Zeitschrift: Studia philosophica : Schweizerische Zeitschrift für Philosophie = Revue suisse de philosophie = Rivista svizzera della filosofia = Swiss journal of philosophy

Herausgeber: Schweizerische Philosophische Gesellschaft

Band: 55 (1996)

Artikel: Un aspect du cartésianisme en philosophie de l'esprit

Autor: Corazza, Eros / Dokic, Jérôme

DOI: https://doi.org/10.5169/seals-882984

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 09.12.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Studia Philosophica 55/96

EROS CORAZZA ET JÉRÔME DOKIC

Un aspect du cartésianisme en philosophie de l'esprit*

0. Introduction

Nous discutons dans cet article d'une forme de « cartésianisme » dont a hérité la philosophie de l'esprit contemporaine. Nous allons nous concentrer en particulier sur la thèse cartésienne de l'indépendance des pensées par rapport au monde « extérieur » – à ce que Descartes nomme la *res extensa* – en montrant, d'une part que cette thèse n'est pas logiquement liée au dualisme cartésien – c'est-à-dire à la distinction ontologique entre la *res extensa* et la *res cogitans* – et d'autre part que la plupart des formes du cartésianisme contemporain acceptent la thèse de l'indépendance sans toutefois souscrire au dualisme cartésien. Au cartésianisme contemporain qui est engagé à l'égard d'une forme d'internalisme, nous chercherons à opposer une version de l'externalisme selon laquelle les pensées ne sont pas indépendantes de l'environnement au sein duquel elles sont formées. Nous établirons un contraste entre deux formes d'externalisme : un externalisme fort selon lequel les pensées dépendent de l'objet (ou de son identité numérique) et un externalisme modéré selon lequel les pensées, tout en dépendant de facteurs « externes », ne dépendent pas de l'objet sur lequel elles portent. Contre certains développements anti-cartésiens récents, nous prétendrons que l'externalisme modéré est suffisant pour rejeter le cartésianisme[1].

* Nous tenons à remercier les participants au colloque *Descartes. Et après?* pour leurs remarques, dont nous avons essayé de tenir compte en rédigeant cette version.

1 En philosophie de l'esprit contemporaine, les adjectifs « cartésiens » et « anti-cartésiens » sont très largement utilisés, et nous ne saurions présenter ici, faute d'espace, toutes les nuances que ces adjectifs revêtent aujourd'hui. Notre cible, ici, sera en fait deux études post-ryliennes récentes : MCCULLOCH (1995) et RUDDER BAKER (1995).

1. Les principes de l'internalisme cartésien

Dans cette première section, nous allons proposer une reconstruction de l'internalisme cartésien sous la forme de quatre principes qui nous semblent caractéristiques de cette position. Ce n'est pas l'exégèse de Descartes en tant que telle qui nous intéresse surtout ici, mais plutôt une certaine conception des attitudes psychologiques (expériences, pensées, intentions, etc.) que beaucoup de philosophes contemporains attribuent directement à Descartes ou, plus prudemment, qualifient de « cartésienne ». En formulant les principes en question, nous portons notre attention plus particulièrement sur la perception, mais nous envisagerons à la fin de cette section une façon de les étendre également à la pensée conceptuelle, et notamment à la pensée singulière, c'est-à-dire celle qui porte sur des objets numériquement déterminés et dont l'expression comporte typiquement des termes singuliers (noms propres, termes indexicaux, etc.). Rappelons que chez Descartes, le terme de « pensée » englobe à la fois les expériences sensorielles et les pensées conceptuelles[2].

Considérons le passage célèbre suivant, tiré des *Méditations Métaphysiques* :

> « Enfin je suis le même qui sens, c'est-à-dire qui reçois et connais les choses comme par les organes des sens, puisqu'en effet je vois la lumière, j'ouïs le bruit, je ressens la chaleur. Mais l'on me dira que ces apparences sont fausses et que je dors. Qu'il en soit ainsi ; toutefois, à tout le moins, il est très certain qu'il me semble que je vois, que j'ouïs et que je m'échauffe ; et c'est proprement ce qui en moi s'appelle sentir, et cela, pris ainsi précisément, n'est rien autre chose que penser » (1989 vol. II ; *Méditation Seconde* : 421-2).

Si je vois une tasse de café devant moi, je puis être certain qu'il me *semble* voir une tasse de café devant moi, même si je ne suis pas sûr que mon expérience visuelle soit véridique, c'est-à-dire renvoie à un état de choses réel – un état de choses distinct de moi et localisé dans l'espace qui m'environne. Il pourrait me sembler voir une tasse de café devant moi, dans le même sens, si j'étais victime d'une illusion, à savoir si mon expérience ne renvoyait à aucune tasse de café

2 « Par le nom de 'pensée', je comprends tout ce qui est tellement en nous, que nous en sommes immédiatement connaissants. Ainsi toutes les opérations de la volonté, de l'entendement, de l'imagination et des sens, sont des pensées » (1989 vol. II : *Secondes Réponses*, §1).

réelle[3]. En d'autres termes, si je fais l'expérience visuelle que *p*, qu'elle soit véridique ou non, la proposition « Il me semble voir que *p* » a un statut sémantique et épistémique particulier : elle est vraie, et sa vérité m'est connue avec certitude. Voici donc un premier principe internaliste (des remarques similaires valant pour d'autres modalités sensorielles) :

P1. La vérité de la proposition « Il me semble voir une tasse de café » (telle qu'elle est utilisée par moi) m'est connue avec certitude (au moment où je l'utilise), et ne dépend ni du fait que mon expérience visuelle est véridique (si elle est véridique), ni du fait qu'elle est illusoire (si elle est illusoire).

Le principe P1 est d'ordre *sémantique* : il concerne les conditions de vérité d'une (auto-)attribution du type « Il me semble voir que *p* ». P1 est relativement neutre quant à la nature *ontologique* des états mentaux qui vérifient l'attribution. On sait qu'une telle attribution décrit une expérience spécifique (visuelle, douée d'un certain contenu propositionnel), mais elle ne nous livre pas toutes les conditions d'identité de celle-ci.

Selon un deuxième principe typiquement internaliste, et cette fois d'ordre ontologique, les expériences décrites par des propositions du type « Il me semble voir que *p* » décrivent des *apparences* qui, comme la vérité des propositions en question, ne sont pas concernées par la différence entre une expérience véridique et une illusion. La *même* apparence, décrite par une proposition du type « Il me semble voir que *p* » peut correspondre tantôt à une expérience véridique, tantôt à une illusion.

Surtout, de telles apparences sont *purement mentales* au sens où elle ne dépendent pas essentiellement de la réalité extérieure. Elles peuvent sans doute dépendre *causalement* de la réalité extérieure, si on admet qu'elles peuvent être des effets mentaux de causes matérielles. Mais la cause n'est pas nécessairement matérielle : Descartes semble admettre que ces apparences, normalement causées par la réalité extérieure, auraient pu avoir des causes mentales – par exemple,

3 Pour simplifier la discussion, nous ne faisons pas ici la distinction bien connue entre une hallucination, où rien n'est à proprement parler perçu, et une illusion, ou quelque chose est perçu mais avec des propriétés qu'il n'a pas réellement (par exemple, un bâton droit plongé dans l'eau mais perçu comme courbé). Par « illusion » nous entendons ici le cas radical, où rien n'est à proprement parler perçu.

elles auraient pu être directement causées par un malin génie. Elles sont donc indépendantes, au sens ontologique, du monde extérieur :

P2. Les propositions du type « Il me semble voir que *p* » (telles qu'elles sont utilisées par moi) décrivent un monde purement mental d'*apparences*, indépendantes de la réalité extérieure.

A proprement parler, P1 n'implique pas logiquement P2. Il est théoriquement possible d'accepter P1 mais de rejeter le principe plus controversé P2. Ainsi, on pourrait affirmer que la vérité d'une proposition du type « Il me semble voir que *p* » ne dépend pas du statut illusoire ou véridique de l'expérience décrite (P1), et nier que la proposition décrive des expériences qui soient *en elles-mêmes* indifférentes à la réalité extérieure (¬P2). De telles expériences pourraient dépendre de diverses façons essentielles de la réalité extérieure, même si elles étaient toutes correctement décrites de la même manière, c'est-à-dire au moyen de la proposition « Il me semble voir que *p* ». Simplement, il s'agirait d'une description *incomplète*, parce que générique, de leur nature spécifique.

Du point de vue cartésien, le principe P2 a l'avantage de fournir une réponse au moins provisoire aux questions posées par le scepticisme radical. Selon ce principe, en effet, si nous doutons de l'existence du monde extérieur, nous pouvons pour ainsi dire battre en retraite, et nous réfugier dans un monde intérieur d'apparences parfaitement définies, décrit par des propositions du type « Il me semble voir/penser/etc. que *p* ». Si quelqu'un remet en question notre capacité générale à connaître, et en particulier à connaître le monde extérieur, nous pouvons au moins rétorquer que nous connaissons des faits qui relèvent du monde intérieur des apparences, quelle que soit la relation exacte entre ces apparences et le monde extérieur s'il existe[4].

Toutefois, les deux principes P1 et P2 ne suffisent pas à définir entièrement l'internalisme cartésien. Le fait que je peux décrire des apparences indépendantes du monde extérieur en utilisant une proposition du type « Il me semble voir que *p* » n'implique pas que mes *autres* expériences, celles que je ne décris pas explicitement en utilisant une proposition de ce type, aient le même type d'indépendance par rapport au monde extérieur. En effet, il est assez naturel de

4 McDowell (1986 : 148), qui s'appuie sur Burnyeat (1982), voit là l'une des innovations cartésiennes par rapport au scepticisme antique : « [L]a concession, de la part des sceptiques antiques, selon laquelle les apparences ne peuvent pas être remises en question devient l'idée d'une palette de faits connaissables de manière infaillible par le sujet impliqué dans ceux-ci ».

supposer que lorsque j'utilise une proposition du type « Il me semble voir que *p* », je forme une pensée *de deuxième ordre*, c'est-à-dire une pensée qui porte par introspection sur mon expérience visuelle. Mais alors, comment exclure que lorsque je tourne ma pensée vers mon expérience visuelle, celle-ci m'apparaisse modifiée de façon plus ou moins importante, au lieu de se présenter telle qu'elle est en elle-même ? Car si l'introspection modifie radicalement mon expérience visuelle, il est possible que le monde d'apparences que me révèle le doute cartésien ne soit accessible *que* si j'adopte une attitude introspective, et s'évanouisse dans le cas contraire. En revanche, si la pensée introspective ne modifie pas l'expérience, elle permet au sujet de prendre conscience d'apparences qui ne dépendent pas de l'attitude introspective elle-même, et qui subsistent même lorsque je ne suis pas d'humeur à introspecter[5].

Bernard Williams critique Descartes pour avoir confondu la conscience immédiate, par exemple la conscience d'un objet externe de perception, et la conscience réflexive, par exemple la conscience de la perception elle-même[6]. Ce que Williams appelle « conscience réflexive » est précisément ce que nous avons appelé « conscience introspective », c'est-à-dire une conscience de deuxième ordre qui porte sur une conscience de premier ordre. Mais en réalité, il est plus probable qu'il ne s'agisse pas d'une confusion de la part de Descartes, mais d'une thèse substantielle qui met en rapport interne perception attentive et pensée réflexive. Invité à clarifier ses vues sur le rapport entre la vision des animaux et la nôtre, Descartes précise en effet, dans une lettre du 3 octobre 1637 :

> « [M]on opinion n'est pas que les bêtes voient comme nous lorsque nous sentons que nous voyons, mais seulement qu'elles voient comme nous lorsque notre esprit est appliqué ailleurs » (1989 vol. I : 786).

Parfois, je vois alors que « mon esprit est appliqué ailleurs », c'est-à-dire quand je ne suis pas attentif aux éléments de mon champ perceptif. Par exemple, je suis au volant de ma voiture mais je me concentre sur ce que me dit mon passager. Je *vois* la voiture qui me précède tourner à gauche, puisque je ralentis pour ne pas l'emboutir, mais je ne forme pas la pensée que je vois la voiture (ni celle qu'il me

5 Il nous semble que MCDOWELL (1986) néglige les problèmes liés à la réflexivité dans sa reconstruction – par ailleurs remarquable – de l'internalisme cartésien.

6 Cf. WILLIAMS (1978 : 226 et 286).

semble la voir). Dans une situation de ce genre, on peut rapprocher, selon Descartes, mon expérience perceptive de celle des animaux.

La thèse de Descartes s'accorde bien avec la perspective internaliste. La vraie vision est la vision attentive et réflexive, la seule qui fasse intervenir l'esprit, alors que l'expérience non attentive, celle qui n'engage pas la pensée réflexive et que nous partageons avec les animaux, est en réalité purement « mécanique ». (Ici, Descartes est bien entendu fidèle à sa conception des animaux comme des automates naturels.) La réflexivité ne serait donc pas une propriété accidentelle de l'expérience perceptive, mais un élément constitutif de celle-ci, responsable du fait qu'elle peut être considérée comme un état mental conscient, et non un simple état corporel[7].

Mais si la réflexivité est déjà constitutive de toute perception immédiate, il est beaucoup moins pertinent d'affirmer que lorsque je forme expressément une pensée réflexive sur mon expérience, je modifie radicalement celle-ci. En fait, selon Descartes, mon expérience reste la *même* (immédiate et objet de réflexion à la fois, pour ainsi dire) que je la décrive ou non en utilisant une proposition du type « Il me semble voir que *p* ». Par suite, nous pouvons parler d'un monde d'expériences et d'idées *transparent à la conscience*. Si nous combinons ce dernier point aux deux principes internalistes déjà formulés, nous obtenons le principe suivant :

> P3. Les apparences accessibles au doute cartésien correspondent à l'aspect purement mental de nos expériences perceptives ordinaires, de premier ordre[8].

Ce que la conscience révèle, c'est le côté purement mental de toute expérience de premier ordre, qu'elle soit illusoire ou véridique. Je perçois une tasse de café. Mon expérience pourrait être véridique, mais elle pourrait aussi être une illusion. En fait, quel que soit l'état du monde extérieur, elle aurait le *même* aspect mental accessible à la conscience.

7 On peut laisser ouverte ici la question difficile de la structure de la réflexivité chez Descartes. Nous ne voulons certainement pas attribuer à Descartes la thèse selon laquelle toute pensée est réflexive au sens où elle est l'objet d'une *autre* pensée d'ordre supérieur. On pourrait suivre ici Arnauld qui établissait une distinction entre la réflexion virtuelle, une propriété de toute pensée qui la rend consciente au sujet, et la réflexion expresse, une opération de deuxième ordre sur une pensée. Sur cette distinction, cf. BAERTSCHI (1992 : 380sq).

8 Cf. WILLIAMS, 1978 : 284-5.

Le principe P3 est théoriquement compatible avec l'idée que les expériences, en plus d'avoir un aspect purement mental, ont aussi un aspect « externe », inaccessible à la conscience. Si cet aspect externe participe de la nature intrinsèque ou essentielle de l'expérience, cette idée serait bien entendu contraire à la philosophie cartésienne. Pour compléter notre reconstruction de l'internalisme cartésien, il faut ajouter que l'aspect purement mental de nos expériences perceptives *constitue* leur nature essentielle en tant qu'expériences. Si les expériences ont un aspect externe, celui-ci ne peut être qu'accidentel. Nous obtenons enfin la thèse de l'autonomie des états mentaux par rapport au monde extérieur :

> P4. Nos expériences ordinaires, de premier ordre, sont autonomes par rapport au monde extérieur au sens où elles pourraient rester les mêmes, du point de vue de leur nature intrinsèque, quel que soit l'état du monde extérieur[9].

Le principe P4 est d'ordre modal : l'aspect purement mental de nos expériences définit leur nature essentielle. A la question de savoir quel dénominateur commun il y a entre ma perception véridique d'une tasse de café et l'illusion possible de percevoir une tasse du même type, la réponse correcte est qu'il s'agit d'un état mental à part entière, indépendant du monde extérieur.

La plupart des commentateurs analytiques de Descartes lui attribuent – implicitement ou explicitement – les quatre principes internalistes qui viennent d'être formulés. Ceux-ci concernent spécifiquement la perception, mais des principes similaires peuvent être envisagés pour la pensée singulière. Considérons l'exemple radical suivant (sur lequel on reviendra dans la section 3). Dans le cas de la perception, l'internalisme cartésien implique que la différence entre une expérience illusoire et une expérience véridique est non pertinente du point de vue mental. Selon une thèse analogue appliquée à la pensée singulière, ce qui est non pertinent du point de vue mental, c'est la différence entre une pensée qui porte sur un objet réel, numériquement déterminé et une pensée qui porte en apparence sur un tel objet mais qui en réalité est *vide*[10]. Or certains philosophes contem-

9 Cf. McDowell (1986 : 151). Le principe P4 est aussi ce que McCulloch (1995) appelle la doctrine de la « self-containedness ».

10 Notons que tout le monde s'accorde à dire que la différence entre une pensée vraie et une pensée fausse est non pertinente du point de vue mental, au sens où elle n'affecte pas la nature essentielle de la pensée en tant que structure

porains prétendent qu'il est possible d'essayer de former une pensée singulière à un moment donné sans que l'on puisse s'assurer en examinant sa conscience à ce moment si la tentative est une réussite ou un échec, c'est-à-dire si la pensée est véritablement formée ou non. Par exemple, Gareth Evans rejette la thèse, qu'il attribue à Russell, selon laquelle les objets de nos pensées auraient ce qu'il appelle précisément la « propriété cartésienne » :

> « Tous les objets avec lesquels un sujet peut être en relation directe [*acquainted*] ont ce qui peut être appelé la propriété cartésienne : il n'est pas possible qu'un sujet pense qu'il y a un objet d'un genre approprié avec lequel il est en relation directe [...] sans qu'il y ait un objet de ce genre » (1982 : 44).

Dans le cas précis, la remarque est dirigée contre la notion de relation directe telle qu'elle est interprétée par Russell (1956) (comme mettant en relation le sujet avec des données sensorielles privées), mais elle concerne en principe toute théorie qui admet comme Descartes la transparence absolue de la pensée à la conscience. Selon Evans, il est possible de penser *à tort* avoir réussi à former une pensée d'un certain genre. Par exemple, si je suis victime d'une illusion, je peux essayer de former la pensée « Voici un petit homme vert », mais ma tentative échoue fatalement : je ne suis pas en relation mentale directe et appropriée avec un petit homme vert. La conscience ne nous révèle pas la nature mentale ultime de nos propres expériences : selon la théorie de la pensée défendue par Evans, la nature d'une pensée véritable, qui porte sur un état de choses réel, diffère de celle d'une pensée vide, mais la différence ne se manifeste pas forcément à la conscience.

2. Le monisme cartésien

Peu de gens acceptent aujourd'hui le dualisme du mental et du physique tel qu'il est interprété par Descartes, c'est-à-dire l'idée selon laquelle les états mentaux sont ceux d'une substance immatérielle, distincte des substances physiques sur lesquelles ils portent éventuellement. Une difficulté bien connue du dualisme cartésien concerne le mystère qui entoure l'interaction entre le mental et le physique : manifestement, l'esprit interagit causalement avec le monde extérieur,

propositionnelle (ce qui n'implique pas, bien sûr, que la vérité d'une pensée soit inconnaissable).

et vice-versa ; or il faut expliquer comment deux substances ayant des natures totalement différentes peuvent entretenir des relations causales.

Mais le point crucial, pour nous, est que l'internalisme cartésien, défini par les quatre principes formulés dans la section précédente, est indépendant de la thèse immatérialiste de Descartes. Certes, on préfère de nos jours supposer que toutes les entités réelles – les chaises, les explosions, les sentiments, les attitudes propositionnelles – sont localisées dans le même univers spatio-temporel, et donc embrasser contre Descartes une forme de *monisme* ontologique. La conception du mental représentée par les principes P1-P4 (étendus à l'ensemble des états mentaux) survit toutefois à cette critique du dualisme cartésien. L'idée d'un « monisme cartésien » – c'est-à-dire d'un internalisme cartésien combiné à un monisme ontologique – n'est pas à première vue incohérente. Dans cette section, nous allons rappeler brièvement quelles sont aujourd'hui les versions les plus répandues de ce que nous appelons le monisme cartésien. Toutes ont ceci de particulier que la ligne de partage entre le mental et le non-mental n'est plus métaphysique, comme chez Descartes, mais littéralement *spatial* : si tout est physique, les états physiques que l'on peut également qualifier de « mentaux » – les pensées, les perceptions, etc. – sont circonscrits dans une région particulière du monde, c'est-à-dire à l'intérieur de la tête, plus précisément dans le cerveau. Les adjectifs « intérieur » et « extérieur » doivent être entendus en un sens quasi littéral. Mais les états mentaux « intérieurs » en ce sens spatial sont toujours autonomes par rapport au monde « extérieur », c'est-à-dire au reste du monde physique : la notion modale d'indépendance s'applique ici (presque) aussi bien qu'à la position de Descartes[11].

Bien entendu, le moniste cartésien ne considère pas que la conscience ait toujours le dernier mot sur l'identité des états mentaux – une

11 Cf. MCDOWELL (1986 : 153), MCCULLOCH (1995), RUDDER BAKER, (1995). Elle s'applique *presque* aussi bien car la physique contemporaine (et en particulier la physique quantique) nous interdit peut-être d'affirmer que les états cérébraux sont *absolument* indépendants du reste du monde physique. Mais cette précision, bien qu'importante, ne change rien aux grandes lignes de la thèse du monisme cartésien qui nous intéresse ici, selon laquelle les états cérébraux sont, *en tant qu'états mentaux*, relativement autonomes par rapport au reste de la réalité extérieure. Le monisme cartésien reconnaît par conséquent l'existence d'une ligne de partage tranchée entre les états mentaux « intérieurs » et la réalité « extérieure » sur laquelle ils portent éventuellement.

inspection neurophysiologique éclairée du cerveau serait sans doute plus appropriée. Mais les résultats d'un « examen de conscience » restent, sinon un critère, du moins un symptôme fiable de l'identité et de la différence des états mentaux. Si deux états mentaux paraissent exactement similaires du point de vue de la conscience, le moniste cartésien doit expliquer cette similarité, et il le fera presque certainement en postulant une structure neurophysiologique commune (peut-être d'ordre supérieur).

On peut dire que dans la philosophie contemporaine, le monisme cartésien se manifeste sous au moins cinq formes qui correspondent à autant de doctrines philosophiques[12] :

1. *Les théories de l'identité-type.* Selon ces théories, les types des états mentaux sont identiques à des types d'états cérébraux. Ainsi, pour chaque type A d'attitude psychologique (croire/désirer/souhaiter que *p*) et pour chaque sujet S, il y a un type d'état cérébral E tel que S est dans E si et seulement s'il a l'attitude A[13].

2. *Les théories de l'identité-occurrence.* Selon ces théories, les occurrences des états mentaux sont identiques à des occurrences d'états cérébraux. Ainsi, pour chaque occurrence *a* d'une attitude psychologique, et pour chaque sujet S, *a* est l'occurrence d'une attitude de S si et seulement s'il y a une occurrence *e* d'état cérébral telle que $a = e$[14].

3. *Les théories constitutives.* Selon ces théories, les attitudes psychologiques sont constituées par des états cérébraux. Toutefois, les occurrences des états cérébraux ne sont pas nécessairement identiques aux occurrences des attitudes, dans la mesure où un autre état cérébral aurait pu instancier la même occurrence d'attitude psychologique. Un état cérébral constitué d'autres molécules aurait pu instancier la même

12 Cf. RUDDER BAKER (1995 : ch 1).

13 Cf. ARMSTRONG (1980).

14 Cf. PUTNAM (1975). Davidson (1993) est un cas à part, dans la mesure où il défend une théorie de l'identité-occurrence (toute occurrence d'état mental est identique à une occurrence d'état cérébral), mais ne souscrit pas à l'internalisme cartésien. Il défend en fait au moins une forme d'externalisme modéré, au sens de la section suivante. Toutefois, comme Davidson semble refuser systématiquement d'introduire comme nous les notions modales de dépendance et d'indépendance pour caractériser le débat entre l'internalisme et l'externalisme, sa position apparaît comme un compromis peu stable – elle ne permet pas en particulier de poser la question des propriétés essentielles ou constitutives de la pensée.

occurrence d'attitude psychologique. Il s'agit d'une théorie matérialiste non réductionniste[15].

4. *Les théories fonctionnalistes.* Selon ces théories, les attitudes psychologiques sont définies par des rôles fonctionnels occupés par des états cérébraux. Les états mentaux sont expliqués en fonction de leurs causes et de leurs effets. Un état mental constitue une attitude de type A en vertu de ses relations aux stimulations sensorielles (entrées), de ses relations avec les autres états internes et de ses relations avec le comportement (sorties)[16].

5. *Le matérialisme éliminativiste.* Cette théorie met en doute l'existence des attitudes psychologiques mais admet que si elles existaient, elles seraient identiques à des états cérébraux. Seulement, les états cérébraux ne peuvent pas être identifiés au moyen d'attributions d'attitudes psychologiques de la forme « S croit que *p* » (où S est un sujet et « *p* » une proposition)[17].

Rudder Baker (1995 : 12-18) reconnaît trois motivations ou arguments très simples en faveur de la conception reçue selon laquelle les attitudes psychologiques sont, d'une façon ou d'une autre, des états cérébraux. Pour simplifier la discussion, nous pouvons nous concentrer sur le cas de la croyance, mais des remarques similaires sont valables pour d'autres attitudes psychologiques, dont les expériences.

15 Les théories constitutives sont moins répandues ; elles reposent sur une définition appropriée de la notion de constitution, qui n'implique pas l'identité (*x* peut constituer *y* sans être à proprement parler identique à lui). Sur ces théories, cf. BAKER (1995 : 8-10).

16 Cf. LOAR (1981) et LEWIS (1983). Deux remarques sur le fonctionnalisme sont peut-être utiles. Premièrement, le fonctionnalisme est une espèce d'internalisme cartésien seulement si les « entrées » et les « sorties » qui définissent les états mentaux sont elles-mêmes internes : ce sont par exemple des stimulations sensorielles et des réponses motrices comprises comme des modifications localisées à la surface du cerveau (cf. HORNSBY, 1986). Toutefois, certaines versions du fonctionnalisme sont dites « larges » car les entrées et les sorties pertinentes ne sont pas internes : ce sont par exemple des objets perçus et des comportements compris comme des éléments du monde extérieur. (HARMAN, 1982 défend une telle version du fonctionnalisme). Deuxièmement, dans la mesure où un fonctionnaliste défend la doctrine selon laquelle avoir un esprit signifie avoir une constitution interne d'une sorte ou d'une autre, il va sans dire qu'il n'est pas contraint d'affirmer que les attitudes psychologiques sont des états cérébraux. Il se peut qu'une machine, un ordinateur par exemple, ait des rôles fonctionnels suffisamment complexes pour avoir des attitudes mentales. De même, le fonctionnalisme est en principe compatible avec le dualisme cartésien.

17 Cf. CHURCHLAND (1989).

Un argument tiré de la métaphysique :

(1) Les croyances (si elles existent) sont soit des états d'un esprit immatériel, soit des états cérébraux.
(2) Les croyances ne sont pas des états d'un esprit immatériel.
(3) Donc : les croyances sont des états cérébraux.

Le présupposé implicite de cet argument est la thèse de l'indépendance, inhérente au monisme cartésien. L'alternative présentée dans la première prémisse suppose en effet que les attitudes psychologiques sont des états internes indépendants du monde extérieur au sens du principe P4 (étendu à la croyance).

Un argument tiré de la science :

(1) La science a confirmé la conception des croyances comme des états cérébraux.
(2) Toute conception confirmée par la science est correcte.
(3) Donc : la conception des croyances comme des états cérébraux est correcte.

L'idée sous-jacente est que les sciences de base, comme la physique ou la neurophysiologie, démontrent empiriquement que les attitudes psychologiques sont des états cérébraux ou du moins reposent exclusivement sur ceux-ci.

Un argument tiré du pouvoir causal :

(1) Si les croyances n'étaient pas des états cérébraux, elles ne pourraient pas expliquer causalement le comportement.
(2) Les croyances peuvent expliquer causalement le comportement.
(3) Donc : les croyances sont des états cérébraux.

L'enjeu est ici le suivant : si Sophie saisit la canette de bière qui se trouve dans le frigo, c'est (typiquement) qu'elle désire étancher sa soif et croit qu'il y a une canette dans le frigo. Il semble qu'une situation de ce genre implique des causes mentales et des mouvements corporels. Toutefois, pour que ces causes mentales puissent déterminer le mouvement corporel, il faut qu'elles soient, d'une façon ou d'une autre, de nature corporelle. Or les meilleurs candidats pour jouer un tel rôle causal semblent être les états cérébraux[18].

18 Fodor (1988 : 30) écrit à ce sujet : « Mais on peut arguer que toute notion d'état psychologique utile d'un point de vue scientifique devrait respecter la surimposition. Après tout, la surimposition esprit/cerveau (et/ou l'identité esprit/cerveau) est la meilleure idée que quiconque ait eue jusqu'à présent sur la question de savoir comment la causalité mentale est possible. » Rappelons

Ces trois arguments sont cartésiens dans la mesure où ils souscrivent, au moins implicitement, au principe P4, c'est-à-dire à la thèse selon laquelle nos expériences ordinaires sont autonomes par rapport au monde «extérieur». C'est par référence à un principe de ce genre que le matérialiste peut finalement assimiler nos attitudes psychologiques à des états cérébraux. Si, en revanche, nous rejetons le principe P4 de l'autonomie, nous ne pouvons pas réduire les attitudes à des états cérébraux, dans la mesure où celles-ci ne seraient plus autonomes par rapport au monde «extérieur». Autrement dit, si les attitudes ne sont pas autonomes par rapport au monde «extérieur», elles ne peuvent pas être assimilées aux états cérébraux qui, par définition, sont indépendants de la réalité «extérieure» – c'est-à-dire le monde hors du crâne.

La stratégie anti-cartésienne, ou externaliste, la plus prometteuse consiste sans doute à rejeter la première prémisse des trois arguments. Elle consiste en somme à rejeter l'idée selon laquelle les croyances et les autres attitudes sont assimilables d'une façon ou d'une autre à des états cérébraux. Nous pouvons considérer la stratégie anti-cartésienne (ou externaliste) comme étant non réductionniste au sens où elle ne réduit pas ontologiquement les attitudes aux états cérébraux des agents. Par ailleurs, lorsqu'il s'agit d'expliquer le comportement en termes d'attitudes, la conception externaliste est aussi non réductionniste: les attitudes qui guident causalement le comportement du sujet ne sont pas réductibles aux états cérébraux de ce dernier. Une explication causale du comportement ne doit pas, selon la position anti-cartésienne, postuler que l'un des termes de la relation causale soit un état cérébral.

Jusqu'à présent, nous avons seulement montré, premièrement, que l'internalisme cartésien n'implique pas logiquement le dualisme ontologique et, deuxièmement, que le monisme ontologique n'implique pas logiquement l'internalisme cartésien. Ce dernier point signale l'existence possible d'une position intermédiaire qui rejette à la fois le dualisme ontologique et l'internalisme cartésien. Mais quels arguments peut-on invoquer contre l'internalisme cartésien?

qu'il y a surimposition de A sur B lorsque toute différence dans A entraîne nécessairement une différence dans B. Par exemple, s'il y a surimposition de l'esprit sur le cerveau, toute différence mentale ou psychologique entraîne quelque différence dans les états cérébraux.

3. Les formes de l'externalisme

Pour répondre de manière précise à cette question, nous devons distinguer trois formes d'engagement contre l'internalisme cartésien:
(1) La nature intrinsèque des états mentaux n'est pas entièrement accessible à la conscience.
(2) La nature intrinsèque des états mentaux dépend de facteurs «externes» (c'est-à-dire non mentaux selon les critères internalistes).
(3) La nature intrinsèque des états mentaux dépend de l'existence ou de l'identité numérique d'un objet «externe», qui devient alors l'objet de ces états.

Les relations logiques entre ces trois thèses sont les suivantes. La thèse 3 implique la thèse 2 (si A dépend de l'existence ou de l'identité numérique de l'objet «externe» B, il dépend plus généralement de facteurs «externes»). Par contre, la thèse 2 n'implique pas la thèse 3 (A peut dépendre de facteurs «externes» sans dépendre de l'existence de son objet «externe»). Comme on le verra, la thèse 1 est logiquement indépendante des autres thèses, bien que l'adoption de n'importe quelle thèse 2 ou 3 puisse fournir une raison d'adopter la première.

1. La première thèse n'est pas forcément externaliste, puisqu'un internaliste peut rejeter la conscience comme critère du mental. Par exemple, il peut soutenir que certains états mentaux sont «internes» mais inaccessibles à la conscience, comme le sont les états de l'inconscient chez Freud et ceux qui constituent la compétence linguistique chez Chomsky. (Un tel internaliste est toutefois tenu de nous livrer un critère substantiel du mental). De façon plus générale, il peut reconnaître entre des expériences des différences mentales qui ne se retrouvent pas au niveau de la conscience (du moins au moment où le sujet a les expériences en question). Ce n'est pas parce que deux expériences ont les mêmes propriétés conscientes qu'elles ont la même identité mentale[19].

Une autre façon de rejeter la première thèse consiste à renoncer à la thèse cartésienne selon laquelle la réflexivité est constitutive de la

19 Par exemple, McDowell (1986) présente une position possible, qu'il qualifie d'«innocente» par rapport à celle de Descartes, selon laquelle les propositions du type «Il me semble voir que *p*» décrivent parfois des expériences véridiques, et d'autres fois de simples apparences illusoires. Comme la différence entre une expérience illusoire et une expérience véridique n'est pas purement interne, la position de McDowell est bien sûr externaliste.

pensée en général. Lorsqu'on parle d'états mentaux « accessibles à la conscience », on fait référence à une conscience réflexive qui, selon ce point de vue, est toujours introspective – elle fait intervenir une pensée d'ordre supérieur. De ce fait, elle est susceptible de modifier plus ou moins radicalement la nature des états mentaux pris pour objets. Ainsi, même si la réflexion révèle un aspect essentiel commun entre deux expériences apparaissant comme exactement similaires, il ne s'ensuit pas que cet aspect soit également commun et essentiel lorsque les expériences en question *ne* sont *pas* des objets de la conscience réflexive. Il serait donc théoriquement possible que la conscience réflexive, en tant qu'elle implique de la part du sujet un changement d'attention radical du monde extérieur à ses propres états mentaux, le trompe sur la nature profonde de ces derniers. Cette position est, du moins en principe, compatible avec une attitude internaliste.

En rejetant la transparence de la pensée à la conscience, on fait un premier pas en direction de l'externalisme, mais ce pas n'implique pas encore que les états mentaux soient dépendants, relativement à leur existence et/ou à leur essence, de facteurs « externes » au sens cartésien. Il est en effet compatible avec l'idée que les aspects qui résistent à la conscience (réflexive) soient purement mentaux, au sens où ils seraient indépendants de l'état du monde extérieur. Le mental serait autonome par rapport au monde extérieur, spatio-temporel, mais divisé en deux : on trouverait d'un côté les aspects accessibles à la conscience, et de l'autre ceux qui lui résistent.

2. Nous proposons de caractériser la thèse 2 comme étant une forme d'*externalisme modéré* et la thèse 3 une forme d'*externalisme fort*. Comme nous le verrons, l'externalisme modéré, contrairement à ce que semblent présupposer des anti-cartésiens récents comme McCulloch et Rudder Baker, constitue déjà une forme d'opposition cohérente à la philosophie de l'esprit cartésienne. En effet, la thèse 2, qui revendique le fait que la pensée est ancrée sur des facteurs externes et plus précisément contextuels, est incompatible avec l'internalisme cartésien.

Considérons les conditions dans lesquelles nous exerçons d'ordinaire notre pensée. Comme l'a bien vu Wittgenstein (1953), la pensée est étroitement liée à des *capacités* déterminées (et il est probable que cela vaille également pour la perception). Une capacité est de l'ordre d'une disposition, et on sait combien Ryle a également insisté sur le caractère dispositionnel de la pensée dans le cadre de sa critique de

Descartes. Si la pensée relève au moins partiellement d'une disposition, elle ne peut pas être identifiée, comme le fait apparemment Descartes, à un processus ou état conscient[20]. Par exemple, la compréhension d'un mot n'est pas déterminée, ou du moins pas exclusivement, par les états de conscience du sujet; elle doit pouvoir se *manifester* dans ce que le sujet est capable de faire. Le sujet linguistiquement compétent doit savoir comment agir *en conséquence* de sa compréhension, et en fonction de ses intérêts. Pour reprendre l'exemple de Ryle, la croyance que la glace d'une patinoire naturelle est dangereusement mince « consiste à ne pas hésiter à se répéter ou à dire à autrui qu'elle est mince, à agréer lorsque d'autres personnes l'affirment, à refuser d'accepter des énoncés disant qu'elle ne l'est pas, à tirer les conséquences de la proposition première et ainsi de suite. Mais c'est également avoir tendance à patiner prudemment, à frémir en imaginant tous les désastres possibles et à avertir d'autres patineurs » (1978 : 131). Ce sont les conséquences théoriques et pratiques de la croyance qui en déterminent le contenu, et non pas les états de conscience du sujet de cette croyance.

La thèse selon laquelle la pensée est une capacité à agir de multiples façons représente sans doute une difficulté pour une interprétation spécifique de la notion cartésienne de conscience, mais à première vue et en tant que telle, elle est compatible avec l'internalisme cartésien. En effet, le philosophe cartésien essaiera d'interpréter le « comportement » étroitement lié à la pensée selon l'argument de Ryle et de Wittgenstein en termes non pas d'actions corporelles, mais de volonté et d'intentions spirituelles. La compréhension d'un mot serait bien liée à toutes sortes d'intentions d'agir, mais c'est l'aspect purement mental de ces intentions qui serait pertinent, et non pas leur dimension corporelle, extrinsèque aux intentions elles-mêmes selon l'internalisme cartésien. Nous ne prétendons pas que cette position soit véritablement satisfaisante, mais que le problème le plus sérieux que pose la thèse dispositionnelle pour l'internaliste cartésien est ailleurs.

Voici le problème. Dans certains cas, il est clair que les capacités qui contribuent à définir la pensée ne peuvent s'exercer que dans des conditions « externes » favorables. Pour que nos pensées puissent seulement s'exercer, certaines conditions contextuelles doivent être

20 Toutefois, Descartes assimile à un moment les idées innées à des dispositions : cf. ses *Notae in programma* (1989 vol. III : 807sqq).

remplies; il faut que le contexte «extérieur» manifeste une relative stabilité à travers le temps. On peut donner ici trois exemples assez différents, mais qui mettent tous en évidence cette dépendance de la pensée par rapport au contexte.

Le phénomène de l'indexicalité. Il est probable que la maîtrise des concepts indexicaux, c'est-à-dire les concepts exprimés typiquement par des termes comme «ici» ou «maintenant», soit liée à la capacité de s'orienter dans l'espace et dans le temps – à «garder la trace» d'un endroit dans l'espace, ou d'un moment dans le temps. Pour appliquer ces concepts, le sujet doit exercer une capacité de ce genre. Mais la réussite de cet exercice dépend non seulement du talent du sujet pensant, mais aussi et surtout de circonstances «externes»: sommairement, le sujet doit disposer d'un certain nombre de repères spatio-temporels dans un environnement relativement stable pour que sa pensée indexicale puisse seulement avoir prise. Il s'ensuit que le sujet peut se tromper sur la façon dont s'exerce sa propre capacité; il peut penser qu'elle s'exerce correctement alors qu'en fait, ce n'est pas le cas. Par suite, comme l'exercice de la capacité est une condition nécessaire de la formation de la pensée, le sujet peut se tromper sur la nature de sa propre pensée[21].

Les concepts de couleur. Wittgenstein note à un certain moment que l'utilisation des termes de couleur, et par suite nos pensées mêmes sur la couleur, dépendent de la présence stable, dans l'environnement, de certains objets qui peuvent être considérés comme des échantillons conservant leur couleur sur une période substantielle. Comme le suggère Wittgenstein, si tous les objets changeaient constamment de couleur, nous serions incapables d'avoir des pensées déterminées sur la couleur des objets. Ce n'est pas simplement que nos pensées sur la couleur d'objets particuliers changeraient de valeur de vérité d'un instant à l'autre; c'est le contenu même de nos pensées qui serait menacé[22].

21 Cf. EVANS (1993: 127): «Qu'y a-t-il de si absurde dans l'idée [non cartésienne!] qu'il y a des pensées que l'on ne peut former que parce que l'on occupe une position particulière dans l'espace, ou dans le temps, ou parce que l'on est en train de percevoir un objet?». Notons que le point soulevé ici par Evans est en principe compatible, contrairement à ce qu'il paraît affirmer, avec la théorie de Perry qu'il critique. (Sur la théorie de l'indexicalité de ce dernier, cf. 1993).

22 Cf. WITTGENSTEIN (1989). Pour une discussion de ce point chez Wittgenstein, cf. PEARS (1993: 370).

La division du travail linguistique. D'autres formes de dépendance contextuelle existent, moins spécifiques mais peut-être plus importantes. Certaines pensées – selon un certain point de vue, *toutes* les pensées – dépendent de l'existence d'un langage public. Il est fort possible que ce que chaque locuteur d'un langage est capable d'exprimer et de comprendre ne soit pas indépendant des ressources à la disposition d'autres locuteurs du langage, de sorte qu'il existerait, comme le dit Putnam, une «division du travail linguistique» (1975) ou, selon l'expression de Kaplan, un «partage sémantique» (1989: 601). Mais si les pensées, ou du moins certaines d'entre elles, sont essentiellement linguistiques, elles dépendent d'un environnement social déterminé. Elles restent des états mentaux au sens précis où elles sont formées par un individu qui possède une certaine compétence linguistique, mais celle-ci ne peut s'exercer que si l'individu est correctement placé dans une communauté sociale et linguistique.

Pour résumer ce deuxième point, la thèse caractéristique de l'internalisme cartésien selon laquelle la pensée est indépendante de tout facteur «externe» a été récemment attaquée sur deux fronts différents, et souvent par référence explicite à Descartes. D'une part, on fait valoir que la formation de certaines pensées est étroitement liée à des capacités cognitives déterminées dont l'exercice dépend de la position unique du sujet dans son environnement spatio-temporel. D'autre part, on cherche à montrer que certaines pensées sont essentiellement linguistiques, de sorte que leur formation dépend du rôle joué par le sujet dans la communauté de locuteurs à laquelle il appartient. De manière générale, la nature intrinsèque de la pensée dépend de facteurs «externes», c'est-à-dire non mentaux selon les critères internalistes. Du point de vue externaliste, le sujet pensant n'est pas autonome par rapport au monde extérieur: il doit pouvoir *compter*, la plupart du temps de manière tacite, sur le fait que son environnement extérieur est structuré de diverses façons, sans quoi certaines pensées lui seraient inaccessibles. Le fait que nous sommes capables de telles pensées atteste indirectement de l'existence d'un monde extérieur structuré, que les structures en question soient des formes de vies (Wittgenstein, McCulloch), des jeux de langage (Wittgenstein, Putnam, Burge), des contingences factuelles (Evans), des essences naturelles (Putnam), ou tout cela à la fois[23].

23 On peut mentionner aussi la thèse plus controversée de l'externalisme biologique, selon laquelle l'identité de nos pensées dépend de certaines fonctions

3. Un externaliste peut en principe faire sienne la thèse 2 mais rejeter la thèse 3. Toutefois, cette position n'est possible que si les conditions « externes » nécessaires à la formation de la pensée selon la thèse 2 n'incluent ni la condition que l'objet de la pensée existe, ni directement l'objet de la pensée lui-même. Les tenants de la thèse 3 insistent sur le caractère *relationnel* de la pensée. Selon eux, la pensée n'est pas un état mental autonome par rapport à la réalité extérieure ; au contraire, c'*est* une relation robuste entre le sujet et l'objet de sa pensée. Elle est robuste au sens spécifique où si la pensée est donnée, l'objet pensé est, par nécessité, également donné. Il est vrai que la thèse 2 implique déjà que la formation de la pensée dépend essentiellement de certaines relations au contexte « externe » ; toutefois, elle n'implique pas que la pensée elle-même soit une relation robuste dont les termes seraient respectivement le sujet et l'objet pensé[24].

On peut distinguer deux versions de l'externalisme fort. Selon une version, la pensée dépend de l'existence de l'objet pensé : comme toute relation, l'existence de la pensée dépend de celle de ses termes. Jean ne peut pas être plus âgé que Pierre si Pierre ou Jean n'existent pas. Selon une version plus forte, la pensée dépend de l'*identité numérique* de l'objet pensé dans le cas de la pensée singulière, et de l'espèce naturelle dans le cas de certaines pensées générales. Par exemple, ma pensée singulière sur Pierre ne pourrait *en aucune situation contrefactuelle* porter sur Jean, parce que Pierre et Jean sont deux personnes numériquement distinctes. Cette dernière version, surtout lorsqu'elle concerne les pensées singulières, est aujourd'hui associée aux noms de Gareth Evans et de John McDowell. On peut noter en passant que ces auteurs tiennent à souligner leur opposition à Descartes, par exemple lorsque le second écrit que l'une des raisons qui motivent la version forte de la thèse 3 est « l'intérêt de ses implications radicalement anti-cartésiennes » (1986 : 146).

L'externalisme modéré (thèse 2) et l'externalisme fort (thèse 3) peuvent diverger sur la question des pensées vides (ou, dans le cas perceptif, des illusions). La thèse générale de la dépendance de la

biologiques qui dépendent à leur tour d'une histoire évolutionniste appropriée : cf. MILLIKAN (1984 et 1993) et PAPINEAU (1987). Pour une discussion critique de cette thèse, cf. ENGEL (1992 : ch. 5).

24 Pour être précis, donc, il faut établir une distinction entre la notion d'une pensée qui *est* une relation robuste à son objet externe (externalisme fort) et celle d'une pensée qui *implique* des relations robustes au monde extérieur (externalisme modéré).

pensée par rapport à des facteurs « externes » est compatible avec une certaine théorie des pensées vides ou des illusions, c'est-à-dire des états mentaux sans objet (par exemple, la pensée du jeune enfant selon laquelle la Bonne Fée va échanger sa dent de lait contre un sou, ou mon hallucination d'un petit homme vert). Il est en principe possible d'affirmer que les états mentaux sans objet sont les *mêmes*, du point de vue d'une théorie substantielle de l'identité de la pensée, que d'autres états mentaux dirigés vers des objets réels (par exemple, la même pensée de l'enfant si la Bonne Fée existait, et ma perception du petit homme vert s'il était réel). L'existence d'un objet ne serait pas une condition *nécessaire* de la formation de la pensée ou de la présence d'une expérience. Selon l'externalisme fort, la nature des états mentaux dépend essentiellement de l'existence de leur objet, quel qu'il soit. Une pensée non vide (respectivement, une perception véridique) ne pourrait jamais être identique, dans une autre situation, à une pensée vide (respectivement, à une illusion).

De manière générale, les deux formes d'externalisme divergent sur la question de savoir quelles situations contrefactuelles sont pertinentes pour déterminer la nature de notre pensée. Considérons l'exemple de Putnam, dans lequel le terme « eau » désigne une espèce naturelle, qui possède la propriété essentielle d'être composée surtout de molécules H_2O. Par cet exemple, Putnam attire notre attention sur l'intuition suivante : si nous énoncions la phrase « Voici de l'eau » sur Terre-Jumelle[25], en présence d'un liquide ayant les mêmes propriétés phénoménales que l'eau (c'est-à-dire les propriétés d'être incolore, inodore, d'étancher la soif, etc.) mais une autre structure moléculaire (disons X_YZ, où $X_YZ \neq H_2O$), notre énoncé serait littéralement *faux*. Ce n'est pas l'apparence phénoménale de l'eau qui fixe l'extension du terme « eau », mais son essence en tant qu'espèce naturelle – sa structure moléculaire. Putnam présente sa fiction de Terre-Jumelle comme un contre-exemple à la conception internaliste de la pensée comme étant invariante sous les changements contextuels. Seules les deux formes d'externalisme que nous avons distinguées semblent capables de rendre compte de l'intuition signalée par Putnam. Autrement dit, elles ratifient comme vrai l'énoncé contrefactuel suivant :

25 Terre-Jumelle est une planète fictive inventée par Putnam (1975), exactement similaire à la Terre, sauf qu'à la place de l'eau, on y trouve un liquide qui lui ressemble mais qui possède une toute autre structure moléculaire. L'exemple de Putnam est devenu un *locus classicus* dans la philosophie du langage et de l'esprit contemporaine.

C1. Si j'étais transporté dans un autre environnement, où les océans seraient composés surtout d'autres molécules que celles dont sont composés surtout les océans terrestres, ma phrase « Voici de l'eau », utilisée dans le sens qu'elle a effectivement, pourrait être fausse.

Selon la position externaliste, le concept que *nous* exprimons avec le terme « eau » est ancré sur un contexte terrien, qui fixe son identité en tant que concept. Le désaccord entre les tenants des différentes formes d'externaliste porte sur la question de la relation entre le contexte et l'entité désignée par « eau » : l'externaliste modéré prétend, contrairement à l'externaliste fort, que le contexte n'inclut pas l'existence de cette entité. Faisons l'hypothèse que le terme « eau » aurait été *vide* si nos rivières, nos lacs et nos océans avaient contenu un liquide semblable à l'eau mais doté d'une structure moléculaire instable[26]. La question tourne alors autour du statut de l'énoncé contrefactuel suivant :

C2. S'il y avait eu dès le départ dans notre environnement un liquide d'une composition moléculaire instable, nous aurions peut-être acquis le même concept d'eau, sauf qu'il aurait été vide.

L'externalisme modéré peut accepter cet énoncé contrefactuel, à condition que l'ancrage au contexte du terme « eau », dans la situation contrefactuelle envisagée, partage certaines propriétés pertinentes avec l'ancrage au contexte du terme « eau » dans notre situation réelle. L'attitude modérée n'est pas en soi incohérente : nous « gardons la trace » de l'eau en « gardant la trace » des régions terrestres dans lesquelles nous croyons trouver le même liquide. Le fait que notre concept de l'eau soit vide ou qu'il désigne effectivement une espèce naturelle n'a rien à voir avec le contexte au sens de l'externaliste modéré, mais avec ce qu'on va réellement y trouver. Par opposition, l'externaliste fort doit rejeter l'énoncé contrefactuel C2, en arguant que celui qui le défend sous-estime la *robustesse* de la relation entre le concept « eau » et ce qu'il désigne[27].

26 Le lecteur est prié de changer l'exemple s'il le considère comme peu plausible. L'une des difficultés de l'expérience de pensée de Putnam est qu'elle fait intervenir un terme ordinaire (« eau ») avant de disposer d'une théorie substantielle du concept exprimé par ce terme. Mais les enjeux devraient apparaître clairement malgré le caractère éventuellement arbitraire de l'exemple.

27 Notons que les deux versions de l'externaliste *fort* divergent sur le statut d'un autre énoncé contrefactuel, du type : « S'il y avait eu dès le départ dans notre environnement un liquide d'une autre composition moléculaire, nous aurions peut-être acquis le même concept d'eau ». Si on pense que le concept d'eau dépend seulement de l'existence d'une espèce naturelle déterminée, on peut à

Nous ne saurions entrer ici dans les détails de la controverse entre les différentes formes d'externalisme. Il suffit peut-être de mentionner que si l'énoncé contrefactuel C1 correspond bien à une intuition de notre part, mise en évidence par Putnam, il n'est pas clair qu'il en aille de même pour C2. C1 suppose une lecture *intra-mondaine* de l'exemple de Putnam, dans laquelle Terre et Terre-Jumelle coexistent dans le même monde possible, mais pas au même endroit, alors que C2 suppose une lecture *inter-mondaine* de cet exemple, dans laquelle Terre et Terre-Jumelle existent dans des mondes possibles différents, mais au même endroit. Or notre intuition est que si nous étions *transportés* sur Terre-Jumelle, notre énoncé « Voici de l'eau » en présence du liquide de Terre-Jumelle serait faux. Notre intuition est moins catégorique lorsqu'il s'agit d'imaginer un autre contexte *à la place* de celui dans lequel nous avons effectivement introduit le terme « eau ».

Nous pouvons à présent résumer notre discussion des trois thèses spécifiquement dirigées contre l'internalisme cartésien. Nous avons vu que la première thèse, qui concerne le rapport entre le mental et la conscience, n'est pas forcément externaliste, du moins si nous admettons contre la position cartésienne que la conscience réflexive n'a pas le dernier mot sur la nature intrinsèque des états mentaux. Mais il est clair que cette thèse est un premier pas en direction de l'externalisme représenté par l'ensemble des thèses 2 et 3, ou par l'une d'entre elles. Un externaliste modéré qui n'accepte que la thèse 2 est déjà en mesure de remettre en question l'image cartésienne d'un espace mental intérieur non concerné par les variations extérieures. Les configurations mentales que l'on trouve dans cet espace soi-disant intérieur dépendent en fait essentiellement de l'environnement extérieur, naturel ou social. Si nous acceptons la thèse 2 alors, comme le dit McDowell, « nous sommes forcés de nous imaginer que les domaines internes et externes s'interpénètrent, et ne sont pas séparés par une frontière typiquement cartésienne » (1986 : 150). Il n'y a pas de ligne de partage tranchée entre ce qui est mental ou intérieur, et ce qui est extérieur[28].

la limite accepter ce type d'énoncé contrefactuel. On doit le refuser, en revanche, si on pense que le concept d'eau dépend non seulement de l'existence d'une espèce naturelle déterminée, mais aussi du fait que cette espèce ne peut être qu'un liquide composé surtout de molécules d'H_2O.

28 La formulation de McDowell suggère par ailleurs que la distinction entre internalisme et externalisme n'est pas la même que celle entre réalisme et idéalisme. L'externaliste laisse ouverte la question du statut ontologique de

Les différences entre les thèses 2 et 3 concernent surtout la question de savoir comment individuer les pensées. Cette question taxinomique est liée au problème des *critères d'identité* de la pensée, et des intérêts objectifs que nous avons lorsque nous considérons que deux pensées particulières sont les mêmes ou pas. Il vaut peut-être la peine de mentionner en passant l'idée récemment avancée par Barwise et Perry (1983) selon laquelle il n'existe pas de critères et d'intérêts *uniques* en la matière. Ces auteurs nous mettent en garde contre ce qu'ils appellent « le sophisme de l'information mal placée », selon lequel il y aurait une seule notion de pensée qui comprendrait tous les rôles que l'on fait normalement jouer à nos états mentaux. Peut-être la pensée que l'on attribue au sujet pour expliquer son comportement n'est-elle pas forcément celle qu'on lui prête pour évaluer ses croyances comme étant vraies ou fausses, par exemple. Selon Perry, l'explication du comportement fait référence à un « rôle cognitif » censé par exemple expliquer pourquoi tous les sujets qui pensent « Je suis en train d'être attaqué par un ours » se comportent de la même manière (*caeteris paribus*), c'est-à-dire en prenant ses jambes à son cou. Mais nous devons aussi faire référence au « rôle sémantique » de la pensée, censé expliquer pourquoi certaines pensées exprimées par « Je suis en train d'être attaqué par un ours » sont vraies et d'autres fausses. Cette théorie admettrait *plusieurs* niveaux d'analyse du mental qui correspondraient à des intérêts objectifs différents.

Mais la théorie est *ipso facto* anti-cartésienne si elle admet qu'au moins *un* des niveaux d'analyse du mental dépend de facteurs « externes ». Or le rôle sémantique de la pensée, qui chez Perry prend la forme d'une proposition singulière qui contenant directement les objets et les propriétés pensés par le sujet, est au moins partiellement « externe » en ce sens. En fait, selon une certaine interprétation de Perry, le rôle cognitif est également « externe », puisqu'il est identifié essentiellement par les phrases du langage public que le sujet pensant est disposé à accepter. De ce point de vue, même le terme indexical « je », tout égocentrique soit-il, dépend de la capacité du sujet à se comporter de manière appropriée dans sa communauté linguistique, par exemple d'utiliser le terme « je » à bon escient, ou de comprendre autrui lorsqu'il utilise le même terme. Contrairement à ce qu'affirme

ce qui est « extérieur » ; il milite seulement contre l'idée qu'on puisse séparer nettement les éléments « extérieurs » et « intérieurs » de la réalité quelle qu'elle soit.

McCulloch (1995: 191), la division de la pensée en rôle cognitif et rôle sémantique n'est pas *en tant que telle* foncièrement cartésienne. En tout cas, une telle division, si elle pouvait se justifier, nous permettrait peut-être de concilier différentes formes d'externalisme pour un domaine donné de pensées. S'il y a plusieurs aspects mentaux d'une pensée, il n'est pas garanti *a priori* qu'ils dépendent de la même façon du monde extérieur[29].

L'un des intérêts de reconnaître plusieurs aspects du mental réside précisément dans la possibilité de rendre compte de certaines intuitions cartésiennes sans assimiler les attitudes psychologiques à des états intérieurs au sens cartésien. En particulier, on peut en principe expliquer comment les attitudes ont un pouvoir causal sur le monde extérieur. Le comportement d'un agent dépend causalement de ses attitudes. Ces dernières peuvent expliquer le comportement sans toutefois être assimilées à des états cérébraux. Deux agents se comportent de la même façon, *caeteris paribus*, s'ils forment les mêmes pensées-types. Notre explication psychologique pourrait prendre la forme suivante: deux agents forment les mêmes pensées-types s'ils sont disposés à employer les mêmes phrases pour décrire et expliquer leur comportement. Puisque les phrases sont des entités sociales qui dépendent de la signification de signes publics, on ne revient pas à l'internalisme cartésien[30].

On voit donc que d'un certain point de vue, les réactions récentes à l'internalisme cartésien, par exemple celles que l'on trouve chez Rudder Baker et McCulloch, paraissent exagérées au sens où elles négligent la possibilité d'un externalisme modéré. Il n'est pas nécessaire pour l'anti-cartésien d'embrasser la position de ces auteurs, à savoir ce que nous avons appelé l'externalisme fort, ou la thèse selon laquelle le contenu de pensée dépend de l'existence ou de l'identité

29 Bien entendu, il est aussi possible de concilier différentes formes d'externalisme en affirmant qu'elles ne concernent pas toutes le même domaine de pensées. De manière générale, la possibilité et l'intérêt d'une conciliation dépendent d'une théorie substantielle de la pensée.

30 Nous ne prétendons pas régler par ces brèves remarques le problème épineux du rôle causal des attitudes psychologiques. Nous insistons seulement sur le fait qu'on ne peut pas inférer directement de la prémisse que les attitudes psychologiques jouent un rôle causal la conclusion qu'elles sont internes au sens cartésien. L'externaliste peut admettre que les attitudes dépendent essentiellement de facteurs externes, mais leur reconnaître un rôle causal susceptible de rendre compte de certaines similarités de comportement dans différents contextes.

numérique de l'objet pensé. Du moment que l'on reconnaît l'existence de facteurs « externes » qui concourent à la formation d'une attitude mentale, on accepte une certaine forme d'externalisme et, par là-même, on est anti-cartésiens. La thèse P4 de l'autonomie de la pensée est alors rejetée, ce qui constitue déjà, à première vue, une opposition cohérente au cartésianisme.

4. Une réaction cartésienne ?

De la part d'un défenseur de l'internalisme cartésien, une réaction naturelle contre les différentes formes d'externalisme présentées dans la section précédente consiste à adopter ce qu'on pourrait appeler une « stratégie de déconstruction ». Celui qui adopte cette stratégie va chercher à démêler, parmi les conditions nécessaires à la formation de la pensée, celles qui sont externes et celles qui sont internes. Par exemple, à une analyse externaliste du type :

(1) S a la pensée que *p* si et seulement si C,

où S est un sujet pensant et C des conditions qui dépendent d'une certaine façon de l'environnement social ou naturel du sujet, l'internalisme tentera de substituer :

(2) S a la pensée que *p* si et seulement si C_i et C_e,

où C_i et C_e correspondent à des conditions logiquement indépendantes qu'il juge respectivement « purement internes » et « purement externes ». On peut donner ici un ou deux exemples de déconstruction.

Considérons la thèse externaliste selon laquelle la pensée est essentiellement linguistique et dépend d'échanges sociaux entre locuteurs d'une même communauté linguistique. La stratégie de déconstruction, dans ce cas, consiste à séparer dans ce qui apparaît comme une dépendance mentale par rapport à des facteurs sociolinguistiques, les aspects « externes » et les aspects « internes ». Les aspects « internes » concernent la pensée privée que le sujet *associe* personnellement à ces mots, alors que les aspects « externes » pourraient inclure les mots en question et leur usage dans la communauté linguistique pertinente. Selon l'internaliste, les aspects internes sont tout à fait indépendants des aspects externes. Autrement dit, là où l'externaliste défend une thèse du type :

(3) S exprime la pensée que *p* en utilisant la phrase *s* avec une intention littérale seulement s'il appartient à une certaine communauté linguistique, dans laquelle la phrase *s* a la signification publique que *p*.

l'internaliste propose une déconstruction de ce genre :

(4) S exprime la pensée que *p* en utilisant la phrase *s* avec une intention littérale seulement s'il associe à *s* une certaine pensée privée, et si la pensée que *p* est déterminée d'une certaine façon comme la « somme » de cette pensée et des pensées privées associées à *s* par les autres membres de la communauté linguistique en question (quand ils utilisent la phrase *s* avec une intention littérale).

(4) n'implique pas forcément que chaque locuteur compétent associe une pensée similaire à la même phrase. Dans cet exemple, la stratégie de déconstruction débouche sur une position internaliste spécifique communément appelée *mentalisme* ou *psychologisme*. De nos jours, cette position et ses variantes diverses ont été critiquées avec virulence par Dummett (1991), qui les appelle les conceptions du langage comme un code.

Considérons, sur un plan plus général, un autre exemple de déconstruction, dirigée contre la thèse de l'externalisme fort selon laquelle la pensée est relationnelle (thèse 3 ci-dessus). Là où l'externaliste affirme :

(5) S pense à *x* si et seulement si R(S,*x*),

où R est une relation qui définit la nature intrinsèque de la pensée, l'internalisme propose la déconstruction suivante :

(6) S pense à *x* si et seulement si M(S) et R'(S,*x*).

où M est un état mental « interne » du sujet, entièrement spécifiable indépendamment de son objet *x*, alors que R' est une relation externe appropriée (par exemple, causale) entre le sujet et l'objet de sa pensée[31] (Dans cet exemple également, les conditions M(S) et R'(S,*x*) sont censées être logiquement indépendantes). L'internaliste substitue cette dernière analyse à la proposition externaliste en stipulant que toutes deux doivent avoir les mêmes conditions de vérité.

Pour reprendre l'exemple de Putnam, la jalousie est un état mental relationnel, au sens où l'on est toujours jaloux *de* quelqu'un. Or l'internaliste cherchera à montrer que lorsqu'on attribue la jalousie à quelqu'un, on lui prête en réalité un état affectif autonome par rapport au monde extérieur, en ajoutant que cet état est causalement ancré sur une personne particulière, qui devient l'objet de la jalousie. Soit dit en

31 Cf. EVANS, 1982 : 200-1.

passant, Putnam lui-même exprime quelque scepticisme sur la possibilité d'identifier un tel état affectif.

Il est peut-être utile de signaler que l'externalisme modéré peut proposer à son tour une déconstruction entre les facteurs sémantiques et les facteurs cognitifs ou mentaux d'une attitude donnée. Cependant il ne suppose pas, contrairement à l'internaliste, que ces derniers sont indépendants de tout facteur «extérieur». Dans la déconstruction de l'externaliste modéré, l'état mental M du sujet S dépend de facteurs contextuels. M peut être spécifié indépendamment de l'objet *x* sur lequel il porte, mais pas de manière purement «interne», sans recourir aux facteurs contextuels qui rendent sa formation possible. On voit une fois de plus que cette forme d'externalisme est suffisante, *pace* Rudder Baker et McCulloch, pour rejeter le cartésianisme.

Notons pour terminer quelques difficultés liées à la stratégie internaliste :

(i) La réussite de la stratégie internaliste n'est pas garantie *a priori*. Il est théoriquement possible que la relation R dans l'analyse externaliste soit primitive, et donc incapable d'être décomposée selon la recette internaliste dans toutes les situations dans lesquelles elle est exemplifiée. L'internalisme devient alors au mieux un programme de recherche, et non plus une thèse *a priori*.

(ii) A supposer que la décomposition internaliste de la relation R soit possible, il faut donner un critère convaincant qui permette de qualifier l'état M de «mental». Si M n'est pas un état mental, la contre-proposition internaliste manque sa cible, qui était de préserver un espace mental intérieur indifférent aux contingences mondaines. Ce deuxième point nous semble être la difficulté la plus sérieuse liée à la position internaliste.

Premièrement, l'état M n'est pas mental conformément aux principes de la psychologie ordinaire [*folk psychology*], puisqu'il ne peut pas justifier à lui seul l'attribution de la pensée que *p* (la relation «externe» R' doit également être exemplifiée). Du point de vue internaliste, une proposition du type «S pense que *p*» ne décrit pas *stricto sensu* un état mental, mais un ensemble hétérogène de conditions dont certaines seulement sont mentales.

Deuxièmement, la déconstruction internaliste s'expose à une objection d'ordre phénoménologique[32]. Certaines descriptions phénomé-

32 Cf. McCulloch, 1995 : 195.

nologiques, par exemple, semblent être incompatibles avec une analyse de type internaliste : la jalousie se présente comme un sentiment essentiellement dirigé vers une personne particulière, et non pas comme un état affectif personnel qui engagerait exclusivement le sujet jaloux à défaut de la personne jalousée. (Il s'agirait donc, dans le cas de la jalousie, d'une position externaliste forte).

A ce stade, le Cartésien pourrait objecter que dans le processus mental qui consiste à « simuler » la jalousie, c'est-à-dire à imaginer *en pensée* un contexte dans lequel la jalousie a prise, ce sentiment est bien réel même si la personne jalousée n'existe pas, ou si d'autres conditions externes nécessaires au sentiment de jalousie ne sont pas remplies. Ce n'est donc pas le contexte réel qui est pertinent, mais le contexte simulé ou imaginé (qui peut mais qui ne doit pas correspondre à un contexte réel). Cette objection cartésienne suppose toutefois l'idée controversée selon laquelle la jalousie imaginée est qualitativement similaire à la jalousie réelle, ce qui reste précisément à démontrer. (Des remarques analogues sont valables pour la pensée conceptuelle).

Troisièmement, l'état M ne peut pas être qualifié de « mental » uniquement parce qu'il est accessible à la conscience réflexive, comme on l'a vu plus haut. Ce n'est pas parce que nos expériences « paraissent » identiques au sujet conscient qu'ils le sont réellement, c'est-à-dire au point de vue de leur nature mentale intrinsèque.

On voit donc que le critère du mental sur lequel se base l'internalisme cartésien n'est pas clairement établi. Or en l'absence d'un critère adéquat, l'externalisme sous une forme ou une autre reste un concurrent philosophique sérieux.

BIBLIOGRAPHIE

ARMSTRONG, D., 1980, *The Nature of Mind*, University of Queensland Press, Brisbane.

BAERTSCHI, B., 1992, *Les rapports de l'âme et du corps. Descartes, Diderot et Maine de Biran*, Vrin, Paris.

BARWISE, J., et Perry, J., 1983, *Situations and Attitudes*, MIT, Cambridge, Mass.

BURNYEAT, M., 1982, « Idealism and Greek Philosophy: What Descartes Saw and Berkeley Missed », in : G. VESEY (éd.), *Idealism Past and Present*, CUP, Cambridge.

CHURCHLAND, P.M., 1989, *A Neurocomputational Perspective : The Nature of Mind and the Structure of Science*, MIT, Cambridge, Mass.

DAVIDSON, D., 1993, *Actions et événements*, trad. franç. P. ENGEL, PUF, Paris.

DESCARTES, R., 1989, *Œuvres philosophiques*, trois volumes, F. ALQUIÉ (éd.), Bordas, Paris.

DRETSKE, F., 1969, *Seeing and Knowing*, University of Chicago Press, Chicago.

DUMMETT, M., 1991, *Les origines de la philosophie analytique*, trad. franç. M.-A. LESCOURRET, Gallimard, Paris.

ENGEL, P., 1992, *Etats d'esprit. Questions de philosophie de l'esprit*, Alinea, Aix-en-Provence.

EVANS, G., 1982, *The Varieties of Reference*, OUP, Oxford.

— 1993, « Comprendre les démonstratifs », in : E. CORAZZA et J. DOKIC, *Penser en contexte. Le phénomène de l'indexicalité*, Editions de l'Eclat, Combas.

FODOR, J., 1988, *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*, MIT, Cambridge, Mass.

HARMAN, G., 1982, « Conceptual Role Semantics », *Notre Dame Journal of Formal Logic*, xxii.

HORNSBY, J., 1986, « Physicalist Thinking and Conceptions of Behaviour », in : J. MCDOWELL et P. PETTIT (eds).

KAPLAN, D., 1989, « Afterthoughts », in : ALMOG, J., PERRY, J., WETTSTEIN, H. (eds.), *Themes from Kaplan*, OUP, Oxford.

LEWIS, D., 1983, « Mad Pain and Martian Pain », in *Philosophical Papers*, vol. 1, OUP, Oxford.

LOAR, B., 1981, *Mind and Meaning*, CUP, Cambridge.

MCCULLOCH, G., 1995, *The Mind and its World*, Routledge, Oxford.

MCDOWELL, J., 1986, « Singular Thought and the Extent of Inner Space, » in : J. MCDOWELL et P. PETTIT (eds).

— et PETTIT, P. (eds.), 1986, *Subject, Thought, and Context*, Clarendon Press, Oxford.

MILLIKAN, R.G., 1984, *Language, Thought, and Other Biological Categories*, MIT, Cambridge, Mass.

— 1993, *White Queen Psychology and Other Essays for Alice*, MIT, Cambridge, Mass.
PAPINEAU, D., 1987, *Reality and Representation*, Blackwell, Oxford.
PEARS, D., 1993, *La pensée-Wittgenstein : du Tractatus aux Recherches Philosophiques*, trad. franç. C. CHAUVIRÉ, Aubier, Paris.
PERRY, J., 1993, *The Problem of the Essential Indexical and Other Essays*, OUP, Oxford.
PUTNAM, H., 1975, « The Meaning of 'Meaning' », in : *Mind, Language, and Reality : Philosophical Papers*, vol. 2, CUP, Cambridge.
RUDDER BAKER, L., 1995, *Explaining Attitudes*, CUP, Cambridge.
RUSSELL, B., 1956, *The Philosophy of Logical Atomism*, in *Logic and Knowledge*, R. C. MARSH (ed.), Allen et Unwin, Londres.
RYLE, G., 1978, *La notion d'esprit*, Payot, Paris.
WILLIAMS, B., 1978, *Descartes*, Penguin Books, Harmondsworth.
WITTGENSTEIN, L., 1953, *Philosophical Investigations*, CUP, Cambridge.
— 1989, *Notes sur l'expérience privée et les « sense data »*, trad. franç. E. RIGAL, 2e édition, TER, Mauvezin.