

Pascal Ludwig – Putnam sur l'expérience consciente

Introduction

Le but de cet article est d'examiner le changement de perspective d'Hilary Putnam sur le rôle épistémologique des expériences sensorielles conscientes entre 1981 et 1994¹. Alors qu'il rejette le réalisme externe en 1981, défini comme la thèse selon laquelle les objets physiques possèdent une nature indépendamment de l'esprit humain, Putnam adhère en 1994 à une forme de réalisme qu'il appelle le « réalisme naturel ». La stratégie argumentative de Putnam en 1981 est de faire naître dans un premier temps ce que James Conant a décrit comme une « angoisse cartésienne », puis de montrer qu'une angoisse plus profonde, une « angoisse kantienne », découle de l'angoisse cartésienne. Conant distingue en effet entre deux formes que le scepticisme contemporain peut prendre². L'angoisse « cartésienne » naît selon lui d'une interrogation portant sur la question de savoir si nous possédons ou non, actuellement, des connaissances. Ainsi, je peux me demander, à la façon de Descartes dans la première *Méditation métaphysique*, si je suis en train de voir ma main, et en conséquence si je possède la connaissance, issue de la perception, selon laquelle j'ai bien une main, ou si je suis en train de rêver. Mais de manière plus radicale, je peux aussi m'interroger sur la *possibilité même* de la connaissance. Le second chapitre de *Raison, vérité et histoire* a pour fonction de susciter une angoisse kantienne à partir de l'angoisse cartésienne, en établissant la vérité du conditionnel suivant: « si le réalisme externe est vrai, nous faisons nécessairement face à une angoisse kantienne impossible à dissiper ». L'argument modèle-théorique joue un rôle central dans la justification de ce conditionnel.

Or Putnam revient sur cet argument, qu'il révisé de façon profonde en 1994. Cette révision repose sur la remise en question de ce qu'il considère alors comme un présupposé majeur, mais injustifié de la stratégie argumentative mise en place au début des années 1980, un présupposé qui porte sur l'expérience sensorielle consciente et sur son rôle dans la justification des croyances. En 1981, Putnam partage avec les

- 1 Entre *Reason, Truth and History* (Cambridge, Mass., Cambridge University Press, 1981 ; *Raison, vérité et histoire*, tr. fr. A. Gerschenfeld, Paris, Minuit, 1984) et l'article « Sense, Nonsense, and the Senses », *The Journal of Philosophy*, 91, 1994, p. 445-417, qui reprend les conférences Dewey données à l'université de Columbia en mars 1994. Ces conférences ont également été publiées sous une forme plus développée dans *The Threefold Cord: Mind, Body, and World*, New York, New York, Columbia University Press, 1999, dont on dispose d'une traduction française : *La triple corde*, trad. fr. P. Fasula, R. Ehrensam, S. Plaud, & J.-M. Roux, Paris, Vrin, 2017.
- 2 Conant J., « Two Varieties of Skepticism », *Rethinking Epistemology*, volume 2, (éd. A. Günther & J. Conant), Berlin, Walter de Gruyter, 2012, p. 1-73.

philosophes de l'âge classique, comme Descartes et Locke, l'idée selon laquelle une expérience sensible consciente véridique procure au sujet qui l'éprouve exactement la même justification qu'une expérience sensible consciente non-véridique. Pour saisir le lien qui existe entre cette conception classique de la fonction épistémique des données sensibles et la stratégie argumentative mise en place dans *Raison, vérité, histoire*, partons du trilemme suivant³ :

1. Les objets physiques de taille moyenne existent indépendamment de notre esprit, de nos états mentaux, et de nos cadres conceptuels (réalisme externe) ;

2. Nos états sensoriels conscients portent sur des objets, et leurs rôles épistémiques dépendent de ces objets (intentionalisme) ;

3. Nos états sensoriels conscients jouent le même rôle épistémique lorsqu'ils sont véridiques et lorsqu'ils ne sont pas véridiques (conception classique des données sensibles).

Dans *Raison, vérité et histoire*, Putnam accepte l'énoncé 3 de ce trilemme, ce qui produit une tension, sinon même une contradiction, entre 1 et 2. Si des états conscients peuvent constituer des raisons de croire indépendamment de leur caractère véridique, il en découle qu'un sujet en train de voir une porte jaune juste devant lui a accès à exactement la même donnée probante qu'un autre sujet ayant une hallucination visuelle d'une porte jaune juste devant lui. En ce sens, les données probantes que nous procurent nos sensations sont internes : elles ne dépendent que de nos états cérébraux, et pas des relations que ces états entretiennent avec la réalité. Mais en même temps, ces états ont des objets selon l'énoncé 2, et leur rôle rationnel dépend de leurs objets. Ainsi, l'hallucination visuelle d'une porte jaune porte bien sur un objet intentionnel, et sa fonction rationnelle – par exemple, le fait qu'elle constitue une bonne raison de croire en la présence d'une porte jaune dans l'environnement – dépend de cet objet. Cela implique qu'il existe une relation référentielle reliant certains aspects de nos expériences sensorielles conscientes et les objets externes.

La fonction argumentative principale de l'argument dit « modèle-théorique » est de montrer que l'existence d'une unique relation référentielle n'est pas compatible avec l'énoncé 1 du trilemme. Mais, du moins selon Putnam lui-même lorsqu'il se relit en 1994, l'argument repose sur la présupposition formulée en 3. Je me propose de commencer par expliquer en détail pourquoi c'est le cas, en essayant d'analyser le plus précisément possible le concept d'expérience sensorielle consciente utilisé par Putnam en 1981, et de

3 Je m'inspire ici de Bill Brewer, sans reprendre exactement ses formulations. Voir Brewer B., « Perception and Content », *European Journal of Philosophy*, 14 (2), p. 165-181.

décrire son rôle dans l'argument anti-réaliste développé dans les trois premiers chapitres de *Raison, vérité, histoire*. Dans un second temps, j'expliquerai en quoi le réalisme « naturel » que Putnam défend en 1994 diffère du réalisme interne, et je montrerai que cette différence découle pour l'essentiel d'une nouvelle conception du rôle rationnel des expériences sensorielles conscientes. Pour finir, je soutiendrai que si le réalisme naturel de Putnam est compatible avec l'intentionalisme, il se démarque nettement du représentationalisme, dont Putnam développe une critique qui reste pertinente aujourd'hui.

1. La place de l'expérience consciente dans *Raison, vérité et histoire*

Les sensations conscientes ont une place singulière dans la métaphysique des états mentaux défendue par Putnam en 1981. Les pensées conceptuelles, comme les croyances ou les intentions, sont identifiées à des rôles causaux métaphysiquement indépendants de leurs implémentations cérébrales, conformément aux grandes lignes du fonctionnalisme que Putnam a été l'un des premiers philosophes à défendre dans les années 1960⁴. Les sensations conscientes constituent néanmoins selon lui des exceptions au fonctionnalisme, en raison, nous allons le voir, des propriétés purement *qualitatives* qu'elles possèdent⁵.

1.1 Inversion de spectre et transparence de l'expérience

Comme Putnam le souligne en 1994 à propos de sa position de 1981 : « Selon moi, les sensations avaient un aspect "qualitatif" qui ne pouvait pas être caractérisé de façon fonctionnelle, et cet aspect devait être identifié à certaines propriétés physiques de notre cerveau⁶ ». C'est avant tout l'argument du « spectre inversé » qui conduit alors Putnam à reconnaître l'existence de propriétés purement qualitatives des expériences conscientes accessibles par l'introspection⁷. Cet argument repose sur la concevabilité de *scenarii* dans lesquels des sujets diffèrent systématiquement relativement aux propriétés qualitatives de leurs expériences conscientes accessibles par introspection, sans différer relativement à ce que leurs sensations représentent, et sans qu'aucun des sujets soit

4 Voir en particulier Putnam H., « Psychological Predicates », in W. H. Capitan & D. D. Merrill (éd.), *Art, Mind, and Religion*, Pittsburgh, University of Pittsburgh Press, 1967, p. 37-48.

5 Putnam écrit explicitement qu'identifier une « qualité présentée », par exemple l'expérience consciente d'une nuance particulière de bleu, à un rôle causal ou fonctionnel (*functional state*), n'est « pas plausible ». Voir Putnam H., *Reason, Truth and History*, p. 79.

6 Putnam H., *Reason, Truth and History*, p. 19.

7 Voir Putnam H., *Reason, Truth and History*, p. 80.

dans l'erreur. Par exemple, il pourrait sembler concevable que Nina et Houria diffèrent, en raison de la façon spécifique dont les mécanismes de la vision sont implémentés dans leurs cerveaux respectifs, quant aux couleurs qui causent en elles les sensations visuelles du rouge et du vert. Il semble en effet qu'on puisse imaginer qu'à chaque fois qu'une tomate rouge est présente dans l'environnement visuel de Nina, la qualité visuelle associée à la sensation de rouge lui soit présentée, mais qu'à chaque fois qu'Houria voit une tomate, ce soit la sensation de vert qui lui soit présentée. Et qu'à chaque fois que Nina voit une salade verte la sensation de vert lui soit présentée, là où la sensation de rouge est présentée à Houria. Si ce scénario d'inversion sans erreur est concevable, la sensation possédant la qualité du vert ne joue pas le même rôle causal dans l'architecture mentale de Nina et dans celle d'Houria : dans un cas, cette sensation est produite par la présence de la couleur verte dans l'environnement, dans l'autre par la présence de la couleur rouge. En 1981, Putnam tient pour acquis que ce type de scénario est bel et bien concevable, et c'est pour cela qu'il ne peut pas identifier les sensations, conçues comme des états mentaux dotés de propriétés qualitatives accessibles directement par introspection, à des rôles fonctionnels.

Je voudrais souligner d'emblée que la conception de l'expérience sensorielle consciente défendue en 1981 par Putnam est incompatible avec un principe important, le principe de transparence des expériences conscientes. La source de ce principe se trouve dans cette remarque suivante de G. E. Moore :

« [...] au moment même où nous essayons de focaliser notre attention sur la conscience et de voir ce qu'elle est de façon distincte, elle semble disparaître : tout se passe comme si nous avions un vide pur et simple face à nous. Lorsque nous essayons de connaître la sensation de bleu par introspection, tout ce que nous parvenons à voir c'est le bleu : l'autre élément apparaît diaphane⁸. »

Moore souligne dans ce texte que l'expérience semble transparente, au sens où il semble difficile d'en faire une cible pour l'attention. Lorsque nous essayons de prêter attention aux propriétés intrinsèques d'une sensation de couleur, par exemple, il semble que nous prêtons en fait attention à une couleur externe, la propriété visible d'une surface.

Cette intuition de la transparence de l'expérience a fait l'objet de très nombreuses discussions récentes. Je considère pour ma part qu'on peut la formuler au travers du principe suivant, portant sur la *connaissance introspective*

8 Voir Moore G. E., « The Refutation of Idealism », *Mind*, 12 (48), p. 450.

Principe de transparence de l'expérience : lorsque nous formons une connaissance introspective portant sur les propriétés de nos expériences sensorielles conscientes, les seuls faits directement accessibles par introspection sont des faits incluant les objets physiques de taille moyenne que nous sommes en train de percevoir (dont notre propre corps et ses parties). Les expériences elles-mêmes ne nous sont pas présentées directement dans l'introspection⁹.

Les *scenarii* d'inversion de spectre ne sont concevables qu'à condition de rejeter le principe de transparence. Reprenons l'exemple de Nina et d'Houria, et supposons qu'elles soient toutes les deux en train d'observer une rose rouge. Selon le principe de transparence, elles n'ont accès à rien d'autre, lorsqu'elles forment une croyance introspective portant sur leur sensation visuelle, qu'à la rose réelle qu'elles observent et à sa couleur apparente. Bien sûr, la couleur de la rose peut *sembler subjectivement* différente à Nina et à Houria, par exemple si Houria observe la rose dans des conditions anormales d'illumination. Mais dans ce cas l'une des deux pourra être considérée comme ayant une perception non-véridique, ou anormale, de la rose. Or les *scenarii* d'inversion insistent sur le fait qu'aucun des sujets ne perçoit anormalement le monde.

En 1981, Putnam semble néanmoins rejeter le principe de transparence, et c'est la raison pour laquelle les *scenarii* d'inversion de spectre lui paraissent concevables. Il considère en effet que *les qualités des expériences nous sont « présentées¹⁰ »*, c'est-à-dire qu'elles sont directement accessibles par l'introspection, indépendamment de ce à quoi les sensations sont causalement associées dans le monde extérieur. À cette époque, Putnam considère qu'il existe une différence essentielle entre la pensée conceptuelle d'un côté, et les états mentaux sensoriels conscients de l'autre, relativement à l'introspection. Les pensées, et les concepts qui en sont des constituants, ne sont pas directement « présentés » dans l'introspection. Il s'agit d'une des conséquences de la réfutation de la théorie magique de la référence au chapitre 1 de *Raison, vérité et histoire*. Les concepts ne sont pas accessibles à l'introspection, car la possession d'un concept suppose que l'agent soit inséré dans un environnement et que ses états mentaux soient reliés causalement aux entités présentes dans cet environnement. En revanche, les sensations sont « présentées » directement au sujet par l'intermédiaire de l'introspection, comme il ressort de ce passage : « lorsque nous nous livrons à l'introspection, nous ne percevons pas de “concepts” traversant notre esprit en tant que

9 Pour une formulation semblable, voir Speaks J., *The Phenomenal and the Representational*, Oxford, Oxford University Press, 2015.

10 Voir Putnam H., *Reason, Truth and History*, p. 19.

tels. Arrêtons le cours de nos pensées quand et où nous le voudrons, nous y trouverons des mots, des images, des sensations, des sentiments ». Supposons que je trouve, par introspection, une certaine image dans mon esprit. Selon le premier chapitre de *Raison, vérité et histoire*, l'introspection seule ne me permet pas d'établir ce à quoi réfère cette image ou ce qu'elle représente — s'il s'agit, pour reprendre un exemple de Putnam, d'une image d'arbre ou d'une image d'un « objet étrange de nature et de fonction inconnues¹¹ ». Néanmoins l'introspection permet au moins d'avoir accès aux propriétés *qualitatives et intrinsèques* de l'image. Putnam ne dit rien de la manière dont cette acquisition introspective de croyances portant sur les sensations peut s'effectuer, mais le fait qu'il utilise l'expression de « présentation » à propos des qualités des sensations suggère que cette acquisition repose sur une forme d'accointance, de relation épistémique spécifique et directe entre le sujet et ses propres expériences¹².

1.2 L'image cartésienne/matérialiste de l'esprit conscient

Lorsqu'il revient, en 1994, sur la stratégie argumentative de *Raison, vérité et histoire*, Putnam souligne que cette démarche présuppose une conception cartésienne/matérialiste de l'esprit : « mon image de notre fonctionnement mental n'était autre que l'image cartésienne/matérialiste, une image selon laquelle il ne peut que paraître magique que nous ayons accès à quoi que ce soit d'extérieur à nos "entrées" [*inputs* – ces "*qualia*" dont je pensais qu'on pouvait les identifier à des occurrences physiques¹³ ». Il s'agit d'une théorie cartésienne car les qualités des sensations sont directement accessibles par l'introspection ; mais c'est aussi une théorie matérialiste, puisque ces propriétés qualitatives sont identifiées à des propriétés physico-chimiques d'états cérébraux.

Pour illustrer la conception « cartésienne/matérialiste » de l'expérience sensorielle consciente, partons de l'expérience de pensée des cerveaux encuvés, et comparons les données auxquelles Younès, un sujet ayant des perceptions véridiques, a accès, et les données auxquelles sa réplique encuvée Younès* – une personne ayant exactement les mêmes états mentaux que Younès, mais en tant que cerveau encuvé — a accès pour sa

11 *Ibid.*, p. 4.

12 L'expression « accointance » (« *acquaintance* ») désigne une relation épistémique spécifique et irréductible que le sujet est censé entretenir avec certaines entités, dont des entités ayant un rôle psychologique comme les *sense-data*. Voir Russell B., *The Problems of Philosophy*, Londres, Oxford University Press, 1912 ; trad. fr. F. Rivenc, *Les problèmes de la philosophie*, Paris, Payot, 1989. L'existence de cette relation est controversée. Pour une revue de la littérature, voir Gertler B., *Self-Knowledge*, Londres, Routledge, 2011.

13 Putnam H., « Sense, Nonsense, and the Senses », p. 464.

part. Supposons que Younès soit en train de regarder un arbre, et que Younès* ait exactement les mêmes sensations visuelles que Younès, mais à l'intérieur de sa cuve. Selon Putnam en 1981, Younès et Younès* possèdent les mêmes données sensibles, puisqu'ils sont conscients, de façon introspective, des mêmes sensations, ayant les mêmes qualités. Autrement dit, les cerveaux encuvés sont, sur le plan sensoriel, des répliques parfaites des personnes réelles : «leurs images, leurs mots, etc., sont qualitativement identiques aux images, aux mots, etc., qui représentent des arbres dans notre monde¹⁴».

Cela implique aussi qu'une expérience sensorielle consciente constitue une donnée probante, susceptible de fonder une croyance perceptive, indépendamment des relations causales qu'elle entretient avec le monde extérieur. En conséquence, le fait que l'expérience soit véridique ou non ne change rien à son statut évidentiel. Putnam est très clair sur ce point en 1994 : «selon la philosophie de la perception cartésienne/matérialiste, de même que selon la conception dualiste cartésienne bien sûr, les entrées perceptives sont les limites externes de nos processus cognitifs ; tout ce qui se trouve au-delà de ces entrées n'est connecté que causalement, mais pas cognitivement, à nos processus mentaux¹⁵». Lorsque Younès forme l'image visuelle d'un arbre en regardant un arbre, sa perception est véridique ; en revanche, ce n'est pas le cas de Younès* lorsqu'il forme la même image sensible dans sa cuve. Pour autant, ces deux sujets possèdent, selon la conception cartésienne/matérialiste, les mêmes données sensibles. On peut résumer cette position par les deux thèses de survenance suivantes. La première est une thèse métaphysique (qui traduit la dimension « matérialiste » de la conception cartésienne/matérialiste), la seconde, une thèse épistémologique (qui traduit sa dimension « cartésienne ») :

(i) *Survenance des expériences sensorielles conscientes relativement aux propriétés physico-chimiques intrinsèques* : deux sujets A et B indiscernables du point de vue de leurs propriétés physico-chimiques intrinsèques sont également indiscernables du point de vue des expériences sensorielles conscientes qu'ils éprouvent : ils ont exactement les mêmes états sensoriels conscients.

(ii) *Survenance des données sensorielles relativement aux expériences sensorielles conscientes* : deux sujets A et B indiscernables du point de vue des expériences sensorielles conscientes ont accès exactement aux mêmes données sensorielles.

Par transitivité, ces deux thèses prises conjointement impliquent (iii) :

14 *Reason, Truth and History*, p. 12-13.

15 «Sense, Nonsense, and the Senses», p. 460.

(iii) *Survenance des données sensorielles relativement aux propriétés physico-chimiques intrinsèques* : deux sujets A et B indiscernables du point de vue de leurs propriétés physico-chimiques intrinsèques ont accès exactement aux mêmes données sensorielles.

Le scepticisme de type cartésien (au sens de Conant) découle de la thèse (iii). Supposons que les seules données dont nous disposons pour affirmer l'existence du monde extérieur soient des données sensorielles. D'après (iii), un cerveau encuvé dispose exactement des mêmes données sensibles qu'un sujet percevant normalement le monde. Or il est difficile de voir comment les données sensorielles dont dispose un cerveau encuvé pourraient fonder des connaissances. La connaissance du monde extérieur semble donc menacée, ce qui suscite l'angoisse cartésienne¹⁶.

Mais c'est l'angoisse kantienne, bien plus que l'angoisse cartésienne, qui est au premier plan dans *Raison, vérité et histoire*. Il ne semble faire aucun doute, pour Putnam en 1981, que les sensations possèdent des propriétés représentationnelles à côté de leurs propriétés qualitatives. Par exemple, l'image sensorielle qui se présente à l'esprit de Younès lorsqu'il regarde un arbre est une image d'arbre. En ce sens, on peut dire qu'elle porte sur un arbre, qu'elle réfère à un arbre. Qu'en est-il cependant de l'image qualitativement identique qui se présente dans l'esprit du cerveau encuvé Younès* ? L'argument modèle-théorique développé au chapitre 2 de l'ouvrage vise à établir que si le réalisme externe est vrai, c'est-à-dire si les objets physiques de taille moyenne comme les arbres existent indépendamment de notre esprit et de nos schèmes conceptuels, cette question ne possède aucune réponse déterminée¹⁷. Nous pouvons certes affirmer que l'image dans l'esprit de Younès réfère à un arbre, mais uniquement relativement à un *contexte* dans lequel certaines relations causales relient un arbre à certains types d'activation du cortex visuel de Younès. C'est la raison pour laquelle Putnam défend, dans le chapitre 3 de *Raison, vérité et histoire*, une forme de relativisme conceptuel selon laquelle ce n'est que relativement à des pratiques d'utilisation des signes dans des communautés linguistiques que ceux-ci ont des référents déterminés¹⁸. Et, de la même façon, d'après la théorie défendue alors, ce n'est que relativement à un contexte dans lequel un concept est utilisé dans des inférences, elles-mêmes liées à des pratiques et à

16 Sur ces points, voir Huemer M., *Skepticism and the Veil of Perception*, Lanham, Rowman & Littlefield Publishers, 2001.

17 Voir Button T., *The Limits of Realism*, Oxford, Oxford University Press, 2013.

18 Voir Putnam H., *Reason, Truth and History*, p. 52. Le relativisme conceptuel implique que les objets auxquels nous faisons référence et les espèces que nous reconnaissons comme naturelles sont relatifs à des cadres conceptuels. En revanche, Putnam rejette le relativisme de la vérité, qu'il considère comme une position incohérente.

des normes d'usage, donc à un « schème conceptuel », que ce concept porte sur un objet extérieur¹⁹.

Cependant, les arguments conduisant tant à l'angoisse cartésienne qu'à l'angoisse kantienne reposent sur le présupposé (iii) ci-dessus, selon lequel les données sensorielles surviennent relativement aux propriétés physico-chimiques intrinsèques des sujets. Le point de départ de l'argument modèle-théorique repose en effet sur un ensemble de questions portant sur la référence des symboles, identifiés de manière syntaxique ou formelle. Étant donné une communauté linguistique dans laquelle les mots « chats » et « cerises » sont utilisés, on peut ainsi s'interroger sur ce qui explique que « chats » réfère aux chats, et « cerises » aux cerises. Pour que cette question ait un sens, il faut que l'on puisse identifier les symboles « chats » et « cerises » indépendamment de leurs référents, uniquement donc par des propriétés formelles. Par exemple, on pourrait proposer d'identifier ces symboles par des propriétés phonologiques, ou par des propriétés orthographiques. Dans la suite de la discussion, je nommerai « véhicule » tout symbole ou état représentationnel identifié indépendamment de ses propriétés sémantiques.

Pour susciter véritablement une angoisse sceptique, il faut que l'argument modèle-théorique soit généralisé du niveau des signes linguistiques à celui des pensées. Il faut, autrement dit, que l'on puisse identifier l'équivalent des véhicules linguistiques dans le domaine de la pensée. Cela n'implique pas un engagement vis-à-vis de l'idée d'un langage mental, mais simplement vis-à-vis de l'idée selon laquelle il existe des occurrences mentales, ayant des propriétés sémantiques, mais identifiables indépendamment de ces propriétés — bref, un engagement vis-à-vis de l'existence de véhicules mentaux de nos états mentaux accessibles par introspection.

Considérons de nouveau l'image mentale d'un arbre. L'occurrence mentale correspondant à cette image, selon la théorie défendue dans *Raison, vérité et histoire*, est un état cérébral possédant des propriétés conscientes directement accessibles à l'introspection. Si l'arbre que je suis en train d'observer a des feuilles rouges, le fait que mon image visuelle instancie la propriété « sensation de rouge » est ainsi directement accessible par introspection, et ne dépend de rien d'autre que de mes propriétés cérébrales. À cet égard, Putnam considère que la distinction entre le *véhicule* d'un symbole linguistique et ses propriétés sémantiques peut être appliquée aux expériences conscientes. Dans le cas que nous discutons, on peut séparer un véhicule possédant à la

19 Cela conduit aussi Putnam à affirmer que « les “objets” n'existent pas indépendamment des schèmes conceptuels » (*Reason, Truth and History*, p. 52).

fois des propriétés conscientes accessibles par introspection et des propriétés cérébrales, et l'objet dénoté dans le monde extérieur par ce véhicule dans un certain contexte — en l'occurrence un certain arbre que je suis en train d'observer. L'introspection me permet de former des croyances justifiées portant sur mon expérience, mais uniquement en tant que véhicule : l'accès direct aux propriétés qualitatives de mes états sensoriels me permet par exemple d'affirmer de façon justifiée que j'ai une sensation de marron, mais pas que cette sensation porte sur une feuille d'arbre. L'introspection ne peut pas me garantir que ma sensation est véridique plutôt qu'illusoire ou hallucinatoire, ni même qu'elle porte sur un arbre plutôt que sur un autre objet. Seule la prise en considération du contexte, en particulier de certaines relations causales, permet d'inférer que la sensation est bien une sensation d'arbre.

1.3. L'image cartésienne/matérialiste et le rôle évidentiel des expériences conscientes

Le présupposé central de l'approche de Putnam en 1981 est donc le principe de survenance selon lequel les données sensorielles surviennent relativement aux propriétés cérébrales intrinsèques des sujets. Ce principe est cohérent avec l'hypothèse d'inspiration cartésienne selon laquelle les expériences conscientes sont directement accessibles par introspection en tant que véhicules de propriétés représentationnelles. Et dans ce contexte, l'image que *Raison, vérité et histoire* donne de la fonction épistémique des expériences sensorielles conscientes peut sembler à la fois cohérente et plausible.

Partons en effet de l'idée internaliste selon laquelle l'expérience sensorielle confère des raisons de croire qui sont *accessibles* au sujet. Selon l'image cartésienne/matérialiste de l'introspection, accéder à ces raisons de croire issues de la perception revient à accéder introspectivement, de façon directe, à des expériences occurrentes, indépendantes des relations causales avec l'environnement. En tant que raison de croire, l'image d'un arbre est exactement identique pour moi et pour ma réplique encuvée, car nous n'avons tous les deux accès introspectivement qu'à des « entrées perceptives », qui ne sont connectées que « causalement », et pas « cognitivement », au monde extérieur.

On trouve dans la littérature épistémologique de nombreuses définitions de l'internalisme compatibles avec cette idée. Ainsi, Laurence Bonjour écrit-il que « selon l'approche la plus généralement acceptée, [...] une théorie de la justification est internaliste si et seulement si elle requiert que tous les facteurs nécessaires à la justification d'une croyance d'une certaine personne soient accessibles cognitivement à

cette personne²⁰ ». Ce concept d'accessibilité cognitive renvoie à l'introspection, comme le précise Robert Audi dans le passage suivant : « certains exemples suggèrent que la justification se fonde entièrement sur ce qui est interne à l'esprit, dans le sens de ce qui est accessible à l'introspection ou à la réflexion pour un sujet — une conception que l'on pourrait appeler internalisme de la justification²¹ ». Si l'on applique maintenant ces idées internalistes à l'expérience sensible et à son statut évidentiel, on parvient à l'hypothèse suivante :

Mentalisme : le support évidentiel d'une croyance perceptive est toujours un *état mental du sujet* dont c'est la croyance, à savoir une expérience consciente accessible par introspection pour ce sujet.

Putnam adhère à cette idée en 1981, et surtout il en donne une interprétation elle-même internaliste au sens non plus épistémologique, mais sémantique et métaphysique. Seule l'expérience conçue comme un état *intrinsèque* du sujet, indépendant métaphysiquement des relations causales qu'il peut entretenir avec son environnement, est en effet accessible par introspection d'après la conception cartésienne/matérialiste de l'esprit. Il existe un contraste essentiel, du point de vue épistémologique, entre une image d'arbre causée dans mon esprit par un mécanisme causal normal de perception visuelle, et une image d'arbre causée par un mécanisme anormal, par exemple dans le cadre d'une hallucination. Néanmoins, la conception cartésienne/matérialiste considère que *nous n'avons pas accès à cette différence par l'introspection*. Lorsque je perçois un arbre véridiquement, j'ai accès exactement à la même expérience sensorielle consciente que lorsque j'ai une expérience visuelle d'un arbre dans le cadre d'une hallucination. Pour le dire autrement : être conscient de ma raison de croire qu'il y a un arbre devant moi lorsque je vois un arbre, ce n'est pas être conscient de la présence d'un arbre devant moi, mais plutôt être conscient, introspectivement, d'une expérience sensorielle consciente occurrente. C'est la présence de cette occurrence dans mon esprit, et pas la présence de l'arbre, qui constitue ma raison de croire.

Cette remarque nous ramène au principe de transparence de l'expérience. En raison de nos intuitions concernant les inversions de spectres, nous devons, selon Putnam en 1981, rejeter ce principe de transparence. L'introspection nous permet donc d'acquérir, de façon directe, des croyances portant sur les aspects purement qualitatifs de nos

20 Bonjour L., « Externalism/Internalism », in Dancy J. & Sosa E. (éd.), *A Companion to Epistemology*, Oxford, Blackwell, 1992, p. 364.

21 Audi R., *Epistemology: A Contemporary Introduction to the Theory of Knowledge*, Londres, Routledge, 1998, p. 233-234.

expériences sensorielles conscientes, et ceci complètement indépendamment des relations causales que ces expériences sensorielles peuvent entretenir avec les objets extérieurs, et donc indépendamment de la présence réelle de ces objets dans le contexte de perception. Ce sont à ces aspects qualitatifs de nos expériences sensorielles, et uniquement à eux, que nous avons accès lorsque nous réfléchissons au support évidentiel que la perception nous procure. L'angoisse cartésienne et l'angoisse kantienne en découlent : puisque l'introspection ne nous dit rien de ce qui distingue une expérience véridique d'une expérience non-véridique, le problème de la fixation des référents des états mentaux apparaît insoluble, ce qui conduit à abandonner le réalisme externe.

2. Du réalisme interne au réalisme naturel

Les conférences Dewey, prononcées par Putnam en mars 1994, marquent l'abandon définitif de la conception cartésienne/matérialiste de l'expérience consciente en faveur du « réalisme naturel », c'est-à-dire d'une forme de réalisme direct. L'hypothèse centrale de ce réalisme naturel porte sur les objets de la perception : « un réaliste naturel en mon sens considère que les objets de la perception (normale, véridique) sont des choses externes, ou, de façon plus générale, des aspects de la réalité extérieure²² ». Pour employer la terminologie de Locke : ce sont les objets réels externes, et non les idées, qui sont présents à mon esprit lorsque je perçois le monde de façon normale.

Selon cette perspective, l'erreur essentielle de la vision cartésienne/matérialiste consiste à introduire des *intermédiaires* entre le sujet percevant et le monde, sous la forme d'un voile de représentations mentales. De quoi sommes-nous en effet directement conscients dans la perception ? Je l'ai souligné plus haut : l'idée centrale de l'approche cartésienne/matérialiste réside en ce que le support évidentiel d'une croyance fondée sur la perception est toujours un *état mental* du sujet dont c'est la croyance. Dans ce cadre, le sujet n'est directement conscient que de ses propres états mentaux, par exemple de ses propres sensations et de leurs qualités sensibles, qui constituent l'unique base évidentielle de ses croyances portant sur le monde extérieur. Un réaliste naturel considère au contraire que la perception est *essentiellement* une conscience directe des objets externes. Dans la perception véridique, ces objets sont immédiatement présents à l'esprit du sujet, sans passer par l'intermédiaire des

22 Putnam H., *The Threefold Cord: Mind, Body, and World*, p. 10.

sensations ou des représentations mentales²³.

L'erreur principale commise en 1981, du point de vue du réalisme naturel de 1994, est certainement de ne pas avoir rejeté suffisamment clairement la confusion, inhérente à la conception cartésienne/matérialiste, entre *l'objet* d'une expérience perceptive véridique, et le *véhicule* de cette expérience²⁴. À cet égard, d'ailleurs, Putnam insiste sur la proximité qui existe entre les philosophes de l'époque moderne, comme Descartes ou Locke, pour lesquels les idées sont définies comme les objets immédiatement présents à l'esprit, et les philosophes représentationalistes matérialistes dont Fodor constitue l'exemple par excellence. Selon ces derniers, nos états mentaux intentionnels ont pour véhicules des représentations mentales qui sont réalisées physiquement dans des états cérébraux. Ces représentations apparaissent à Putnam comme les analogues des impressions sensorielles des classiques :

« Dans les sciences cognitives contemporaines, il est à la mode de postuler l'existence de "représentations" dans l'ordinateur cérébral. Si l'on suppose que l'esprit est un organe, et que l'on identifie l'esprit au cerveau, l'on en vient irrésistiblement à penser (1) que certaines des représentations sont les analogues des "impressions" des théoriciens classiques (l'ordinateur cérébral, c'est-à-dire l'esprit, tire des inférences sur la base au moins de certaines des "représentations", les sorties des processus de perception, exactement comme l'esprit tire des inférences sur la base des impressions selon la théorie classique) et (2) que ces représentations sont reliées aux objets présents dans l'environnement de l'organisme uniquement de façon causale, mais pas de façon cognitive (exactement comme les impressions n'étaient liées que causalement, mais pas de façon cognitive, aux "objets externes")²⁵. »

Ce texte est important pour mon propos, car il permet de saisir à la fois la continuité et l'évolution de la pensée de Putnam sur la question de la référence des pensées de 1981 à 1994. Continuité d'abord. En 1994 comme en 1981, Putnam rejette une certaine forme de représentationalisme, qu'il associe souvent aux sciences cognitives et à Jerry Fodor en

23 Putnam attribue à John McDowell le mérite d'avoir clairement identifié, dans ses conférences John Locke, l'erreur fondamentale de la conception cartésienne/matérialiste de l'esprit. Voir McDowell J., *Mind and World*, Cambridge, Mass., Harvard University Press, 1994 ; trad. fr. C. Alsaleh, *L'esprit et le monde*, Paris, Vrin, 2007.

24 Sur ce point, voir Huemer M., *Skepticism and the Veil of Perception*, p. 80.

25 Putnam H., *The Threefold Cord: Mind, Body, and World*, p. 8.

particulier, selon laquelle des relations causales ou nomologiques seraient susceptibles de déterminer le contenu des pensées. En 1981, l'argument modèle-théorique lui permet d'écarter cette forme de représentationalisme, puisque cet argument implique qu'il n'y a pas une unique façon de déterminer la relation référentielle si celle-ci repose sur la causalité. La conclusion qu'il atteint alors est que le réalisme externe doit être abandonné, et avec lui le matérialisme représentationaliste : ce n'est que relativement à un cadre conceptuel que nos pensées possèdent des conditions de vérité objectives²⁶.

En 1994, Putnam demeure tout aussi sévère avec cette approche, à laquelle il reproche de réduire la référence à « une relation uniquement causale » alors qu'il s'agit d'une relation « cognitive ». Ce qu'il veut dire par là, c'est que lorsque nous percevons le monde, la perception n'est pas un simple mécanisme causal produisant des occurrences de représentations sensorielles dans notre cerveau. C'est aussi un processus cognitif au sens fort, c'est-à-dire un processus de production de connaissances, au travers duquel le sujet devient conscient [*aware*] de la présence des objets dans son environnement et de leurs propriétés perceptibles.

Mais pourquoi Putnam considère-t-il que la forme de représentationalisme qu'il dénonce est incompatible avec le réalisme naturel ? Quel argument donne-t-il pour motiver son jugement selon lequel cette approche considérerait la référence comme une relation « uniquement causale », et pas « cognitive » ? Dans le texte que nous venons de citer, il raisonne en développant une analogie entre d'une part le réalisme indirect en philosophie de la perception, et d'autre part la « mode » consistant à interpréter les mécanismes perceptifs en postulant des inférences inconscientes. Selon cette analogie, une philosophie computationaliste de l'esprit considère que les représentations mentales construites par les mécanismes sensoriels sont aussi les entrées d'inférences conduisant à la fixation de croyances. Supposons, conformément à la conception représentationaliste de l'expérience consciente, que nos sensations conscientes soient des représentations construites par des mécanismes computationnels modulaires. Si ces représentations sont ensuite les entrées des mécanismes de fixation de croyances, ne s'ensuit-il pas, comme le suggère Putnam, que l'approche computationaliste de l'esprit implique au fond une forme de philosophie indirecte de la perception — qu'elle en est réduite à confondre les *véhicules* représentationnels de nos expériences conscientes et

26 Un cadre conceptuel fixe un certain nombre de règles d'usage, et relativement à ces règles les mots et les concepts ont bien un référent déterminé. En revanche, il est impossible, selon l'argument modèle-théorique, de « fixer » la relation référentielle indépendamment de tels cadres conceptuels : « Ou bien l'usage du langage fixe l'interprétation au préalable, ou bien rien ne peut la fixer ». Voir Putnam H., « Models and Reality, I », *The Journal of Symbolic Logic*, 45 (3), p. 482.

les *objets* de ces expériences ?

Ce n'est pas le cas, et je vais soutenir que l'argument esquissé par Putnam repose sur une confusion. Lorsqu'on parle d'inférences inconscientes en sciences cognitives, on désigne par là des mécanismes computationnels qui n'ont pas pour fonction de justifier des croyances, mais de produire certains effets dans un programme. Un représentationaliste peut donc *à la fois* adhérer à l'image computationaliste de l'esprit, c'est-à-dire considérer que les processus perceptifs sont des processus causaux explicables dans un cadre computationnel, *et* adhérer au réalisme direct, c'est-à-dire à la thèse selon laquelle ces mécanismes débouchent sur la perception directe des faits par les agents cognitifs. À cet égard, l'expression «représentationalisme» est malheureusement ambiguë. Elle désigne d'une part un programme de recherche en philosophie de l'esprit, qui repose sur l'hypothèse selon laquelle un certain nombre de processus psychologiques peuvent être expliqués en postulant l'existence de représentations mentales. Mais elle désigne aussi une forme de réalisme indirect en philosophie de la perception, qui considère que les agents cognitifs ne sont jamais directement conscients des objets de leurs représentations sensorielles, mais uniquement des véhicules représentationnels. Afin de bien clarifier ce point, je propose de distinguer ces deux formes de représentationalisme à l'aide d'indices :

Représentationalisme₁ : programme philosophique d'esprit réductionniste, qui prétend que les processus psychologiques pourront à terme être expliqués de façon réductive comme des mécanismes computationnels impliquant de façon essentielle des représentations mentales réalisées dans le cerveau.

Représentationalisme₂ : thèse selon laquelle l'expérience sensorielle consciente fait intervenir de façon essentielle des représentations mentales portant sur des objets extérieurs, et selon laquelle les sujets percevants ne sont pas directement conscients des objets extérieurs même lorsqu'ils les perçoivent de façon véridique, mais uniquement des véhicules représentationnels qui portent sur ces objets.

Lorsqu'on formule les deux versions du représentationalisme de la sorte, il devient clair qu'elles sont indépendantes l'une de l'autre du point de vue logique. En particulier, l'adhésion au représentationalisme₁ n'implique pas de devoir adhérer *aussi* au représentationalisme₂.

Si Putnam suggère le contraire en 1994, c'est sans doute en raison de l'interprétation qu'il donnait lui-même, en 1981, de la conception cartésienne/matérialiste. Selon cette interprétation, comme je l'ai souligné plus haut, le sujet percevant n'a pas accès à

d'autres évidences sensibles qu'à ses sensations, conçues comme des états internes, réalisés dans le cerveau, et ayant des propriétés purement qualitatives. Mais pourquoi un philosophe computationaliste et représentationnaliste devrait-il accepter cette thèse quant au rôle évidentiel des états sensoriels ? Ne pourrait-il pas avoir au fond le beurre et l'argent du beurre, en adhérant au représentationnalisme₁ en tant que programme de recherche naturaliste plausible, tout en rejetant l'idée d'un voile des représentations mentales servant d'interface épistémique entre les agents cognitifs et leur environnement, donc en embrassant le réalisme direct ?

Putnam caractérise le réalisme naturel comme une philosophie de sens commun²⁷, et il fait référence tout autant à William James, à Ludwig Wittgenstein, et à John L. Austin qu'à John McDowell. En première approche donc, le réalisme naturel est défini négativement, par le rejet du représentationnalisme₂, c'est-à-dire par le rejet de l'idée d'une *interface* entre la perception et le monde²⁸. Mais comme je l'ai souligné, le représentationnalisme₁ est compatible avec cette caractérisation négative, puisque rien n'implique logiquement qu'un représentationnaliste₁ adhère à l'idée d'une telle interface.

En fait, de nombreux philosophes représentationnalistes₁ rejettent même explicitement le représentationnalisme₂ dans leurs écrits. Le principal argument que Putnam donne en faveur du réalisme naturel est un argument phénoménologique, qui repose sur la thèse de la transparence : il semble que lorsque nous réfléchissons à nos expériences conscientes de façon introspective, nous tournions notre attention vers le monde extérieur, et pas vers quelque intermédiaire représentationnel que ce soit. Mais la plupart des philosophes représentationnalistes considèrent aujourd'hui que la thèse de la transparence est précisément un argument *en faveur* du représentationnalisme, et pas un argument *contre* cette position²⁹. On considère en général que le représentationnalisme, appliqué au domaine des expériences sensorielles conscientes, implique la thèse de survenance suivante :

Survenance des propriétés phénoménales relativement aux propriétés représentationnelles : nécessairement, deux sujets qui diffèrent quant à leurs propriétés phénoménales diffèrent quant à leurs propriétés représentationnelles.

27 Il reprend l'expression de James « the natural realism of the common man », « le réalisme naturel du sens commun » : cf. par exemple (*The Threefold Cord: Mind, Body, and World*, p. 101).

28 Mon souci, affirme-t-il, est de « proposer une façon de concevoir l'expérience qui n'en fasse pas une interface entre nous et le monde, un écran de cinéma dans nos têtes ou nos esprits dont nous ferions l'expérience directe [...] » (*The Threefold Cord: Mind, Body, and World*, p. 169).

29 Voir par exemple Harman G., « The Intrinsic Quality of Experience », *Philosophical Perspectives*, 4, 1990, p. 31-52 ; Dretske F., *Naturalizing the Mind*, Cambridge, Mass., MIT Press, 1995 ; Tye M., *Consciousness Revisited: Materialism Without Phenomenal Concepts*, Cambridge, Mass., MIT Press, 2009.

L'idée derrière cette thèse est que les propriétés qualitatives des expériences conscientes peuvent être expliquées par les propriétés représentationnelles de ces expériences. Ainsi, le fait d'avoir une sensation de rouge sera expliqué par le fait d'avoir une représentation visuelle d'une surface ayant la couleur rouge. Non seulement l'idée cartésienne/matérialiste d'un intermédiaire interne n'intervient pas dans le représentationalisme présenté ainsi, mais elle paraît incompatible avec l'esprit de ce programme de recherche, qui vise précisément à expliquer les propriétés qualitatives des expériences à l'aide d'une relation entre les agents et le monde, la relation de représentation, plutôt qu'à l'aide d'une relation épistémique reliant le sujet à des objets internes (par exemple, à des *sense-data*). C'est d'ailleurs à Gilbert Harman, un représentationaliste éminent, que l'on doit la formulation la plus souvent citée de l'intuition de la transparence :

«Lorsqu'Héloïse voit un arbre devant elle, les couleurs dont elle fait l'expérience sont toutes senties [*experienced*] comme des traits de l'arbre et de son environnement. Aucune n'est sentie comme un trait intrinsèque à son expérience. Pas plus d'ailleurs qu'elle ne sent quelque trait de quoi que ce soit comme un trait intrinsèque de son expérience. Et cela vaut aussi pour vous [...] Regardez un arbre, et essayez de porter votre attention sur un trait intrinsèque de votre expérience visuelle. Je prédis que vous allez découvrir que les seuls traits vers lesquels votre attention se porte sont des traits de l'arbre³⁰.»

Harman rejette explicitement dans ce texte l'idée d'un intermédiaire qualitatif entre le sujet percevant et le monde : lorsque nous réfléchissons aux propriétés qualitatives de notre expérience, ce n'est pas à l'aide d'une relation épistémique directe avec cette expérience, mais uniquement en prêtant attention à ses objets extérieurs et à leurs propriétés perceptibles – une position apparemment incompatible avec le représentationalisme₂.

Faut-il donc conclure que la critique du représentationalisme par Putnam repose sur un malentendu, sur une confusion entre représentationalisme₁ et représentationalisme₂ ? Je ne le pense pas, et dans une dernière partie de ma discussion je voudrais essayer de montrer qu'on peut trouver chez Putnam au moins l'esquisse d'un argument extrêmement intéressant contre le représentationalisme³¹.

30 Harman G., « The Intrinsic Quality of Experience », p. 31-52.

31 Mon interprétation converge à ce propos avec celle de David Macarthur. Voir Macarthur D., « Putnam's Natural Realism and the Question of the Perceptual Interface », *Philosophical Explorations*, 7(2), p. 167-

3. Représentationalisme, intentionalisme, et réalisme naturel

3.1 Un réalisme direct purement verbal ?

Considérons le passage suivant des conférences Dewey :

« Il y a moins qu'il n'y paraît dans certaines versions du réalisme direct. Cette expression est parfois appliquée à toute position qui rejette la thèse selon laquelle les objets de la perception "véridique" sont des *sense-data*. Un tel usage rend le fait d'être un réaliste direct bien trop facile. Tout ce que l'on a à faire pour être un réaliste direct en ce sens, c'est de dire, par exemple à propos de l'expérience visuelle, "nous ne percevons pas les expériences visuelles, nous les avons". Une simple réforme linguistique, et voilà ! On compte comme un réaliste direct³². »

Ce texte montre que Putnam n'ignore pas que certains philosophes représentationalistes₁ rejettent le représentationalisme₂, qu'ils ne confondent pas le véhicule d'une représentation avec son objet, puisqu'ils considèrent que nous *avons* les expériences conscientes qualitatives lorsque nous percevons (il s'agit des véhicules représentationnels de la perception), mais que nous ne les *percevons* pas (et en ce sens donc elles ne sont en général pas des *objets* de perception). Comme il cite John Searle à la fin de ce passage, il est intéressant de comparer le réalisme direct dont Searle se revendique avec le réalisme naturel de Putnam, afin de parvenir à saisir ce qui distingue selon lui le réalisme direct « bien trop facile » du réalisme naturel. Searle soutient que l'on peut concilier la thèse selon laquelle la perception est intentionnelle, c'est-à-dire selon laquelle une expérience sensorielle consciente, par exemple visuelle, *porte sur* les objets présents dans l'environnement cognitif d'un agent, avec le réalisme direct, et donc avec le rejet du représentationalisme₂³³ :

« Lorsque je vois une voiture (...), j'ai une certaine sorte d'expérience visuelle. Lors de la perception visuelle de la voiture, je ne vois pas l'expérience visuelle, je vois la voiture ; mais lorsque je vois la voiture, j'*ai* une expérience

181.

32 Putnam H. *The Threefold Cord: Mind, Body, and World*, p. 10.

33 À vrai dire, Searle lui-même distingue sa position représentationaliste de ce qu'il nomme le « réalisme représentatif », c'est-à-dire la thèse selon laquelle « nous percevons des représentations du monde réel plutôt que les objets réels eux-mêmes ». Il est donc conscient que l'expression « représentationalisme » est ambiguë. Voir Searle J., *Seeing Things as They Are: a Theory of Perception*, New York, Oxford University Press, 2015, p. 15.

visuelle, et cette expérience visuelle est une expérience visuelle *de* la voiture [...]»³⁴.

L'expérience visuelle d'une voiture possède une propriété intentionnelle puisqu'elle *porte sur* une voiture réelle lorsque l'expérience est véridique. Searle considère de plus que les expériences sensorielles conscientes sont des attitudes propositionnelles³⁵, qui possèdent des conditions de satisfaction : « l'expérience visuelle a des conditions de satisfaction exactement au sens où les croyances et les désirs ont des conditions de satisfaction »³⁶. Enfin, et c'est le point central pour notre propos, Searle considère que deux expériences visuelles A et B peuvent être exactement du même type, même si A est véridique, mais pas B. Supposons que l'expérience visuelle A me permette de voir une voiture réellement présente en face de moi, mais que l'expérience B, une hallucination visuelle, représente de façon non véridique le fait qu'une voiture se trouve devant moi. Dans ce cas, selon Searle, « l'expérience visuelle [B] a le même contenu intentionnel que l'expérience véridique, mais il n'y a pas d'objet de l'expérience visuelle »³⁷. Les expériences visuelles ne sont donc pas des objets de perception, et en ce sens Searle rejette le représentationalisme₂. Néanmoins, il s'agit d'états subjectifs, au sens où une expérience sans objet, par exemple une hallucination visuelle, peut partager exactement le même contenu intentionnel qu'une expérience de perception véridique. Autrement dit, une expérience non-véridique peut avoir le même contenu intentionnel qu'une expérience véridique, même si la première, contrairement à la seconde, ne porte sur aucun objet : il ne faut pas confondre, insiste Searle, l'objet d'une expérience avec son contenu³⁸.

Putnam semble être en accord avec la conception intentionnelle de l'expérience sensorielle défendue par Searle. Il considère en effet, comme John McDowell d'ailleurs³⁹, que les expériences sensorielles possèdent un contenu, et donc une intentionalité⁴⁰. Elles ont une fonction rationnelle, puisqu'elles peuvent justifier des croyances. Le fait de voir

34 Searle J., *Intentionality: an Essay in the Philosophy of Mind*, Cambridge, Cambridge University Press, 1983, p. 37-38 ; voir aussi Searle J., *Seeing Things as They Are*, p. 11.

35 Le contenu d'une expérience visuelle, écrit-il, comme le contenu d'une croyance, « est toujours une proposition complète » (Searle J., *Seeing Things as They Are*, p. 40).

36 *Ibid.*, p. 39.

37 *Ibid.*, p. 19.

38 *Ibid.*, p. 26-27.

39 Voir McDowell J., *Mind and World*. McDowell a clarifié sa position dans un article de 2013, dont le titre est parfaitement explicite : l'expérience perceptive est *à la fois* relationnelle, et pourvue d'un contenu intentionnel. Voir McDowell J., « Perceptual Experience: Both Relational and Contentful », *European Journal of Philosophy*, 21 (1), p. 144-157.

ou d'entendre un chien aboyer constitue ainsi une bonne raison de croire qu'un chien est en train d'aboyer. Or, souligne Putnam⁴¹, l'existence de telles relations rationnelles entre les expériences et les croyances n'est possible qu'à condition que les premières possèdent, comme les secondes, un contenu⁴². Putnam partage donc avec Searle la thèse de l'intentionnalité des expériences sensorielles. Sa motivation pour accepter cette thèse est cependant différente : selon Putnam, les expériences ont des contenus car elles peuvent jouer une fonction rationnelle ; elles sont des *raisons de croire*, et en tant que raisons elles rendent des contenus accessibles à la pensée⁴³. C'est un point très important : si les expériences sont bien des raisons, on peut supposer que des expériences ayant les mêmes contenus doivent aussi jouer le même rôle rationnel, qu'elles doivent constituer les mêmes raisons du point de vue du sujet⁴⁴.

Pour Searle, une expérience visuelle véridique peut avoir le même contenu intentionnel qu'une expérience non-véridique, par exemple une hallucination. Selon ce point de vue, la version encuvée de mon cerveau a, en ce moment, des expériences visuelles qui sont exactement du même type, et qui ont exactement le même contenu intentionnel, que celles que je suis en train d'avoir. Il en découle qu'elles doivent pouvoir jouer la même fonction rationnelle que mes expériences véridiques. Mais alors, ne sommes-nous pas reconduits à l'angoisse cartésienne, puis à l'angoisse kantienne ? Pour sortir de l'angoisse sceptique, il faut pouvoir établir que j'ai accès à des raisons de croire,

40 On en trouve la preuve textuelle dans le passage suivant : «Une expérience non-véridique et une expérience véridique peuvent en effet partager une partie de leur contenu. Par exemple, les deux expériences peuvent “nous dire” (de façon incorrecte, dans le cas non-véridique), qu'il y a une porte jaune devant nous» (Putnam H., *The Threefold Cord: Mind, Body, and World*, p. 153). Dans ce passage, Putnam discute l'intentionnalisme de McDowell, et il semble clair qu'il partage sa position. Il précise bien néanmoins que les expériences non-véridiques et véridiques ne partagent *qu'une partie* de leur contenu, ce qui implique qu'elles n'ont pas littéralement le même contenu. Ainsi l'hallucination d'une porte jaune devant moi représente bien une porte, mais il s'agit d'un contenu général, que l'on pourrait identifier à la description « il y a une porte devant moi ». En revanche si je perçois une porte qui se trouve réellement devant moi, ma perception porte sur *cette porte elle-même*, dans sa singularité, et pas sur une porte quelconque.

41 En référence à une discussion des arguments de John McDowell par Richard Heck. Voir Heck R., «Nonconceptual Content and the “Space of Reasons” », *The Philosophical Review*, 109 (4), p. 483-523.

42 Putnam H., *The Threefold Cord: Mind, Body, and World*, p. 156.

43 Pour une analyse approfondie de cette notion d'accessibilité à la pensée, voir J. Speaks, *The Phenomenal and the Representational*, Oxford, Oxford University Press, 2015, p. 121. Putnam affirme que «le concept de conscience (certainement le concept de conscience pertinent en épistémologie) est le concept d'accessibilité à la pensée (*availability to thought*)» (Putnam H., *The Threefold Cord: Mind, Body, and World*, p. 36).

44 Searle ne mentionne en revanche pas la fonction épistémique des expériences parmi les raisons qui motivent l'intentionnalisme. Selon lui, l'argument le plus important en faveur de cette position est le principe de transparence : c'est parce que nos expériences nous présentent intentionnellement des états de choses que l'expérience nous apparaît transparente. Voir *Seeing Things as They Are*, p. 58.

à des données probantes, auxquelles ma version encuvée n'a pas accès. Mais comment cela serait-il possible si nos expériences ont le même contenu ?

3.2 La question du facteur commun

La question de savoir s'il existe ou non un « facteur commun » entre les expériences sensorielles véridiques et les hallucinations permet de clarifier la façon dont Putnam comprend l'intentionalisme et, je l'espère du moins, de mieux comprendre également les raisons qui le poussent à rejeter la conception représentationnaliste de la conscience.

Supposons que deux expériences visuelles A et B me présentent une porte jaune, la première de façon véridique (il y a donc vraiment une porte jaune devant moi), l'autre de façon non-véridique (il n'y a pas vraiment de porte jaune dans mon environnement, j'ai cependant une hallucination qui, lorsque je réfléchis à l'expérience que je suis en train de vivre, ne peut pas être distinguée subjectivement de l'expérience véridique A). Le désaccord entre Putnam et Searle porte sur la question de savoir s'il y a, ou non, un *facteur commun* à A et B.

Putnam rejette l'hypothèse du facteur commun⁴⁵. La discussion que j'ai menée dans le paragraphe précédent suffit à motiver ce rejet. Dans le bon cas, l'accès du sujet à son expérience visuelle est un accès à une raison de croire qu'il y a une porte jaune devant lui ; dans le mauvais cas en revanche, l'accès à l'expérience visuelle ne constitue pas un accès à une raison de croire cela, ou du moins, il s'agit d'une raison qui possède un rôle différent, car elle ne peut pas complètement justifier la croyance⁴⁶. Puisque les expériences ont des rôles différents, elles ont donc aussi des contenus intentionnels différents, et, si l'on prend véritablement l'intentionalisme au sérieux, il faut en déduire qu'elles sont qualitativement différentes, c'est-à-dire qu'il ne s'agit pas d'expériences du même type. C'est bien la conclusion de Putnam : dans le bon cas (A), je vois une porte jaune, ce qui me permet de prendre conscience de sa présence. Dans le mauvais cas en revanche, il me semble simplement que je vois une porte, mais je ne vois en réalité rien. Cela n'implique pas simplement que mon expérience n'a pas d'objet réel — Searle serait d'accord sur ce point —, mais aussi et surtout, que ma prise de conscience n'est

45 Putnam H., *The Threefold Cord: Mind, Body, and World*, p. 152-153.

46 S'il me *semble* simplement qu'il y a une porte jaune devant moi, j'ai accès à une raison de croire qu'il y a une porte jaune qui n'est pas factive, c'est-à-dire qui ne garantit pas la vérité de la proposition selon laquelle il y a une porte devant moi. Dans le cas A, où l'expérience est véridique, j'ai accès à une raison factive car j'ai conscience de la présence de la porte elle-même devant moi, pas simplement d'une représentation mentale. Sur ces questions, voir Pritchard D., *Epistemological Disjunctivism*, Oxford, Oxford University Press, 2012.

littéralement pas la même dans les deux cas : les deux expériences n'appartiennent pas à la même espèce ou au même type psychologique, et on peut dire en ce sens qu'elles ne partagent pas de facteur commun.

C'est un point qui ne semblera contre-intuitif qu'à ceux ayant succombé, comme Searle, à la conception cartésienne/matérialiste de l'esprit. D'un point de vue subjectif, les expériences A et B sont par hypothèse impossibles à distinguer l'une de l'autre. Searle en déduit qu'elles sont identiques, car il lui semble absurde que deux expériences indiscernables l'une de l'autre par l'introspection puissent appartenir à des types différents⁴⁷. Mais, nous l'avons vu, le rejet de la conception cartésienne/matérialiste est la motivation principale de l'adhésion de Putnam au réalisme naturel. Or le rejet de cette conception va de pair avec une remise en question du statut introspectif privilégié accordé aux expériences conscientes en 1981. Comme je l'ai souligné dans la première partie de cet article, Putnam considérait en 1981 que les expériences conscientes pouvaient constituer, contrairement aux concepts, des objets immédiatement accessibles à l'introspection. Si l'on rejette cette hypothèse d'un accès direct du sujet à ses propres sensations, il devient raisonnable de généraliser aux états sensoriels conscients ce que Putnam avait déjà établi à propos des états conceptuels, à savoir que leur identité dépend de relations avec l'environnement. Lorsque je suis victime d'une hallucination indiscernable d'une expérience visuelle véridique, je ne sais pas que l'expérience que je suis en train de vivre n'a pas été causée de façon normale, et qu'elle ne porte donc pas sur un objet extérieur réel. Je ne peux pas la distinguer, par introspection, d'une expérience véridique. Néanmoins l'hallucination diffère bien de l'expérience véridique, y compris sur le plan purement qualitatif. En effet, seule la seconde me permet d'accéder aux apparences visuelles d'une certaine chose. Dans le cas de l'hallucination, rien ne m'apparaît, même s'il me *semble subjectivement* que quelque chose m'apparaît.

Le fait que rien ne m'apparaisse dans les mauvais cas n'implique pas que, dans ces situations, rien ne soit représenté du tout. Rejeter l'idée d'un contenu intentionnel commun aux bons et aux mauvais cas ne veut pas dire pas que les expériences illusoire ou hallucinatoires n'ont pas de contenu intentionnel du tout, ni même qu'il n'y a pas *une partie* du contenu représenté qui est partagé avec les expériences véridiques. Reprenons l'exemple, longuement analysé par Putnam, de l'expérience visuelle d'une porte jaune. Supposons que cette expérience soit hallucinatoire. Son contenu n'est alors pas singulier : une porte *quelconque* est représentée comme occupant une certaine position

47 Voir sa discussion de la conception disjonctive au chapitre 6 de Searle J., *Seeing Things as They Are*.

dans l'espace relativement au sujet, un contenu que l'on peut identifier à une proposition existentielle (« il y a une porte jaune X à l'endroit E »). Comparons ce contenu à celui de la perception, véridique cette fois, d'une porte jaune réelle. Ce contenu est singulier, car l'expérience n'est correcte qu'à condition que la porte particulière qui est perçue soit bien telle qu'elle est représentée (jaune, occupant une certaine position) : « il y a une porte jaune, *telle que* $X=P$, à l'endroit E ». Mais on voit que, comme le souligne Putnam, la proposition générale associée à l'expérience non-véridique d'une part, et la proposition singulière associée à l'expérience véridique de l'autre, ont bien un contenu en commun, puisque la seconde implique logiquement la première.

Tout cela n'est-il pas incompatible avec ce que j'ai dit plus haut du rejet par Putnam de l'hypothèse d'un facteur commun ? Un représentationnaliste typique, comme Fred Dretske, considère (i) que le facteur commun est un contenu représenté, qui est *le même* dans l'expérience véridique et dans l'expérience non-véridique ; et (ii) que les propriétés phénoménales, ou qualitatives, de l'expérience, dépendent *intégralement* de ce contenu. Mais Putnam rejette à la fois (i) et (ii). Il rejette (i), puisque, comme nous venons de le voir, il ne pense pas que les contenus représentés soient identiques dans les deux cas. Et il rejette également (ii), puisqu'il considère qu'il ne suffit pas qu'un certain contenu soit représenté par un état perceptif pour qu'on puisse parler d'apparences subjectives. Dans le cas d'une hallucination représentant une porte, on ne peut dire en aucun sens qu'une porte particulière apparaisse au sujet, même si son expérience lui représente bien une porte quelque part dans l'espace.

3.3 Un désaccord portant sur les apparences

Le fond du désaccord entre Searle et Putnam porte donc sur le concept d'apparence. Selon Searle une apparence est essentiellement quelque chose de subjectif et d'accessible immédiatement par introspection. Selon Putnam après 1994, les apparences possèdent au contraire une dimension objective, puisqu'il ne peut pas y avoir d'apparence sans que *quelque chose* apparaisse pour un sujet. Par ailleurs, l'introspection n'est pas infaillible, puisqu'il peut me sembler subjectivement que je perçois une apparence alors que je ne perçois en fait rien. Putnam adhère donc à une interprétation disjonctive de l'expérience consciente, comme il l'explique dans ce passage :

« [...] William James, John Austin, John McDowell [...] ont proposé une conception disjonctive de la perception. Selon cette conception, il n'y a

littéralement rien “en commun” entre [les cas A et B], c’est-à-dire aucune “qualité” identique. Dans le [cas A] quelqu’un voit une porte jaune. Dans le [cas B] il semble à quelqu’un qu’il voit une porte jaune. Pour la conception disjonctive, “avoir l’expérience visuelle d’une porte jaune” est une expression disjonctive (ou ambiguë). Cela peut vouloir dire ou bien voir une porte réelle ayant une qualité colorée réelle (le jaune), ou bien qu’il semble que cela soit le cas⁴⁸.»

Indépendamment du désaccord avec Searle, ce point est important car il permet de mieux saisir ce que Putnam reproche à la conception représentationnaliste de l’expérience. Répétons-le : Putnam est intentionaliste. Il considère donc, comme les philosophes représentationnalistes, qu’une expérience consciente, y compris non-véridique, possède un contenu propositionnel qui représente (ou « présente », dans sa terminologie) un certain état de choses. Le sujet appréhende donc un certain contenu propositionnel dans la situation B comme dans la situation A. Le point central qui transparait dans cette citation, c’est néanmoins que l’appréhension d’un tel contenu propositionnel n’est pas suffisante, en elle-même, pour rendre le sujet conscient d’une apparence. Le philosophe représentationnaliste a raison de vouloir expliquer le fait qu’il semble au sujet d’une hallucination qu’il est conscient de quelque chose en faisant appel à l’idée d’un contenu propositionnel représenté par l’expérience ; mais il aurait tort d’en déduire qu’il y a un objet commun dont le sujet serait conscient dans le cas A comme dans le cas B. Si c’était le cas, l’intentionalisme se réduirait au fond à une version renouvelée de la conception cartésienne/matérialiste, identifiant les apparences qualitatives à des contenus représentés, indépendamment de la présence réelle des objets dans le contexte de perception. Or, si des apparences qualitatives peuvent nous apparaître en l’absence de tout objet réel présent, nous sommes reconduits à l’angoisse cartésienne. Il est essentiel, pour dissiper cette angoisse, de reconnaître qu’il y a bien une spécificité épistémique de l’expérience véridique, y compris sur le plan subjectif, qui consiste en ce qu’elle nous rend conscients de faits réels.

Même s’il exagère parfois son désaccord avec les philosophes représentationnalistes, Putnam a bien mis le doigt sur une limite intéressante de cette approche. Le représentationnalisme est un programme de recherche réductionniste qui a pour ambition d’expliquer les caractéristiques qualitatives des expériences conscientes dans le vocabulaire des représentations mentales. Or, les mêmes représentations mentales

48 *The Threefold Cord: Mind, Body, and World*, p. 153.

sont impliquées dans le cas A et le cas B. Il semble donc bien qu'un représentationnaliste devrait accorder à Putnam : (i) ou bien que les propriétés qualitatives des expériences conscientes ne sont finalement pas explicables de façon réductive uniquement en termes de représentations mentales ; (ii) ou bien que nous avons accès exactement aux mêmes apparences dans les cas A et B. À mon sens, Putnam a raison de suggérer que la seconde option est une voie risquée : s'il me *semble* simplement que je vois une porte jaune alors que ce n'est pas le cas, je n'ai pas accès à l'apparence jaune de la porte, car cette apparence doit être instanciée par un objet pour exister dans un certain contexte⁴⁹.

Certes, un représentationnaliste peut envisager diverses réponses : soutenir que l'on peut être conscient y compris de propriétés sensibles qui ne sont instanciées par aucun objet⁵⁰, ou que l'on peut être conscient de propriétés sensibles instanciées par des objets possibles, mais en tout cas qui ne sont pas présents dans la situation où l'expérience a lieu⁵¹. Aucune de ces réponses n'est complètement satisfaisante d'un point de vue métaphysique⁵². C'est néanmoins surtout sur le plan épistémologique qu'elles se heurtent à une objection. La motivation principale du réalisme direct est de dissiper l'angoisse cartésienne⁵³. Mais comment le fait de prendre conscience des apparences pourrait-il servir de justification à nos croyances perceptives, si cette prise de conscience ne concernait pas seulement des objets réellement présents dans le contexte ? Dretske, par exemple, soutient que nous sommes directement conscients des qualités sensibles représentées par nos expériences visuelles, que ces qualités soient instanciées par des objets réellement perçus ou non. Si c'est le cas cependant, cette prise de conscience directe n'est pas *toujours* une garantie rationnelle de la vérité des croyances qu'elle motive, et elle ne permet pas de dissiper l'angoisse cartésienne⁵⁴.

49 Sur ce point, voir Pitt D., « The Paraphenomenal Hypothesis », *Analysis*, 77 (4), p. 735-741. Quoiqu'il ne mentionne pas Putnam, Pitt développe dans cet article un argument très proche de ceux que nous avons discutés.

50 Voir Dretske F., *Naturalizing the Mind*.

51 Voir Lycan B., *Consciousness and Experience*, Cambridge, Mass., MIT Press, 1996.

52 Voir Pautz A., « Intentionalism and Perceptual Presence », *Philosophical Perspectives*, 21 (1), p. 495-541.

53 Voir, de nouveau, Huemer M., *Skepticism and the Veil of Perception*.

54 Comme l'a fait remarquer une relectrice ou un relecteur anonyme, on pourrait craindre, ici, que le raisonnement soit circulaire : Putnam semble présupposer que l'angoisse cartésienne *peut* être dissipée, et c'est ce qui lui permet d'écarter l'option représentationnaliste. Mais d'une part, l'angoisse cartésienne est une angoisse typiquement philosophique, qui ne caractérise pas notre expérience perceptive naïve du monde, et il semble raisonnable d'imaginer qu'on puisse bien la dissiper. Et d'autre part, le représentationnalisme de Dretske reste solidaire de l'image matérialiste/cartésienne, pour autant qu'il présuppose une relation fondamentale d'acointance avec les contenus représentés par les expériences conscientes. Lors d'une hallucination visuelle représentant la couleur rouge, le sujet serait ainsi accointé avec la couleur rouge, simplement du fait que son système visuel représente cette couleur. Mais le moins

Au fond, selon Putnam, l'erreur des représentationalistes est de croire que la conscience au sens qualitatif, c'est-à-dire la prise de conscience des qualités sensibles qu'ont les choses, peut s'effectuer en l'absence d'une interaction cognitive entre le sujet et les choses en question⁵⁵. Et ici encore Putnam reste fidèle à l'une des intuitions centrales de *Raison, vérité et histoire*: on ne peut comprendre l'acquisition de connaissance en considérant l'esprit et les objets externes comme des entités complètement séparées et isolées les unes des autres. En ce sens au moins, en 1994 comme en 1981, «l'esprit et le monde constituent conjointement l'esprit et le monde»⁵⁶.

Conclusion

Entre 1981 et 1994, Putnam change radicalement de position à propos de la nature des expériences sensorielles conscientes. Alors qu'il rejette l'intuition de la transparence de l'expérience en 1981, et donne un rôle central à l'argument du spectre inversé, il met en 1994 le principe de transparence au cœur de sa philosophie de la perception, ce qui le conduit à affirmer que les apparences conscientes dépendent des objets extérieurs, et pas seulement de l'état du cerveau à un moment donné. J'ai cependant essayé de montrer qu'il existait également un fil rouge reliant le réalisme interne et le réalisme naturel : le rôle épistémique des expériences conscientes. Si l'on présuppose la conception cartésienne/matérialiste de l'expérience consciente, comme le faisait Putnam en 1981, la seule façon de dissiper les angoisses cartésienne et kantienne est d'abandonner le réalisme métaphysique. Le prix à payer, qui est d'adopter une conception épistémique de la vérité, a cependant paru trop lourd à Putnam entre 1981 et 1994, et il remet alors en question le présupposé cartésien/matérialiste lui-même. Putnam demeure également constant dans sa critique des conceptions causales de la référence mentale, en particulier dans son rejet du représentationalisme. Sur le plan épistémologique, les représentations mentales ne doivent pas être considérées comme des interfaces entre le sujet et le monde : sur ce point, comme je l'ai souligné, la plupart

qu'on puisse dire c'est que cette idée pose problème dans le cadre, par ailleurs naturaliste, du représentationalisme. Pour une référence à un texte où Dretske affirme que nous sommes directement conscients de propriétés non-instanciées dans les hallucinations, voir Dretske F., «The Mind's Awareness of Itself», in F. Dretske, *Perception, Knowledge, and Belief*, Cambridge, Cambridge University Press, 2000, p. 163.

55 Putnam parle plus précisément de «transactions» entre le sujet et l'environnement. D'après lui, ces «transactions» sont essentielles à la perception («la perception a un caractère de transaction [*perception is transactional*]»)(*The Threefold Cord: Mind, Body, and World*, p. 159) et elles ne peuvent pas être expliquées réductivement dans le vocabulaire des sciences naturelles (p. 44).

56 *Reason, Truth and History*, p. xi.

des philosophes représentationalistes contemporains le rejoignent en fait. Néanmoins la critique de Putnam va bien plus loin que le simple rejet de l'idée, largement discréditée aujourd'hui, d'un voile de représentations mentales. Il soutient que l'hypothèse réductionniste selon laquelle les représentations mentales pourraient entièrement expliquer les propriétés qualitatives de nos expériences conscientes est incompatible avec le réalisme naturel. Un réductionniste sera en effet tenté d'affirmer que nous sommes subjectivement conscients des mêmes qualités sensibles lors d'une expérience non-véridique – par exemple une hallucination visuelle – et lors d'une hallucination visuelle, puisque les mêmes représentations sensorielles sont mobilisées dans ces deux expériences. Mais Putnam insiste, de façon convaincante, sur le fait que les sujets ne peuvent être conscients de qualités sensibles que si celles-ci sont instanciées par des objets avec lesquels ils interagissent sensoriellement dans leur environnement. Cela implique que des processus représentationnels peuvent, dans l'esprit de deux sujets, être identiques quant aux véhicules impliqués – au sens où les mécanismes intervenant dans ces processus sont les mêmes – tout en débouchant sur des états conscients qualitativement distincts. Cela implique aussi, et c'est l'une des conséquences les plus originales de sa nouvelle approche, que ces différences qualitatives ne sont pas toujours, et par nature, accessibles à l'introspection⁵⁷.

57 Je remercie vivement deux relectrices ou relecteurs anonymes, ainsi qu'Henri Wagner, dont les remarques m'ont permis d'améliorer la rédaction de cet article.