse exister un lien avec une spécialité émergente dont les objectifs typiques étaient de comprendre comment un individu mémorise une liste de mots, comment il raisonne (cf. chap. XI « Raisonnement ») ou résout le problème des tours de Hanoi, de construire des modèles formels de la syntaxe de l'anglais, ou encore de faire jouer un ordinateur aux échecs et de lui faire traduire des textes écrits... Cette image des sciences cognitives ayant largement cédé la place à une conception beaucoup plus inclusive, comprenant notamment la « cognition chaude » (émotions, sentiment de soi, rapport à

autrui...), il est devenu difficile de s'en tenir à l'idée que les domaines considérés sont disjoints. Restent alors deux manières, pour une branche ou école de psychologie, de rester à l'écart des sciences cognitives : soit en récusant leurs hypothèses fondamentales, ou en contestant qu'elles s'appliquent au domaine de la branche en question (c'est typiquement le cas de la psychanalyse et de ses différentes branches, même s'il existe des tentatives de conciliation ; et d'une bonne partie de la psychologie sociale) ; soit en plaidant la spécialisation ou la différence de niveaux, ce qui laisse du reste la possibilité, pour une telle branche, d'exploiter certains résultats issus des sciences cognitives. Sans pouvoir entrer ici dans la discussion qu'appellent ces différents arguments, ni décrire les évolutions récentes (ainsi une partie de la psychologie sociale a-t-elle fait alliance avec les sciences cognitives), disons simplement que, curieusement du point de vue qui est le nôtre dans le présent volume, la psychologie dans son ensemble n'est que partiellement investie dans les sciences cognitives, lesquelles y occupent cependant une place importante.

Le cas de l'IA diffère à plusieurs égards. En premier lieu, l'IA peut être considérée comme la première forme qu'ont prise les sciences cognitives : elle a en effet été la première « interdiscipline<sup>16</sup> » à formuler, dans les années 1950, le projet de ce qui ne s'est appelé sciences cognitives qu'une vingtaine d'années plus tard. En ce sens, l'IA est constitutivement immergée dans les sciences cognitives. Mais en tant que branche des sciences cognitives, elle a été largement supplantée et a été amenée à se reconstruire dans leur marges, en tant que branche de l'informatique spécialisée en une certaine forme d'ingénierie, chargée de rehausser les logiciels de propriétés d'adaptation au contexte leur conférant une sorte d'« intelligence » (les guillemets indiquant la décision d'éviter la question de savoir s'il existe un rapport avec le référent du terme sans guillemets). Les spectaculaires développements récents de l'IA, notamment grâce au *deep learning*, au *data mining*, à certaines techniques encore peu connues de traitement du langage naturel, ont renforcé à la fois l'autonomie de l'IA et ses liens avec les sciences cognitives, auxquelles certaines de ses branches sont solidement arrimées – ainsi la robotique (cf. chap. VII « L'action ») et la vision artificielle ou le traitement automatique du langage naturel.

#### b. Linguistique et anthropologie

La linguistique et l'anthropologie constituent un deuxième cas de figure. Leurs domaines respectifs sont différents de celui des sciences cognitives, mais lui sont intimement liés. La capacité de langage est certainement l'une des plus remarquables dispositions cognitives humaines, et en ce sens la linguistique peut apparaître comme une province (importante) des sciences cognitives. Inversement, si l'on peut dire, l'anthropologie contient les sciences cognitives dans la mesure où elle se donne pour objet la totalité des productions humaines, qui résultent en premier lieu des capacités cognitives de l'espèce et en constituent la trace la plus tangible. Si sommaire que soit cette caractérisation, on peut s'attendre à ce que les deux disciplines en question soient profondément impliquées dans l'entreprise des sciences

---

16. C'est ainsi que la désigne l'Américain Allen Newell (1927-1992), l'un des fondateurs de la psychologie cognitive aux côtés de son compatriote Herbert Simon (1916-2001, Prix Nobel d'économie en 1978).

cognitives, et qu'inversement celles-ci leur aient imprimé une évolution décisive. Eh bien, ce n'est l'opinion que d'une minorité des chercheurs et des écoles dans les deux disciplines.

Le rôle de la linguistique est très particulier : elle a contribué à l'émergence des sciences cognitives, un peu comme l'IA, à la même époque et dans des cadres institutionnels se recoupant largement. Mais contrairement à l'IA, dont le *telos* (le destin, la vocation) était de donner naissance aux sciences cognitives, la linguistique, plus ancienne de quelques millénaires, suivait son bonhomme de chemin, lorsqu'elle a donné naissance à une nouvelle orientation théorique en son sein et co-engendré les sciences cognitives : l'Américain Noam Chomsky (1928-) est en effet le père d'une école de linguistique, la grammaire générative, qui a l'ambition de constituer le cadre général de la linguistique théorique ; et il est aussi l'un de ceux qui ont le plus contribué à définir le cadre théorique des sciences cognitives naissantes. Le lien entre les deux aspects de la pensée de Chomsky est l'idée que ce qui, au sein de la masse indistincte des phénomènes impliquant le langage, peut faire l'objet d'une science authentique est une certaine capacité cognitive, dont tout humain normalement constitué est doté à la naissance, et qui se développe selon un calendrier essentiellement invariant dans toute communauté linguistique donnée, donnant lieu spontanément, chez l'enfant, à l'acquisition de la langue de cette communauté. La linguistique au sens que lui donne Chomsky est donc une branche des sciences cognitives. Mais il y a plus : la forme que prend cette capacité constitue selon lui le modèle de toute capacité cognitive, modèle que résume le titre d'un de ses ouvrages, *Règles et représentations*, et qui est, dans ses grandes lignes, celui que proposent les autres courants dont la convergence donnera naissance aux sciences cognitives.

On voit donc à quel point les destins de la linguistique et des sciences cognitives s'entrecroisent. Mais ils ne le font que dans la perspective chomskyenne (qui inclut non seulement celle de Chomsky lui-même et de ses disciples, mais aussi celle de nombre de linguistes influents qui s'en sont éloignés, tout en souscrivant à sa conception naturaliste du langage), et cette perspective est vivement contestée dans les deux domaines (pour des raisons qui ne sont qu'en petite partie communes). Au sein de la linguistique, il existe, en premier lieu, une profonde division entre ceux qui voient dans Chomsky le fondateur d'une nouvelle science du langage, celle à laquelle ils se consacrent, et ceux qui ne lui concèdent pas un tel rôle ; en deuxième lieu, des désaccords tenaces sur la contribution de Chomsky aux problèmes classiques de la linguistique ; en troisième lieu, un débat sur la pertinence de la dimension mentale ou cognitive pour l'étude du langage (pour ceux qui l'acceptent, et non pour ceux qui la nient, la linguistique est nécessairement impliquée dans les sciences cognitives) ; en quatrième et dernier lieu, parmi les défenseurs du point de vue mentaliste, de sévères controverses sur ce que doit être une théorie du langage comme capacité cognitive (ainsi, l'école qui porte le titre « linguistique cognitive » regroupe des chercheurs très opposés aux conceptions chomskyennes). Tout cela explique que la linguistique soit pour une part profondément imbriquée dans les sciences cognitives, et que pour une part elle en demeure relativement éloignée.

Les rapports entre anthropologie et sciences cognitives ont évolué de manière encore différente. L'anthropologie comptait parmi les six grandes disciplines généralement présentées comme les fées disposées autour du berceau des sciences cognitives. Il y eut en

effet une « lune de miel » au cours de laquelle une nouvelle forme d'anthropologie, dite cognitive, se développa ; selon l'un de ses principaux acteurs, Roy D'Andrade, la psychologie cognitive explorait les mécanismes de la pensée humaine (*how people think* : comment on pense), laissant à l'anthropologie cognitive le soin d'en caractériser les contenus spécifiques (*what people think* : ce qu'on pense). La nouvelle anthropologie réagissait notamment contre les thèses conjointes du relativisme culturel et de l'autonomie de la culture : elle pensait la culture comme un ensemble de croyances et de pratiques universelles et directement ancrées dans les capacités de l'esprit humain, conçues comme caractéristique de l'espèce (cf. chap. XIV « De la coopération à la culture »). Les variations auxquelles les anthropologues accordent traditionnellement une importance capitale sont, selon elle, superficielles et explicables par des mécanismes voisins de ceux qui produisent, à partir de la capacité de langage théorisée par Chomsky, les quelques milliers de langues naturelles à l'apparence si variée. L'exemple le plus connu de cette approche portait sur les systèmes de classification et de nomination des couleurs : une étude ethnographique menée conjointement par un linguiste et un anthropologue montrait que ces systèmes étaient en réalité en nombre restreint, et se rangeaient dans un tableau très simple. Cette étude montrait, aux yeux de ses auteurs, que ni la culture (considérée comme un facteur causal autonome) ni les facteurs contextuels (climatiques, démographiques, géographiques, historiques...) ne peuvent donner naissance à une variété indéfinie de taxinomie des couleurs, laquelle est sous le contrôle de facteurs cognitifs universels.

Pour différentes raisons, l'anthropologie cognitive première manière s'est étiolée, et l'anthropologie a pratiquement disparu pendant deux décennies de l'horizon des sciences cognitives, dont corrélativement la place dans la discipline a été ramenée à peu de chose (elle n'avait, du reste, pas eu le temps de croître beaucoup). Les choses ont commencé à changer à partir de la toute fin du siècle dernier, sous l'effet de trois facteurs principaux. Le premier est le « tournant social » des sciences cognitives, mû à la fois par une prise de conscience de l'importance de la socialité à tous les niveaux de la constitution et de la genèse des fonctions cognitives, et par la prise en compte de la dimension évolutionnaire des phénomènes. Le deuxième, distinct quoique lié, est ce qu'on peut appeler le « tournant expansif » des sciences cognitives, qui replace la cognition dans un contexte de situations particulières, comportant le corps, le mouvement, les artefacts, les savoir-faire, la communauté d'interactions – il en sera question dans la section suivante. Le troisième est interne à l'anthropologie : il s'agit de l'effort obstiné d'un petit nombre de chercheurs pour remettre à l'ordre du jour le projet initial, mais en apportant des idées nouvelles qui tiennent compte à la fois des limites de la première phase, des progrès et de la maturation des sciences cognitives et de l'apport des approches évolutionnaires. Nous n'en sommes néanmoins qu'au début d'une renaissance qui reste largement programmatique : dans l'état présent, la part cognitive de l'anthropologie reste assez réduite.

### c. Neurosciences et philosophie

Restent les neurosciences et la philosophie, qui occupent deux pôles opposés sur la carte des sciences cognitives et présentent de ce fait des configurations en quelque sorte symétriques. Les neurosciences sont incontestablement la plus « dure » des sciences impliquées, la philosophie la moins « dure », et n'est du reste pas à proprement parler une

science. D'autre part, les premières sont si évidemment liées au projet des sciences cognitives (cette évidence n'étant cependant pas apparue immédiatement) qu'il semble facile de caractériser les rapports entre les deux : comme l'IA, mais sans son caractère hybride de science et de technologie, les neurosciences sont à première vue complètement immergées dans les sciences cognitives. Inversement, la philosophie, pour ce que chacun en connaît, se situe manifestement largement en dehors d'elles. Dans l'un comme dans l'autre cas, cependant, les choses ne sont pas aussi simples.

Commençons par les neurosciences. S'il était littéralement exact qu'elles sont immergées dans les sciences cognitives, on ne comprendrait pas pourquoi le besoin s'est fait sentir, il y a une vingtaine d'années, de forger l'expression « neurosciences cognitives ». La réponse qu'on entend souvent est que les neurosciences cognitives sont précisément la partie des neurosciences qui est immergée dans les sciences cognitives. Mais cette réponse un peu normande appelle une deuxième question : pourquoi ne serait-ce qu'une partie des neurosciences, et non la totalité, qui participe aux sciences cognitives ? en vertu de quoi telle question, telle branche, telle orientation des neurosciences demeureraient-elles en dehors ?

La réponse ne va pas de soi, ce qui est en soi un enseignement : comme il arrive régulièrement lors des premières phases de développement d'une nouvelle spécialité, son domaine co-évolue avec ce qu'elle sait ou pressent de son objet. On dispose cependant de deux éléments de réponse :

1) Les neurosciences cognitives s'intéressent aux propriétés du système nerveux en tant que système assurant des *fonctions de haut niveau*.

2) Les neurosciences cognitives s'intéressent à la manière dont le système nerveux réalise les fonctions *cognitives*, sachant qu'il réalise également des fonctions non cognitives.

La combinaison de ces deux idées fournit la réponse qui était généralement acceptée jusqu'à assez récemment : les neurosciences cognitives étudient les propriétés systémiques du système nerveux responsables de la réalisation des fonctions cognitives. Mais l'évolution scientifique récente a fait perdre à ces distinctions leur netteté initiale : ce qu'est le haut niveau par rapport aux niveaux supposés inférieurs, ce qu'est une fonction cognitive par opposition à une fonction non cognitive, voici qui est moins clair que par le passé. On disait par exemple que les fonctions végétatives, régulant le rythme cardiaque, la respiration, le clignement de paupière ou la digestion, étaient des exemples caractéristiques de ce qui n'est *pas* cognitif : or nous avons appris qu'elles jouent, probablement voire sûrement pour certaines d'entre elles, un rôle dans certains mécanismes cognitifs importants.

Cela étant, même flou, ce critère de démarcation reste utile, à condition d'être compris de manière souple. Sur le terrain, on distingue sans peine des travaux relevant manifestement des sciences cognitives de travaux qui n'en relèvent manifestement pas : les premiers renvoient directement à des phénomènes dont s'occupent les autres branches des sciences cognitives, les seconds n'y renvoient, au mieux, qu'indirectement, et font appel à des disciplines telles que la biologie cellulaire et moléculaire, la neurochimie, la biophysique... qui ne sont pas en interaction fréquente avec les sciences cognitives. Il existe bien des cas douteux, mais cela n'invalide pas le critère.

Pour ce qui est maintenant des rapports entre les neurosciences cognitives et le reste de la « galaxie cognitive », trois remarques doivent être faites :

1) Elles sont en partie jointives avec la psychologie cognitive : tout un ensemble de travaux sont à ce point liés à la fois aux neurosciences et à la psychologie qu'il semble à ceux qui les mènent qu'il n'y a plus de bonne raison de tracer une frontière entre les deux sous-disciplines.

2) Elles comprennent un très important volet d'études animales : non seulement l'animal non humain se prête à des expériences qu'on ne peut faire sur l'homme, mais il donne lieu à des comparaisons qui peuvent être extraordinairement éclairantes. Les neurosciences partagent cette orientation avec les approches évolutionnaires et éthologiques, ce qui renforce leur ancrage dans la « galaxie ».

3) Enfin, elles sont, selon certains chercheurs, bien davantage et finalement encore autre chose qu'une composante de cette galaxie : elles sont selon eux la synthèse longtemps attendue, la fusion sans couture des études sur le cerveau et des études sur l'esprit. Cette fusion se fait sous les auspices et l'autorité des neurosciences, branche de la biologie, mais intègre les concepts et les méthodes des autres disciplines impliquées. En ce sens fort, qui ne fait l'unanimité ni dans la communauté des neurosciences ni dans celle des sciences cognitives, mais est porteur d'une orientation importante en leur sein, les neurosciences cognitives ne sont rien de moins que la nouvelle figure des sciences cognitives, l'évolution en cours devant conduire à terme à l'intégration des autres sous-domaines.

La philosophie donne lieu à un phénomène symétrique, nous allons le voir. Mais il faut d'abord dire pourquoi elle figure parmi les disciplines constituantes, une situation rare, sinon unique, dans les sciences contemporaines : si l'on admet généralement que la philosophie peut collaborer avec différentes sciences sociales, sociologie, économie, sciences politiques ou juridiques..., en se plaçant sur le même plan qu'elles, on s'attend à ce qu'elle prenne une place différente dans ses éventuelles interactions avec les sciences de la nature : elle se situerait en amont (s'intéressant à l'histoire conceptuelle de la discipline) ou en aval (procédant à un examen critique et à une reconstruction rationnelle de ses concepts, méthodes et résultats). En réalité, dans certains domaines, il arrive que s'établisse une collaboration sur un même plan, celui de la clarification et de la résolution des problèmes. Mais il faut reconnaître qu'en pratique les interactions sur le terrain même de l'enquête scientifique sont peu fréquentes et d'étendue limitée. Les sciences cognitives constituent à cet égard une exception tout à fait notable. L'explication est qu'elles se veulent des sciences de la nature, mais appliquées à un domaine qui n'est pas a priori naturel, à savoir l'esprit, ses traces et ses productions, qui sont précisément la cible historique des sciences humaines ; et que ce sont aussi certains des principaux objets de la philosophie, depuis l'aube des temps. Elle apparaît ainsi comme un médiateur privilégié, susceptible d'offrir aux sciences de la nature une prise effective sur la sphère humaine. Et de fait, c'est la philosophie, en particulier la logique, qui en est une branche, mais plus largement toute la tradition rationaliste, qui ont fourni aux sciences cognitives naissantes une bonne part de leurs concepts fondamentaux et de leurs cadres théoriques.

Depuis le début, et à ce jour, des philosophes travaillent au coude à coude avec des chercheurs de chacune des autres disciplines constituantes. Ils ne renoncent pas pour autant au travail caractéristique du philosophe des sciences, qui se situe en principe *après* ou *à côté* du travail scientifique, mais les deux types de recherche sont souvent proches au point de se confondre : philosophie de l'esprit et du langage, philosophie des sciences cognitives,

« philosophie cognitive » sont des activités profondément imbriquées. La dernière locution, de création récente, désigne l'ensemble des recherches philosophiques inspirées par les sciences cognitives ou directement pertinentes pour elles, poursuivant certains de leurs objectifs par des moyens plus abstraits et conceptuels, mais qui sont proches des méthodes scientifiques, en sorte qu'un aller-retour peut s'établir entre les propositions des philosophes et les recherches proprement scientifiques.

La place de la philosophie dans les sciences cognitives est donc assurée, ce qui méritait d'être précisé compte tenu de l'existence d'une conception de la philosophie, très répandue, qui la situe sinon aux antipodes, en tout cas à bonne distance du travail scientifique, et du fait que des interactions directes entre philosophie et sciences de la nature sont rares en dehors du cas qui nous occupe. Mais cela ne répond pas à la question de savoir par où passe la frontière entre ce qui, au sein de la philosophie, est lié aux sciences cognitives (ce que nous venons d'appeler « philosophie cognitive »), et ce qui ne l'est pas. Il n'existe pas de réponse univoque à cette question : les réponses s'échelonnent entre deux extrêmes.

Pour certains philosophes, la philosophie n'a d'autre choix que de se mettre à l'écoute des sciences. C'est ainsi qu'en matière d'ontologie, ce que les physiciens et les cosmologistes ont à nous dire de la constitution du monde est pour le philosophe de la plus haute importance ; de même pour ce que les biologistes ont à nous dire de la nature et de la constitution du vivant. Maintenant que la science a commencé de mettre au jour les bases naturelles de la pensée humaine, les grands domaines de la philosophie qui avaient échappé au regard scientifique – théorie de la connaissance, éthique, philosophie politique, philosophie du droit, esthétique et philosophie de l'art, voire philosophie de la religion... – deviennent à leur tour l'objet d'une enquête qui est à la fois philosophique et scientifique, conceptuelle et empirique. Dans cette perspective *naturaliste*, une partie considérable de la philosophie est directement touchée par les sciences cognitives.

À l'autre extrême se situe un groupe de philosophes, numériquement important (et très largement majoritaire en France notamment), pour lesquels les sciences n'ont que peu de choses à nous apprendre sur les questions philosophiques fondamentales, car celles-ci se situent en amont de toute enquête scientifique, et portent de manière générale sur des sujets qu'aucune science ne peut aborder (penser qu'il en va autrement étant considéré comme relevant du « scientisme »). Dans cette perspective résolument antinaturaliste, les sciences cognitives n'impactent pratiquement pas la philosophie, sinon dans la région, marginale, de la philosophie des sciences et plus particulièrement de la philosophie de la psychologie.

Une position médiane, que je recommande personnellement, consiste à accepter à la fois que la science en général, et les sciences cognitives en particulier, *peuvent* apporter un éclairage sur bon nombre de questions philosophiques, voire modifier en profondeur certaines problématiques ; mais que ce n'est pas toujours *nécessairement* le cas, et que la philosophie, inversement, est parfois en mesure de contester la valeur ou la portée de certaines recherches scientifiques. Dans cette perspective, une partie importante de la philosophie est concernée par les sciences cognitives, sans que cela compromette son autonomie et sa capacité à explorer des idées sur lesquelles la science n'a rien à apporter, du moins autant qu'on puisse en juger pour le moment.



### 3. *L'interaction entre disciplines*

#### a. Le premier stade : la pluridisciplinarité

On l'a dit, une bonne partie des recherches qui nous concernent sont mono-disciplinaires ; si elles nous concernent, c'est parce qu'elles peuvent à tout moment prendre place dans une enquête impliquant d'autres disciplines. Les questions abordées par les sciences cognitives renvoient en effet invariablement à des phénomènes relevant de plusieurs disciplines. Pour prendre un exemple parmi cent (les chapitres qui suivent en sont pleins), si l'on se demande comment un enfant acquiert la compréhension de la suite des numéraux (un, deux, trois...), d'un verbe modal (« il faut... », « on peut... »), ou de l'irréel du présent (« si on avait un ballon, on pourrait jouer au foot sur la plage ») on devra engager des recherches en linguistique, en logique, en psychologie du développement, en neurosciences, voire en philosophie. On constituera ainsi, progressivement et par aller-retour entre ces différents chantiers, d'une part une idée relativement précise de ce que constitue la possession d'un concept, l'acquisition de ce concept, la nature des concepts particuliers en question (numéral, suite de numéraux, verbe, modalité, irréel du présent...), le rapport entre possession d'un concept et maîtrise du terme correspondant, d'autre part un ensemble de mécanismes, caractérisés au niveau psychologique et au niveau neural, réalisant l'acquisition en question (cf. chap. IV « Le développement »).

Dans un cas de ce genre (présenté ici de manière trop schématique), les éclairages apportés par les différentes disciplines se complètent : c'est le même problème que chacune se pose, disposant de sa propre interprétation des données du problème, et auquel elle apporte sa solution partielle. La mise en relation de ces différentes solutions fournit une image du phénomène qui n'est en général pas complètement unifiée, mais qui permet en pratique de progresser sur le plan théorique et parfois sur le plan pratique.

#### b. Le deuxième stade : l'interdisciplinarité

Les sciences cognitives donnent aussi naissance à des situations exigeant une interpénétration plus profonde de deux disciplines ou (plus rarement) davantage. Cela se produit lorsqu'un phénomène se présente, de manière souvent encore indistincte, sans que l'on voie de quelle discipline existante il pourrait relever. Il arrive alors que l'on juge que seule une approche combinant concepts, méthodes, schémas explicatifs de disciplines différentes a des chances de préciser le phénomène et d'en faire la théorie. Ce ne peut se produire qu'au cours de discussions approfondies entre chercheurs de ces disciplines, et une fois la synthèse effectuée, il est difficile de se replacer dans la situation initiale : il semble tout naturel de considérer le phénomène en question comme relevant d'une nouvelle sous-discipline.

Des mariages disciplinaires de ce genre peuvent se produire à différentes échelles, allant de la résolution d'un problème particulier à la constitution d'une nouvelle spécialité scientifique. Donnons-en quelques exemples relevant des sciences cognitives, en notant au passage qu'elles ne sont pas le seul domaine dans lequel ils se produisent.

Au plus haut niveau, les sciences cognitives elles-mêmes constituent, on l'a d'ailleurs indiqué d'entrée de jeu, une entreprise interdisciplinaire. On le constate d'abord à leur objet : la cognition, l'esprit/cerveau, les processus et capacités mentales envisagées sous l'angle de

leurs mécanismes biologiques... autant de désignations dont aucune ne relève d'une seule discipline constituée. On le constate ensuite à leurs concepts fondamentaux : information, computation, représentation, codage, implémentation, modularité, etc., ne sont ni des concepts généraux, partagés par toutes les disciplines antérieurement à leur rencontre, ni des concepts qui appartiennent entièrement à l'une d'elle : ils sont le fruit de cette rencontre. On le constate à leurs méthodes, qui incluent la détermination des contraintes que les mécanismes disponibles à un niveau imposent aux hypothèses concernant les mécanismes aux niveaux adjacents. On le constate enfin à leurs résultats, obtenus ou espérés, qui consistent en une compréhension intégrée ou coordonnée des phénomènes neurocognitifs à différents niveaux, sur différentes échelles de comparaison et selon différentes temporalités.

Un exemple plus restreint est fourni par les travaux en cours sur la conscience et l'attention. Voilà deux concepts extraordinairement difficiles à saisir, bien qu'ils appartiennent à la compréhension commune, et jouent un rôle essentiel dans la pratique (« Était-il conscient du risque qu'il prenait – ou : du crime qu'il commettait ?, de la souffrance qu'il infligeait ? », « Ce patient en coma profond est-il minimalement conscient ? », « Faites attention au lait sur le feu, il risque de déborder... » ; « L'agent de sécurité a-t-il fait attention au voyant d'alarme ? »). Ils ont pendant longtemps été soigneusement évités par toutes les disciplines, sauf la philosophie, qui n'en a pas fait, c'est le moins qu'on puisse dire, une *success story*, à moins de compter comme tel (ce qui, pour la philosophie, ne serait pas inconcevable) le constat, après une longue traversée d'un océan d'écrits, qu'il y a là un mystère insondable. Depuis une vingtaine d'années, la philosophie, les neurosciences, la psychologie, voire la robotique, participent à un effort historique pour circonscrire ces phénomènes avec précision et tenter d'en fournir une description consensuelle et une explication. Les avis divergent sur l'état d'avancement de ce projet, mais son caractère interdisciplinaire ne fait pas de doute (cf. chap. XII « Conscience »).

Une entreprise beaucoup plus avancée est celle de la pragmatique linguistique, branche aujourd'hui établie et prospère, à la frontière de la linguistique, de la logique, de la philosophie du langage et de la psychologie, et l'un des fleurons des sciences cognitives (cf. chap. VIII « Le langage »). Un autre exemple est celui de la décision, un domaine abordé dans plusieurs chapitres du présent ouvrage (cf. chap. IX « Décision et apprentissage », et chap. X « Décision et rationalité »), qui mobilise lui aussi la logique, la philosophie, la psychologie, la théorie des probabilités, l'économie et les neurosciences. Un dernier exemple, également exemplaire, est celui de la vision : les « sciences de la vision » (*vision science*) intègrent les apports de la psychologie expérimentale, de la neuropsychologie et de pratiquement toutes les autres branches des neurosciences, de l'IA et de la robotique, de la philosophie, ainsi que des mathématiques qui jouent ici un rôle particulièrement important (cf. chap. VI « La perception visuelle »).

#### 4. *Le multiculturalisme des sciences cognitives*

Les sciences cognitives ne sont pas seulement intrinsèquement interdisciplinaires : elles sont aussi multiculturelles, ce qu'on ne souligne pas toujours suffisamment. Chacune des

cultures qui s'y entrecroisent, sans toujours se mêler, est définie par un domaine du réel et par un type d'objectif. La liste qui suit fera voir de quoi il s'agit.

Les principales cultures du domaine se divisent en trois groupes :

- Le premier comprend les programmes de recherche scientifique, poursuivis dans le contexte académique. Il se divise à son tour en deux sous-groupes :

- les sciences de la nature : neurosciences et autres branches de la biologie, dont la biologie évolutive ; psychologie expérimentale, psychologie du développement, neuropsychologie, sciences physiques, mathématiques et informatique ;

- les sciences humaines : philosophie (dont la logique philosophique), linguistique, anthropologie, économie ;

- Le deuxième rassemble des programmes d'ingénierie : IA, robotique, facteur humain, ergonomie, internet...

- Le troisième est centré sur la médecine, comprenant celles de ses branches concernées par toutes les formes de déficience mentale, motrice ou perceptive, le vieillissement, les syndromes tels que le spectre autistique, les syndromes envahissants de la personnalité ; la neurologie, la génétique ; les spécialités paramédicales afférentes : bioingénierie des prothèses, rééducation, éducation spécialisée...

Depuis une dizaine d'années, un quatrième groupe commence à se former autour du thème de la décision, individuelle et collective : la culture correspondante est celle des sciences politiques, de l'économie, du droit et des sciences sociales appliquées.

Pourquoi introduire cette typologie des cultures, s'ajoutant à celle des disciplines ? Parce qu'elle éclaire des différences d'attitude, d'ethos professionnel, ainsi que des affinités qui traversent, précisément, les frontières entre disciplines. Une culture constitue à la fois un espace de référence commun pour ceux qui s'y reconnaissent et une façon de se différencier, et parfois de s'opposer à ceux qui s'identifient à une autre culture. Les cultures correspondent assez bien aux divisions de l'enseignement supérieur : les facultés des sciences et des lettres et sciences humaines, pour le premier groupe ; les grandes écoles d'ingénieur, pour le deuxième ; les facultés de médecine et écoles paramédicales, pour le troisième ; le quatrième groupe, en cours de constitution, correspond aux instituts d'études politiques, aux facultés de droit et aux écoles de management.

Il s'agit d'une réalité qui ne s'inscrit pas dans les organigrammes officiels, mais qui joue un rôle sociologique et psychologique important et, à mon sens, très positif. Les sciences cognitives sont ainsi protégées des risques de l'enfermement dans l'une ou l'autre des cultures : sans renier sa formation et ses inclinations, chaque participant se frotte à des collègues qui s'intéressent à des questions semblables ou proches, mais dans une perspective très différente, ce qui le protège, dans une certaine mesure, de l'ossification dans une micro-spécialité, et lui ouvre une palette beaucoup plus large d'objectifs possibles et de collaborations (sans parler, pour les nouveaux diplômés, de débouchés...).

Ce multiculturalisme ne va pas sans difficultés : comme on va le voir, les sciences cognitives sont traversées par des divergences de toutes sortes, dont certaines ont pour origine des différences plus profondes que de simples désaccords sur les faits ou sur les méthodes : ce sont justement des différences de culture, et d'en prendre conscience peut aider à les surmonter.

#### IV. LA SCIENCE EN MARCHÉ : PROGRAMMES, ECOLES, CONTROVERSES, INCERTITUDES...

Il a été question, ici et là, d'incertitudes, de doutes et de débats. Cette dernière section a deux objectifs. Le premier est de rappeler, rapidement, le fait que c'est le cas de tout domaine scientifique, et les raisons pour lesquelles on n'y prête généralement pas attention. Le second objectif est de présenter les principales controverses au sein des sciences cognitives, ainsi que les raisons de ne pas s'en inquiéter outre mesure, et d'y voir surtout des sources de tension productives.

##### *1. Moments et visages de la science*

La science constituée apparaît comme un monument résistant au temps et aux outrages de la bêtise, et le processus scientifique comme un fleuve parfois turbulent mais cheminant avec détermination vers l'océan des faits établis. Cette imagerie est trompeuse ; elle fait beaucoup de tort à la science, et donne des armes à toute une industrie anti-science, de bas ou de haut étage.

La science<sup>17</sup> ne sort pas toute armée de la cuisse de Jupiter : c'est évident. Chacun sait qu'elle résulte d'efforts soutenus, parfois héroïques, de la part d'individus intelligents, parfois géniaux, qui se soutiennent, travaillant souvent en équipes et écoles, et se relaient d'une génération à la suivante ; et qu'elle se perfectionne avec le temps : les premières découvertes sont le plus souvent imparfaites, entachées d'erreurs ou de complications finalement inutiles, d'application trop restreinte, etc. D'autre part, leur genèse peut être en partie le fruit du hasard, en partie celui d'intuitions ou de paris sans base rationnelle ; et qu'elles ne deviennent des résultats scientifiques qu'au prix d'une validation selon les normes pertinentes.

Pour tenir compte de ces aspects, les philosophes des sciences ont introduit deux notions. La première est la faillibilité : les faits et théories produits par la science ne sont jamais à l'abri d'une révision plus ou moins drastique, pouvant aller jusqu'au rejet pur et simple. La seconde est la distinction entre contexte de la découverte et contexte de la justification : la manière dont un scientifique découvre un fait ou conçoit une théorie, les conditions sociales, économiques et politiques qui lui sont faites, tout cela est dépourvu de portée théorique (sauf pour qui s'intéresse à l'histoire et à la sociologie des sciences ou encore à la créativité) ; ce qui compte, et à quoi s'appliquent les normes strictes de la rationalité scientifique, ce sont les résultats eux-mêmes : mesures, faits, théories... En conjuguant ces deux idées, on aboutit à une représentation de la science à la fois plus réaliste que celle d'un monument inaltérable et

---

17. Certains philosophes des sciences accordent beaucoup d'importance à la question de savoir s'il faut parler de *la* science ou *des* sciences. S'agissant uniquement, dans beaucoup de contextes, tel que celui-ci, d'une question verbale, il n'y a pas lieu de s'y arrêter. Si le lecteur préfère le pluriel, il rectifiera de lui-même. Par-delà la convention grammaticale, il y a une vraie question, celle de l'unité des sciences (de toutes les sciences), et pour ce qui nous concerne ici, celle de l'unité des sciences cognitives, dont il a été assez longuement question dans la section 3.

moins fragile que celle d'une construction sociale précaire à la merci des hasards, des opinions divergentes et des ambitions personnelles des hommes et des femmes de science, et de l'influence de leurs réseaux et commanditaires. En effet, les produits de la science ont une base objective, constituée par le réseau de leurs relations internes et de leurs justifications, parmi lesquelles les résultats expérimentaux jouent un rôle déterminant. Ils restent perfectibles, mais ne cèdent leur place dans le stock des connaissances humaines que lorsqu'ils sont remplacés ou améliorés de quelque façon. Ils constituent à tout moment la meilleure approximation disponible de la vérité (ou si l'on préfère la meilleure représentation du fragment de la réalité dont ils parlent).

Cette image de la science est elle-même une bien meilleure approximation de la réalité que les deux modèles qu'elle supprime. Elle demeure néanmoins trop fruste. Tout particulièrement dans le cas d'un domaine jeune et interdisciplinaire, deux aspects ou moments de la démarche jouent un rôle important, le premier tourné vers l'avenir, le second, en un sens, vers le passé.

À tout moment de son développement, une science est animée par une certaine vision de ce qu'elle veut accomplir, des moyens qui le lui permettront, et des formes que prendront ses résultats. Cette vision peut être plus ou moins ambitieuse, plus ou moins précise, embrasser un avenir plus ou moins lointain, et se réclamer de succès plus ou moins éclatants. Mais elle n'est jamais une pure extrapolation de résultats déjà acquis : elle va nécessairement au-delà de ce que la science en question peut justifier. En d'autres termes, la science est guidée par des idées qui sont en partie spéculatives ou métaphysiques. Ces idées font parfois l'objet d'exposés systématiques, de manifestes, à usage soit interne, lorsqu'un scientifique ou (cas le plus fréquent) un petit noyau de chercheurs tente de mobiliser des collègues et des étudiants d'une discipline, ou de plusieurs disciplines, autour de leur vision ; soit externe, lorsqu'une avant-garde scientifique veut signaler l'émergence d'une nouvelle approche ou d'une nouvelle discipline, pour l'information d'un large public mais aussi pour obtenir les moyens matériels et institutionnels qu'elle juge nécessaires pour réaliser son projet. Plus un domaine est jeune, plus la vision qu'il met en avant peut sembler téméraire : les manifestes scientifiques soulèvent presque nécessairement des objections, et paraissent souvent reposer sur des bases fragiles, sur des espoirs sans fondement voire des exagérations délibérées quant aux premiers résultats obtenus. On parle alors de « battage » (*hype* en anglais) – ce qui se justifie a posteriori lorsque le programme s'effondre (des exemples parmi d'autres sont ceux de la thérapie génique ou des systèmes experts<sup>18</sup>), mais qui peut tout aussi bien déboucher sur une authentique percée (la théorie germinale de la maladie, associée aux noms de Pasteur et de Koch, ou plus près de nous la neuro-imagerie). Mais entre le battage et la simple audace la distance n'est pas grande : les chercheurs de disciplines galonnées peuvent bien lever un sourcil désapprobateur ou ironique devant les déclarations des pionniers, ils oublient qu'eux-mêmes doivent leur existence à des pionniers qui en leur temps semblaient tout aussi fantasques à leurs contemporains. Cela ne signifie pas que tout programme audacieux soit intelligent et destiné à réussir – au contraire, les réussites sont l'exception –, ni qu'il faille

---

18. Je distingue les cas d'exagération, dans lesquels un noyau de validité subsiste, quoique de bien moindre portée qu'annoncé, des cas d'erreur ou de tromperie scientifique, telles que les « rayons N » ou la « mémoire de l'eau ».

s'abstenir de critiquer ce qui se donne pour une profonde intuition de l'avenir. Mais qu'il ne faut pas s'étonner ni s'alarmer du fait que les sciences cognitives aient souvent été présentées dans des manifestes de ce genre, dont la tonalité générale, l'argumentation, la rigueur ne sont décidément pas celles d'un article ou d'un traité d'une science constituée. Il faut de même prendre avec équanimité le fait que ces manifestes continuent d'être produits, alors que les sciences cognitives ont largement franchi la soixantaine, et qu'entre tous ces manifestes, d'hier et d'aujourd'hui, les désaccords soient la règle, chacun cherchant à se démarquer des précédents : les sciences cognitives ne parlent pas d'une seule voix, et leurs diverses voix muent au cours du temps.

Ainsi, avant la phase de justification des théories, et même avant la phase de la découverte plus ou moins hasardeuse, se place une phase franchement spéculative, programmatique et sans contenu scientifique assignable, celle des visions guidant un programme de recherche, la création d'un nouveau champ disciplinaire, d'une nouvelle approche. Mais, à l'opposé, il existe, *après* la phase de justification rigoureuse, une phase de *reconstruction rationnelle*, de systématisation et de mise en cohérence de l'ensemble des résultats. Le fruit de cette phase est ce que le philosophe des sciences américain Thomas Kuhn (1922-1996) appelle la science des manuels, et qu'il juge presque aussi éloignée de la science qui se construit (celle de l'expérimentation, des mesures et des justifications) que cette dernière l'est des manifestes et des premières découvertes. La science des manuels a une fonction essentielle, celle de former les étudiants et de fournir aux non-spécialistes des informations utilisables dans leur propre champ. Elle ne peut la remplir qu'au prix de simplifications et d'omissions volontaires d'informations qui risqueraient de brouiller la vue des lecteurs sans bénéfice pour eux. Elle procure ainsi une représentation utile mais déformée de la science qui se fait ; en particulier, elle réduit les dissonances et s'abstient de parler des failles, contradictions et incertitudes qui sont la signature de la science vivante.

S'agissant d'une science mûre telle que la physique ou ses différentes branches, ou des parties les plus solides de la biologie, cette opération de simplification-réduction-lissage obéit à des contraintes objectives suffisamment fortes pour que le résultat n'en soit pas fortement affecté : les manuels, dans ces disciplines, présentent des similarités bien plus fortes que leurs différences (auxquelles leurs auteurs attachent tant d'importance). C'est pourquoi on estime généralement possible, dans beaucoup de contextes, de faire abstraction de la distinction entre science des manuels et science vivante. Mais le cas des sciences cognitives est différent : leur jeunesse et leur interdisciplinarité font de toute réduction une opération fortement dépendante des sensibilités des auteurs et des (jeunes) traditions auxquelles il se rattachent. Ce serait un bien mauvais service à rendre aux étudiants et aux non-spécialistes de sciences cognitives que de le dissimuler. Aucun traité ou manuel de sciences cognitives, et celui-ci ne fait pas exception, ne peut prétendre fournir autre chose qu'une vue *partielle*, non seulement en ce qu'elle ne peut être exhaustive, mais parce qu'elle reflète nécessairement une certaine manière de considérer le champ dans son ensemble : en ce sens, toute présentation des sciences cognitives est inévitablement *partiale*. Si le présent ouvrage se distingue, c'est d'abord parce qu'il ne cherche pas à dissimuler son caractère partiel et provisoire, ensuite parce qu'il s'efforce de présenter des points de vue divers, sur différentes questions, qui collent davantage aux recherches en cours qu'un panorama général, enfin parce qu'il tente

d'établir des passerelles entre ces différents points de vue. Le lecteur est en tout cas dûment averti du problème.

## 2. Les débats internes : un aperçu

Il n'est pas de science sans controverses, et celles-ci ne se terminent pas toujours rapidement, à la manière d'une échauffourée dans une cour de récréation : certaines controverses durent, certaines se figent même en un état de dissensus permanent, qui ne se résout jamais mais finit par se dissoudre à mesure que la discipline se transforme et que les générations se renouvellent. Les sciences cognitives ne font pas exception : au contraire, et toujours pour les mêmes raisons, elles sont le lieu de controverses particulièrement vives, qui les agitent davantage que les disciplines mûres.

On peut distinguer, en première analyse, deux types de controverses. Les controverses du premier type restent localisées à un domaine restreint, du ressort d'une discipline ou d'un programme de recherche. Elles ne constituent pas une menace, elles sont le moteur naturel de tout domaine scientifique. Les exemples en sciences cognitives sont innombrables, et on en trouvera dans les différents chapitres. Les controverses du second type sont globales : elles mettent en cause les fondements d'une discipline entière, les prémisses d'un programme de recherche, ou encore l'architecture générale d'un domaine composite, tel précisément que celui des sciences cognitives. C'est de ces dernières qu'il faut dire quelques mots en conclusion de ce chapitre, parce qu'elles peuvent conduire le nouveau-venu à craindre qu'en suivant telle formation plutôt qu'une autre, en lisant ce manuel-ci et non celui-là, on risque de passer à côté de l'essentiel, voire de s'égarer complètement. Je présenterai donc très rapidement ce qui m'apparaît comme les deux principaux terrains d'affrontement, avant de dire pourquoi la vie scientifique, de la formation initiale aux recherches en cours, n'est généralement pas affectée par ces divergences.

### a. La place des disciplines constituantes

Les disciplines et spécialités qui se rattachent aux sciences cognitives entretiennent des rapports de coopération et de complémentarité et sont engagées dans des collaborations plus ou moins étroites. Mais chacune semblerait par définition maîtresse de son propre domaine : le linguiste a autorité sur ce qui touche au langage, le neuroscientifique sur ce qui touche au cerveau, le roboticien sur ce qui touche aux robots, et ainsi de suite. En réalité, ce partage des tâches et des compétences ne va pas toujours sans difficulté, du fait que la raison d'être de la fédération des sciences cognitives est l'existence d'un noyau commun, et qu'aucune des disciplines constituantes n'est a priori mieux placée que les autres pour accéder à ce noyau. Il se pourrait même que seule une discipline nouvelle en ait les moyens, ou du moins qu'elle détienne une clé sans laquelle les efforts des autres composantes seraient insuffisants. Traditionnelle ou nouvelle, la discipline qui détient ou croit détenir une telle clé revendique une place centrale dans le dispositif ; la métaphore souvent utilisée est celle du « navire amiral » (*flagship*).

Historiquement, plusieurs candidatures à la dignité en question se sont manifestées. La première en date est la cybernétique, mais la situation était particulière : les sciences

cognitives ne s'étaient pas encore constituées, et la cybernétique, englobant et s'efforçant de coordonner différents programmes de recherche en émergence, était sans concurrent, tout en ayant la dimension réduite d'une avant-garde, à peine visible dans le paysage scientifique de l'époque. La phase suivante a été celle de l'intelligence artificielle (IA), qui a assumé le rôle de leader, rôle qu'on ne lui a d'abord pas contesté, mais qu'elle a bientôt perdu sous l'effet conjugué de ses échecs et de la montée en puissance d'autres programmes et disciplines. Depuis l'arrivée de l'imagerie fonctionnelle cérébrale, coïncidant avec d'autres avancées en neurosciences, ce sont les neurosciences cognitives qui revendiquent le statut de navire amiral. Un dernier cas, un peu différent, est celui de la biologie évolutive. L'approche évolutionnaire est considérée, par ses partisans, comme fournissant justement une clé sans laquelle on ne peut progresser sérieusement (cf. chap. v « Évolution et cognition ») ; pour autant, elle ne prétend pas constituer à elle seule le navire amiral de la flottille : elle se conçoit davantage comme un corps d'élite dont chaque navire se doit d'embarquer un bataillon.

Ces revendications ne sont pas ou pas seulement des querelles corporatistes. Elles sont parfois alimentées par une méconnaissance de l'histoire et des bases théoriques des sciences cognitives. Mais elles correspondent surtout à des thèses portant soit sur la nature des phénomènes, soit sur la stratégie à mettre en œuvre pour en rendre compte. Sans pouvoir exposer ces thèses et discuter de leurs mérites respectifs, disons par exemple que l'IA défendait l'idée que la cognition est essentiellement une affaire de traitement de l'information, en donnant à ces termes un contenu précis, et que la meilleure méthode pour la comprendre est de soumettre systématiquement les tâches cognitives accomplies par l'homme à une analyse conceptuelle afin d'en construire des simulations informatiques, à rectifier et améliorer à la lumière des performances réelles de ces programmes. De manière également sommaire, l'idée maîtresse des neurosciences cognitives (lorsqu'elles assument le rôle de navire amiral) est que la cognition est essentiellement constituée des fonctions supérieures du cerveau, fonctions conçues comme gouvernant le comportement, et que c'est l'étude de ces fonctions telles que le cerveau les réalise qui mène le plus directement à la compréhension de la cognition. L'approche évolutionnaire est fondée sur l'idée que ces mêmes fonctions ne peuvent être expliquées seulement par leurs causes *proximales* (la manière dont elles sont physiologiquement réalisées dans le cerveau), pas plus qu'on ne peut expliquer un ordinateur en considérant seulement les processus électroniques ni même les processus informationnels dont il est le siège. Il faut enquêter sur les causes *distales*, c'est-à-dire sur les contraintes sélectives qui ont historiquement conduit à l'inscription dans le patrimoine génétique de l'espèce des processus physiologiques en question (pour l'ordinateur, l'analogie serait constitué par les processus ayant conduit à sa conception).

Il est impossible de rendre justice à ces questions dans le cadre du présent chapitre, mais on devine que les positions que je viens d'esquisser, sans être strictement incompatibles, peuvent néanmoins donner lieu à de sérieuses divergences. Et c'est ce qui s'observe : hier l'IA, aujourd'hui les neurosciences cognitives ou la psychologie évolutionnaire dans ce qu'on appelle parfois, de manière polémique, leurs ambitions *impériales*, heurtent de front les convictions de nombreux participants, psychologues, linguistes, anthropologues, roboticiens, philosophes ; sans compter les neuroscientifiques lorsque l'IA et ses alliés philosophes ou psychologues prétendaient occuper le centre du terrain, et inversement. Les oppositions ne



viennent du reste pas seulement des disciplines ainsi « vassalisées » ; elles s'expriment aussi au sein des disciplines « reines », qui abritent elles-mêmes des conceptions très différentes de ce que sont véritablement les sciences cognitives, ou de ce dont elles sont véritablement la science.

#### b. Le classicisme et ses opposants

Cependant, le débat le plus fondamental et le plus visible aujourd'hui est celui qui oppose les défenseurs et les critiques d'une conception qu'on appelle « classique » ou parfois « symbolique ». C'est celle qui a été proposée, illustrée et conceptuellement fondée au cours de la première grande phase des sciences cognitives, et qui a été présentée de manière très schématique ci-dessus (cf. § II.1), comme le cadre conceptuel ayant permis l'émergence du domaine des sciences cognitives. On voit pourquoi elle est dite « classique » ; elle est aussi « symbolique » parce qu'elle fait de la cognition un système d'opérations sur des représentations symboliques : symboles matériels qui sont aussi des réceptacles ou véhicules de significations renvoyant, pour le dire trop vite, au monde extérieur. On parle aussi parfois de « cognitivisme », mais cette étiquette donne lieu à des malentendus, dans la mesure où les spécialistes de sciences cognitives sont parfois appelés les « cognitivistes », tout à fait indépendamment de leurs options théoriques.

Cette conception n'a évidemment pas fait l'unanimité à l'extérieur des sciences cognitives en émergence. Mais en leur sein elle a également été contestée, et en réalité assez tôt, dès la fin des années 1970. Depuis ce moment, les critiques se sont multipliées, accompagnées de contre-propositions ambitieuses, sous forme de nouveaux « paradigmes » chacun fondé sur une propriété ou une dimension particulière, que le classicisme est réputé négliger. Ces idées relèvent souvent de la phase préscientifique, heuristique évoquée un peu plus haut : elles n'ont pas nécessairement une traduction en forme d'hypothèse testable ; c'est pourquoi certains chercheurs se refusent à leur accorder beaucoup d'importance. Il est cependant utile, du moins je le pense, et j'y reviendrai, de signaler leur existence aux nouveaux-venus.

Les principales dimensions qu'auraient, selon ses critiques, oubliées les défenseurs du programme classique sont les suivantes :

- L'expérience : les compétences d'un système cognitif proviennent de sa capacité à apprendre à partir des cas auxquels il a été exposé, et à traiter un nouveau cas par référence aux cas voisins déjà rencontrés (une autre manière de désigner la dimension manquante est donc l'apprentissage) ;

- Le corps : la cognition est « incarnée », elle est inextricablement liée au corps dont elle est un équipement et qui réciproquement en est un pour elle. On ne peut donc pas séparer la cognition du corps : c'est là tomber dans une vision « intellectualiste » qui fait du corps une manière de robot au service d'un système désincarné de traitement de l'information, lui procurant des informations sensorielles et exécutant les mouvements qu'il a décidés au terme de sa cogitation ;

- L'action : une autre forme d'intellectualisme consiste à interposer entre la stimulation sensorielle et l'action une phase d'élaboration des sensations et de planification ; en réalité, la perception et l'action forment une unité indivisible, la cogitation (les formes classiques de la cognition) n'intervenant que dans des situations particulières, à étudier dans un second temps à la lumière de ce qu'on aura compris du phénomène primordial perception-action ;

- La situation : l'individu, comme du reste tout animal non humain, est invariablement et dès les premières minutes de son existence, impliqué dans des situations concrètes, qui appellent à la fois interprétation, décision et action ; toute tentative pour comprendre la cognition à l'image d'une forme de résolution de problème indépendante du contexte, de l'engagement, est vouée à l'échec, car notre appareil cognitif a évolué précisément de manière à nous permettre de gérer des situations, non de résoudre des problèmes d'école ou de réagir dans les situations artificielles et étroitement contrôlées du laboratoire ;

- L'environnement : la cognition n'est pas un phénomène strictement interne à l'individu (en particulier elle n'est pas le produit d'un système nerveux central, limité à la boîte crânienne) – elle peut se produire à l'interface d'un tel système et de son environnement, lequel contient notamment d'autres systèmes de ce genre ainsi que des artefacts et des repères matériels. Cette perspective prend deux formes qui peuvent être adoptées séparément ou ensemble (cf. chap. III « De l'individu au collectif ») :

- selon la première, la cognition ne se limite pas aux processus internes à l'individu : elle se produit *aussi* ailleurs, à l'occasion de certaines interactions entre agents ou systèmes cognitifs et environnement ; ici encore, mais pour une raison différente de l'objection précédente, explorer la cognition dans l'environnement confiné d'un laboratoire est une mauvaise stratégie : la bonne unité d'analyse est le processus d'interactions entre ces différentes composantes ;

- selon la seconde, plus radicale, il n'existe tout simplement pas de cognition interne, car seule l'interaction avec l'environnement fournit un contenu informationnel (on dit parfois sémantique) aux processus internes, contenu en l'absence duquel ils ne sont que des mouvements de matière ;

- Les émotions : l'approche classique, suivant en cela une certaine tradition rationaliste en philosophie, voit dans les émotions (désignées parfois collectivement comme l'affectivité) un facteur bien distinct de la cognition, qui n'interviendrait dans la vie mentale que pour modifier les désirs et la motivation. Elle serait une source de perturbations, tant dans la vie que dans la recherche sur la cognition, qu'il importe d'éliminer autant que faire se peut. La critique consiste ici à rejeter cette conception en bloc : l'affectivité est en réalité une composante essentielle, et le plus souvent utile, de la cognition, et en faire abstraction est un contresens, comme le montrent non seulement l'expérience quotidienne et l'analyse conceptuelle, mais aussi des considérations évolutionnaires (cf. chap. XIII « Émotions »).

Ces différentes perspectives « hétérodoxes », qui imputent à l'« orthodoxie » un « oubli » capital, croisent des objections qui portent sur la nature même des mécanismes ou processus fondamentaux de la cognition. Le classicisme propose d'y voir des computations, au sens de la théorie des algorithmes, exécutées par une variété indéfinie de mécanismes dont le détail est pour l'essentiel sans importance (de même que l'addition est ce qui compte en arithmétique, non la manière dont elle est exécutée par une machine électronique, un humain équipé d'un crayon ou d'un boulier, un système hydraulique ou autre). Les conceptions rivales comprennent les différentes variétés de connexionnisme, unies par le choix du réseau de neurones comme modèle mécanique fondamental des processus cognitifs, et le dynamicisme, qui se prononce à la fois sur le cadre dans lequel il convient d'étudier les systèmes cognitifs (à savoir la théorie mathématique des systèmes dynamiques) et sur la

nature de la cognition elle-même. Pour compliquer encore, mais aussi enrichir cette situation déjà confuse, beaucoup d'auteurs développant ces idées se réclament de penseurs de générations antérieures et d'inspiration nettement différente de celle des pionniers des sciences cognitives, pères du classicisme : ils invoquent souvent des psychologues et neuroscientifiques de l'école russe, de Nikolaï Bernstein (1896-1966) à Vygotski, des pragmatistes américains tels que le philosophe américain John Dewey (1859-1952), des phénoménologues tels que les Allemands Edmund Husserl (1859-1939) et Martin Heidegger (1889-1976) ou le Français Maurice Merleau-Ponty (1908-1961), ainsi que des théories physiques ou mathématiques relatives aux phénomènes de masse, aux systèmes auto-organisés et à l'émergence.

De toutes ces formes de contestation et de toutes ces perspectives, qui sont, selon qui les défend (chacune ayant plusieurs variantes) compatibles, convergentes ou elles-mêmes rivales, se dégagent un certain nombre de programmes identifiés par des étiquettes, parfois bien difficiles à traduire, telles que « *embodied* », « *embedded* », « *enacted* », « *extended* », « *affective* » (regroupées sous le label « 4EA »), ou encore « *situated* », « *distributed* », « *ecological* » et « *externalist* ». Ce qui les rassemble, c'est une conviction très profonde que le classicisme est insuffisant, et même, pour beaucoup, qu'il a fait fausse route.

### c. Quelle attitude adopter ?

Il n'est pas question de présenter ici de manière systématique ces différentes orientations, encore moins de défendre l'une ou l'autre. Il importe en revanche d'aider le lecteur à prendre la mesure de ce que leur existence implique : que devons-nous faire, vers où ou vers qui se tourner si la situation est à ce point confuse et conflictuelle ? Faut-il choisir l'une des nouvelles hétérodoxies ? Faut-il en tout cas rejeter l'orthodoxie, et se donner pour objectif de déterminer la meilleure alternative ? Faut-il au contraire s'en tenir à l'orthodoxie ? Ma réponse est qu'aucune de ces réactions ne s'impose. Voici pourquoi.

Il faut d'abord savoir que la dite « orthodoxie » ne se conçoit nullement comme monolithique : elle abrite une grande variété d'options théoriques mises en œuvre dans des programmes de recherche concrets, et elle est compatible avec une variété encore plus grande d'options possibles. Aussi ses partisans contestent-ils qu'il soit conceptuellement impossible d'intégrer dans ce cadre la plupart des dimensions ou aspects avancés par leurs adversaires comme des objections fondamentales. Mais ils rejettent aussi certaines de leurs thèses : il ne leur paraît nullement évident qu'il faille faire abstraction de la barrière de la boîte crânienne, par exemple, ni que le fait, peu contestable, que l'individu agit au quotidien dans des situations concrètes d'engagement invalide toute étude en laboratoire de certaines capacités de base ; ni que la perception soit totalement dépendante de l'action ; et ainsi de suite. Malgré l'attrait qu'exercent certaines des critiques et des intuitions « hétérodoxes », la conclusion radicale – le rejet du classicisme, et son remplacement par l'une ou l'autre des nouvelles conceptions proposées – n'est pas acquise : l'issue est très incertaine.

Mais est-il bien certain que cette issue prendra nécessairement la forme du choix d'un des cadres en présence ? Quel est, au fond, le statut de ces cadres ? Ils ne sont guère que des tentatives pour donner une forme générale à un très vaste ensemble de phénomènes dont l'unité n'est pas établie, accompagnées de propositions assez vagues sur la manière de procéder. Les recherches sur des phénomènes particuliers, saisis sous un angle précis qui

produisent la plus grande partie des résultats relativement solides du domaine, ne sont généralement assujetties à aucun de ces cadres généraux : elles n'ont besoin que de notions *locales* d'information, de représentation, de computation, de dynamique, etc. Leurs auteurs n'ont pas l'obligation d'en fournir une définition précise, et encore moins de défendre l'idée que ces notions s'appliquent à tout phénomène cognitif. Cet « opportunisme » doctrinal est parfaitement légitime : c'était aussi celui des mécaniciens médiévaux, ou celui des astronomes de la Renaissance, qui se passaient d'une notion générale de force, capable d'unifier la mécanique et l'astronomie. Dans leur majorité, les chercheurs en sciences cognitives se tiennent à l'écart de ces controverses, ou n'y prennent part qu'en marge de leur activité proprement scientifique.

Cela ne signifie nullement que la quête d'un cadre général, ou de fondements d'une science de la cognition, soit vaine, et les débats actuels dépourvus de sens : la plupart des questions posées sont de bonnes questions, auxquelles il faudra bien trouver tôt ou tard une réponse satisfaisante. L'erreur serait seulement de penser que leur caractère fondamental implique qu'elles sont pertinentes pour la totalité des recherches et résultats actuels : il suffit de jeter un coup d'œil sur les disciplines mûres pour se rendre compte que ce n'est jamais le cas. Autrement dit, le chercheur en sciences cognitives, et *a fortiori* l'étudiant, peuvent à la fois s'intéresser aux questions de fondements et considérer que leur approfondissement et leur résolution éventuelle n'est pas une tâche qui leur échoit. Ils peuvent sans mauvaise conscience se concentrer sur des questions particulières. Et même s'ils s'intéressent aux questions générales de fondement, la bonne manière de les aborder passe par une longue confrontation avec un problème local formulé dans les termes d'une tradition de recherche vivante et productive.

#### POUR ALLER PLUS LOIN

ANDLER, Daniel, « Turing : pensée du calcul et calcul de la pensée », in NEF, Frédéric, VERNANT, Denis (dir.), *Les Années 1930. Réaffirmation du formalisme*, Paris, Vrin, 1998 (disponible sur son site <http://andler.dec.ens.fr/>).

BECHTEL, William, GRAHAM, George (dir.), *A Companion to Cognitive Science*, Oxford, Blackwell, 1999 (en particulier l'article introductif).

CLARK, Andy, *Being There. Putting Brain, Body, and World Together Again*, Cambridge (MA), The MIT Press, 1997.

DENNETT, Daniel C., *Brainstorms. Philosophical Essays on Mind and Psychology*, Cambridge (MA), The MIT Press, 1978.

FODOR, Jerry, *Representations*, Cambridge (MA), The MIT Press, 1981.

FODOR, Jerry, *The Modularity of Mind. An Essay on Faculty Psychology*, Cambridge (MA), The MIT Press, 1983 ; trad. fr. par Abel Gerschenfeld, *La Modularité de l'esprit. Essai sur la psychologie des facultés*, Paris, Minuit, 1986.

HAUGELAND, John (dir.), *Mind Design. Philosophy, Psychology, Artificial Intelligence*, Cambridge (MA), The MIT Press, 1981 ; 2<sup>e</sup> éd *Mind Design II*, 1997.

HEBB, Donald O., *Essay on Mind*, Hillsdale (NJ), Lawrence Erlbaum, 1980.

KIM, Jaegwon, *Philosophy of Mind*, Boulder, CO, Westview Press, 3<sup>e</sup> éd, 2010.

RUMELHART, David E., MCCLELLAND, James L., THE PDP RESEARCH GROUP, *Parallel Distributed Processing. The microstructure of cognition (vol. 1 et 2)*, Cambridge (MA), The MIT Press, 1986.

TURING, Alan M., « Computing Machinery and Intelligence », *Mind*, vol. 59, n° 3, 1950, p. 435-450 ; repris dans COPELAND, Jack (éd.), *The Essential Turing*, Oxford, Clarendon Press, 2004, chap. 11.