

NOS PENSÉES NE SONT PAS DANS NOTRE TÊTE

à propos de *La denrée mentale* de Vincent DES-COMBES, Éditions de Minuit. (Critique), 1995.

Où se situent nos pensées et nos activités mentales (croire, vouloir, espérer...) ? Dans notre tête, et plus exactement dans notre cerveau, répondent les adeptes des sciences cognitives. Forts de cette idée, et au vue des grands progrès que l'on a fait récemment dans la connaissance de cet organe, ces derniers pensent être à même, dans un avenir proche, de comprendre comment est produite la pensée. Si elle est « sécrétée » par le cerveau comme la bile l'est par le foie, il suffit en effet de bien décortiquer cette matière cérébrale pour comprendre la genèse et le fonctionnement de l'activité mentale. Penser, croire, vouloir... sont autant de phénomènes qui vont bientôt pouvoir s'expliquer en termes de physique et de biologie. Ce n'est certes pas la première fois que l'on annonce la possibilité d'une telle naturalisation de l'esprit — cette idée est depuis longtemps courante chez tous les matérialistes —, mais les efforts faits en ce sens de nos jours seraient enfin, au dire des protagonistes, sur le point d'aboutir.

Il se peut toutefois que l'on attende longtemps, très longtemps même. Non pas parce que les cognitivistes risqueraient d'être confrontés à des problèmes techniques insoupçonnés — ce qui est toutefois fort probable —, mais tout simplement parce que leur objet d'étude — l'esprit — ne semble pas se situer là où ils pensent. Plutôt que d'être localisés dans le cerveau, les phénomènes mentaux se situeraient en effet dans le monde extérieur, c'est-à-dire que nous ne penserions pas dans notre tête mais dans l'espace public. C'est du moins la thèse que défend Vincent DES-COMBES dans ce livre très critique envers les pré-supposés des sciences cognitives (voir [sommaire](#)). Et si on le suit, le projet de naturaliser l'esprit apparaît effectivement sans objet. C'est un peu comme si vous vouliez comprendre les règles du football en regardant *dans* la tête des joueurs, alors qu'il semble plus judicieux de les interroger et d'aller les voir jouer sur le terrain.

Pour clarifier ces idées, commençons par une simple expérience de pensée. Imaginons que monsieur Dupont veuille aller à la banque ; son cerveau se trouverait, selon les cognitivistes, dans un état physique correspondant à ce désir, c'est-à-dire que la configuration particulière des neurones et des autres entités matérielles de son cerveau expliquerait ce désir d'aller à la banque. Imaginons maintenant, qu'il y a bien

longtemps, un homme de la préhistoire ait été frappé par la foudre de telle sorte que, sous l'effet de la secousse, son cerveau se soit retrouvé dans la même configuration que celui de Monsieur Dupont. Faudrait-il en conclure que l'homme préhistorique eut le désir d'aller à la banque ? Ceci paraît difficile. Et cette difficulté semble indiquer que l'attribution d'un contenu à un acte mental ne peut pas se faire sans une prise en compte du contexte ; c'est-à-dire qu'un phénomène mental (ici un désir) semble bien ne pas relever de la simple sphère privée des individus, mais de l'espace où ces derniers évoluent. Ce qui montre aussi que ce n'est pas parce que l'on peut, en faisant un effort, *intérioriser* nos pensées qu'elles pourraient finir par être des entités « enfermées » dans notre tête.

D'ailleurs cette thèse que les phénomènes mentaux se situent uniquement dans la tête s'accommode mal avec nos usages linguistiques. Nous ne pourrions pas, en effet, dire d'un livre qu'il contient des idées intéressantes ; ce ne serait au mieux qu'une façon de parler, puisqu'en toute rigueur, il faudrait dire que les idées sont uniquement dans la tête de celui qui écrit ou lit le livre. On ne devrait pas dire, non plus, qu'une personne pense, mais plutôt que c'est son cerveau qui pense. Certes, ceci n'ajoute ou n'enlève aucune valeur à la thèse en question. Mais si l'on tient compte du fait qu'il nous faudrait toutefois continuer à dire que c'est la personne physique tout entière, et non son cerveau, qui marche ou qui mange, il apparaît que cette thèse conduit à une certaine forme de dualisme. Il y aurait en nous deux êtres : le sujet des opérations mentales et celui des actions physiques ; et entre l'un et l'autre, tout un jeu de relations causales.

Or, une fois dissocié le physique du mental, les cognitivistes vont bien sûr chercher à les réunifier. Expliquer les phénomènes mentaux, c'est pour eux expliquer comment, par exemple, un désir engendre un autre désir ou une croyance, mais c'est aussi expliquer comment un désir engendre une action. Quelle relation y a-t-il, en quelque sorte, entre mon désir de boire un verre et mon déplacement effectif vers la cuisine pour boire un verre ? Relation que le cognitiviste cherche à expliquer en termes matériels. Remarquons bien que ces problèmes de causalité ne se posent que parce que l'on a distingué la sphère des activités mentales de celle des actions physiques. Si l'on considère, au contraire, qu'une action, en tant que phénomène intentionnel, est aussi un phénomène mental, il n'est plus question de dire que le désir d'aller à la cuisine *cause* le déplacement en question, puisque le déplacement n'est qu'une expression de ce désir ; ou encore, il ne faut plus voir le déplacement comme un indice de ce désir préalable, mais comme une manifestation de ce désir. De même ce que l'on croit n'est pas un état de notre cerveau qui influencerait causalement nos

actions, mais c'est plutôt le style de ces dernières, leur orientation générale.

En revanche, si l'action n'est pas un phénomène mental, il faut bien expliquer comment elle est entraînée par ce dernier. Et pour le cognitiviste, comme on l'a dit, il faut trouver une explication qui fasse référence à des processus causaux (c'est-à-dire où une cause matérielle précède un effet lui aussi matériel). La méthode qu'il va alors suivre s'appuie sur un certain nombre de présupposés qu'il est utile d'explicitier. D'abord, le cognitiviste suppose que les phénomènes mentaux sont distincts les uns des autres. Ainsi, désirer boire un verre d'eau serait distinct de croire qu'il y a de l'eau dans la cuisine ; ce qui n'est bien sûr pas évident. Ensuite, il affirme que les phénomènes mentaux sont des états du cerveau. Vouloir aller à la banque, par exemple, doit donc se traduire par l'existence d'une configuration particulière du cerveau ; ce qui, comme on l'a vu, pose problème. Et enfin, il affirme que les entités mentales, dont la réalité vient d'être posée, ont une fonction causale. Ces présupposés acceptés, il semble toutefois difficile de constituer une théorie « mécanique » de l'esprit.

Pour le cognitiviste, si je vais dans la cuisine, c'est que j'avais l'intention, le désir, d'y aller. Une intention est par définition orientée vers un but : j'ai l'intention d'aller dans la cuisine parce que je pense que je vais y trouver de l'eau et que cela est censé me procurer du plaisir. Ce plaisir, toutefois, n'existe pas encore et ne peut donc pas influencer mon comportement. Il faut en conclure que c'est la représentation de ce plaisir qui cause ma conduite, ou encore, que c'est l'idée de plaisir qui agit sur mes organes moteurs. Mais ce mode d'explication, où une représentation agit sur des choses, n'est autre que celui qui est utilisé dans la magie où désirer qu'un événement se produise peut effectivement suffire pour qu'il ait lieu. S'il vous suffisait de penser à quelqu'un ou à quelque chose pour l'atteindre d'une façon ou d'une autre, vous vous seriez découvert là un pouvoir extraordinaire. Et c'est bien de ce pouvoir causal des représentations que prétendent bénéficier les sorciers et certains adeptes de la parapsychologie. Il est donc pour le moins remarquable de voir les cognitivistes recourir, sans s'en rendre compte, à un type d'explication qu'ils sont par ailleurs — au vu de leur conception de ce que doit être une explication scientifique — les premiers à rejeter. Certes, ici tout se passerait dans la tête. Mais en quoi un mode d'action magique serait-il moins magique parce qu'il se passe dans la tête.

Il est tout aussi remarquable que cette difficulté à réunifier la sphère des activités mentale et celle des actions soit semblable à celle à laquelle étaient confrontés les spiritualistes, que les cognitivistes ne

cessent pourtant de décrier. Ces penseurs soutenaient que les capacités mentales devaient être attribuées à une partie immatérielle de la personne et défendaient ainsi l'idée d'une distinction réelle, substantielle, de l'âme et du corps. Une fois cette distinction faite, il leur était effectivement difficile d'expliquer l'action de l'âme sur le corps, et réciproquement. De leur côté, les cognitivistes, qui sont matérialistes, refusent cette dualité. Mais c'est pour affirmer immédiatement une autre dualité entre le sujet physique et le sujet mental. La différence entre spiritualistes et cognitivistes ne concerne donc que la nature de l'esprit, pas sa localisation : les premiers affirment que c'est une partie immatérielle d'une personne (son âme) qui pense, alors que les seconds affirment que c'est une partie matérielle (son cerveau). Mais étant ainsi d'accord pour détacher le mental (immatériel pour les premiers, matériel pour les seconds) du monde extérieur (matériel), les uns comme les autres se retrouvent confrontés au difficile problème de leur connexion.

Pour montrer, toutefois, qu'ils sont sur la bonne voie, les cognitivistes exploitent l'analogie de l'ordinateur. Qu'une machine, construite avec des composants électroniques, arrive à manipuler des symboles et à effectuer des processus « intelligents » n'est-ce pas la preuve qu'il est possible de produire de la pensée dans un cerveau uniquement avec des éléments matériels ? Notre cerveau ne serait en quelque sorte qu'un ordinateur, et l'esprit serait au premier ce que le programme informatique est au second. Comme l'exécution d'un programme ne dépend pas fondamentalement des matériaux dont est constituée la machine, il semble permis de considérer comme sans importance, pour la pertinence de l'analogie, le fait que les ordinateurs ne soient pas des systèmes biologiques comme les cerveaux. Seule importe pour un cognitiviste la manière dont des symboles s'enchaînent suivant les instructions d'un programme. Ainsi, le cognitiviste pense avoir trouvé, dans cette invention moderne, le meilleur modèle pour comprendre par quel processus calculatoire interne au cerveau une croyance, un désir, une idée, ou tout autre phénomène mental, entraîne causalement un autre phénomène mental ou une action.

Cette analogie entre l'ordinateur et le cerveau n'est bien sûr pas complètement dénuée de sens, puisque l'un et l'autre manipulent des symboles. Mais est-elle suffisante pour affirmer que l'ordinateur pense ? Et puis, à partir de quel niveau de complexité pourrait-on envisager qu'une machine a effectivement une activité mentale ? Pour répondre à ces questions, les cognitivistes font souvent référence à un test que proposait Alan Turing en 1950 : mettez un ordinateur dans une pièce et une personne avec un terminal informatique dans une autre pièce ; demandez à une

deuxième personne, située dans une autre pièce et ignorant à qui ou à quoi elle a affaire, de poser des questions à partir d'un autre terminal informatique à l'ordinateur et à la première personne ; l'argument du test est que si l'ordinateur peut se faire passer, aux yeux de la deuxième personne, pour un être humain en donnant des réponses qui font l'effet d'être sensées, comme le fait la première personne si on lui a demandé de répondre normalement, c'est que l'ordinateur pense comme un être humain. Si ce test était pertinent, il ne resterait plus aux informaticiens qu'à construire un ordinateur qui le passerait avec succès pour que l'on puisse dire qu'on a compris le mécanisme de la pensée puisqu'on a été capable de construire une machine qui pense comme l'homme.

Cette interprétation est toutefois douteuse. Comprendons bien que le test est basé sur un principe de simulation : l'ordinateur doit faire *comme si* il pensait. Or, dans tout jeu de simulation, le fait qu'un simulateur ait réussi à se faire passer pour quelqu'un d'autre ne prouve pas qu'il est cette autre personne. Par exemple, imaginez que vous preniez la voix d'un de vos amis et que vous appeliez au téléphone un ami commun ; si ce dernier vous prend pour celui que vous imitez, ceci ne veut pas dire que vous l'êtes. Rien n'est donc prouvé quant à la « psychologie » de l'ordinateur ; cela prouverait uniquement que l'on peut se faire mystifier par une machine. Qui plus est, dans un tel test, l'ordinateur ne fait que répondre à des questions ; jamais l'ordinateur ne prend l'initiative, ce qui est pour le moins insuffisant pour prétendre simuler une conversation. Or, pour ceci, il faudrait que l'ordinateur soit capable d'avoir des intérêts, des goûts, des désirs et qu'il soit capable d'avoir de l'*à-propos* ; autant de caractéristiques qu'un ordinateur qui ne participe pas à quelque forme de vie que ce soit ne peut acquérir.

Oublions toutefois ces critiques et comparons le fonctionnement d'un ordinateur et d'un être humain lors d'un travail beaucoup plus spécialisé, comme un calcul. Peut-on, comme le cognitiviste, tirer argument de cette capacité des machines à effectuer des calculs, c'est-à-dire à manipuler des symboles, pour subodorer le caractère mécanique de la pensée, ne serait-ce que dans une opération aussi simple ? Ceci est encore douteux. Pour l'être humain il s'agit, par exemple, d'effectuer une multiplication en suivant les tables du même nom ; pour l'ordinateur il s'agit de suivre « mécaniquement » les instructions d'un programme établi à partir de ces mêmes tables. Or, si pour la machine le passage d'un état à un autre s'explique bien physiquement (une instruction déterminant le passage d'un état à un autre état), le calculateur humain, quant à lui, n'est pas mis en mouvement par les règles du calcul ; ces dernières n'agissent pas au sens mécanique du terme sur son esprit et il doit tou-

jours être vigilant à bien les appliquer. Si l'être humain se trompe, il faut le rappeler à l'ordre ; si c'est la machine, il faut la réparer. Ce qui montre que la présence d'une règle à l'esprit de l'être humain, n'est pas équivalente à la présence d'une instruction dans la mémoire de la machine. Malgré la similitude des deux opérations de calcul, il paraît donc difficile de conclure à leur identité.

Cette comparaison entre l'esprit et l'ordinateur n'a de sens, de toute façon, que si l'on considère que l'esprit manipule des représentations et non des choses physiques existant en dehors de notre tête. Mais cette thèse d'un esprit « représentateur », c'est-à-dire d'un esprit qui n'aurait affaire qu'à des images mentales du monde extérieur, ne va pas de soi. En effet, quand vous regardez une image représentant un objet, vous pouvez comparer l'image à cet objet. Si le monde extérieur ne nous apparaissait que sous forme d'images représentatives, il faudrait pouvoir comparer ces images à ce qu'elles représentent. Mais comment les comparer aux objets représentés, puisque ces derniers nous apparaîtraient justement sous la forme d'images mentales ? De plus, admettre que l'esprit n'ait jamais directement affaire à la réalité extérieure, mais uniquement à sa représentation mentale, conduirait à adopter une conception « solipsiste » de la pensée, au sens où les phénomènes mentaux d'un individu pourraient être les mêmes si cet individu était la seule créature existante. Par exemple, on a l'habitude de considérer que la jalousie amoureuse exige trois personnes (A est jaloux de B au sujet de C). Mais le cognitiviste affirme que la jalousie de A n'est autre qu'un état particulier de son cerveau. Sa jalousie ne dépendrait donc pas intrinsèquement de l'existence des autres protagonistes. Et effectivement, on peut très bien imaginer qu'une personne soit jalouse à l'égard d'un rival imaginaire et au sujet d'une personne qui, elle aussi, pourrait ne pas exister. Mais autant on peut comprendre que quelqu'un soit jaloux alors qu'il n'a pas de raison de l'être, autant il est difficile d'accepter l'idée qu'il n'y a pas de différence entre la lucidité et le délire.

En dehors de la conception représentative de l'esprit, c'est peut-être la notion de symbole qui est aussi source de confusion. Quand le cognitiviste dit que l'ordinateur manipule des symboles, il a raison. Quand il dit qu'ils ont un rôle causal uniquement en vertu de leurs propriétés physiques et non de leur signification, il a encore raison. Et la machine est bien construite pour que les transformations effectuées s'accordent avec ce que donnerait une transformation intelligente qui semblerait tenir compte de ces significations. C'est pourquoi il se croit autorisé à conclure que la machine pense ou, ce qui revient au même, que l'activité mentale n'est qu'un processus physique

interne au cerveau. Mais le cognitiviste oublie que, quand la machine manipule des symboles, c'est notre interprétation qui leur a donné ce statut de symboles. La machine n'a pas la capacité instituante par laquelle quelqu'un peut traiter différentes choses comme des symboles afin de communiquer une pensée à propos d'autre chose. Pour cela, il faudrait qu'elle participe de l'institution par laquelle ces symboles sont des symboles. Il faudrait qu'elle ait une forme de vie, qu'elle se réalise dans un monde qui lui est extérieur. Si le symbole n'est symbole que parce qu'il renvoie à autre chose que lui-même, s'il n'existe qu'en dehors de lui-même en vertu d'une institution, il faut en conclure qu'il n'y a de pensée qu'en dehors de ce qui la rend possible, que ce soit un cerveau ou éventuellement une machine. En fin de compte, l'erreur que commettraient les cognitivistes serait donc une erreur de catégorie. Croyant étudier la pensée, ils n'aborderaient que ses conditions de possibilité.

Après toutes ces critiques, faut-il conclure qu'il est impossible de rendre compte de l'activité mentale par des processus causaux ? Il semble, à lire ce livre, que ce soit le cas. Pour justifier mon déplacement en cuisine, je ne pourrais que mettre en avant des raisons comme mon désir de boire un verre d'eau et ma croyance en la présence d'un verre et d'une bouteille d'eau dans cette pièce, mais ceci sans que l'intention — ou son prétendu substrat matériel — de boire un verre d'eau ne s'inscrive dans une chaîne causale aboutissant au geste en question, ou encore, sans considérer l'intention comme un événement cérébral préalable à l'accomplissement de l'action. En conséquence de quoi, il faudrait distinguer les sciences de l'esprit (psychologie, histoire...) des sciences naturelles (chimie, physique, biologie...). Les premières chercheraient à savoir ce qui motive une action, ou quelles en sont les *raisons*, et les secondes ce qui détermine mécaniquement un processus physique, ou ce qui le *cause*. Cela ne voudrait pas dire qu'il faille donner à l'homme un statut à part dans l'univers au sein duquel il vit. Mais plus simplement que l'étude des phénomènes mentaux ne peut se faire par des méthodes qui n'ont de sens que pour des objets localisés. Cela ne veut pas dire non plus, comme on l'a souvent prétendu, que dans les sciences naturelles on chercherait uniquement à expliquer les processus physiques par des lois générales alors que dans les sciences de l'esprit on chercherait au contraire à comprendre les activités humaines par un phénomène de sympathie. Car la compréhension est bien ce qui est recherché par les unes comme par les autres ; seulement c'est en montrant comment des mécanismes peuvent être responsables des processus physiques que cela se fait dans un cas, alors que c'est en identi-

fiant ce qui motive les comportements que cela se fait dans l'autre.

Voici présenté rapidement un certain nombre de problèmes soulevés par Vincent Descombes dans ce livre exigeant qui trouve dans les réflexions de Ludwig Wittgenstein sa principale source d'inspiration. Faut-il en conclure que les tentatives de naturaliser l'esprit sont sans intérêt ? Malgré la pertinence de ses analyses, il serait hasardeux de l'affirmer. Non seulement parce que la question est difficile et qu'il existe des critiques sensées à cette thèse de l'« externalité » de l'esprit. Mais, peut-être plus fondamentalement, parce que, même si les cognitivistes se trompaient du tout au tout sur le sens de leur travail, celui-ci ne serait pas pour autant inutile pour la compréhension des phénomènes mentaux : tout simplement parce qu'il n'est pas impossible qu'une meilleure connaissance des conditions de possibilité de l'esprit puisse permettre de mieux en saisir, sur certains points, le fonctionnement.

Mais l'intérêt du livre de Vincent Descombes ne se limite pas uniquement à cette réflexion profonde sur la localisation des phénomènes mentaux à partir de laquelle il mène sa critique systématique des présupposés des sciences cognitives. Il réside aussi dans la mise en perspective historique des problèmes philosophiques qu'il soulève. On entraperçoit ainsi ce qui peut rapprocher Wittgenstein d'Aristote et, dans l'autre camp si l'on peut dire, on perçoit tout ce que les psychologues qui ont pris modèle sur les sciences naturelles, jusqu'aux cognitivistes actuels, doivent au grand partage opéré par Descartes entre la sphère des activités mentales et celle de l'action. Et le moins intéressant n'est certainement pas de s'apercevoir que des penseurs comme Lacan et Lévi-Strauss ont été eux aussi confrontés aux difficultés résultant de ce partage, puisqu'ils en acceptaient d'une certaine manière les termes. Le premier en assimilant la notion de signifiant à quelque chose de matériel et en cherchant à définir l'action causale d'une « matière psychique ». Le second en voulant expliquer les aspects sociaux par un inconscient structural et en attribuant une efficacité causale aux symboles.

Ajoutons, pour finir, que ce livre a une suite (*Les institutions du sens*, Éditions de Minit, 1996) qui se propose de combler une insuffisance dans cet exposé de la thèse de l'« externalité » de l'esprit. À considérer que les pensées, croyances et autres phénomènes mentaux ne sont pas dans la tête mais relèvent d'une extériorité, on perd une idée chère aux cognitivistes, à savoir qu'ils sont des entités *distinctes* et, en tant que tels, qu'ils peuvent être identifiés. Cette idée d'entités *distinctes* est certes douteuse puisqu'il paraît difficile de dénombrer les pensées comme on dénombre des

objets, par exemple. En revanche, l'identification est fondamentale puisque, selon une formule consacrée, il n'y a pas *d'entité sans identité*. Et effectivement, s'il n'était pas possible de distinguer ce que pense une personne de ce que pense une autre, il faudrait en conclure qu'il revient au même d'avoir telle ou telle idée. On en viendrait rapidement à considérer que toute pensée n'est qu'une parole creuse ou, si l'on refuse cette option, que la thèse développée dans ce livre est incohérente. C'est pourquoi Vincent Descombes cherche à montrer dans le second livre que, sans concevoir les pensées comme des entités indépendantes, il est toutefois possible de les discerner. Alors seulement, il sera envisageable d'adopter cette vision holiste du mental, c'est-à-dire de reconnaître qu'il n'est pas possible de faire abstraction du contexte pour identifier les pensées.

Thomas LEPELTIER,
le 24 février 2000.

[zIntérieur]