
Braver les portes de la perception Enquête hétérophénoménologique sur la nature de l'illusion de la conscience

Auteur : Zwaenepoel, Pierre

Promoteur(s) : Seron, Denis

Faculté : Faculté de Philosophie et Lettres

Diplôme : Master en philosophie, à finalité didactique

Année académique : 2021-2022

URI/URL : <http://hdl.handle.net/2268.2/15240>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Université de Liège
Faculté de Philosophie & Lettres

Braver les portes de la perception

Enquête hétérophénoménologique sur la nature
de l'illusion de la conscience

Travail de fin d'études présenté par Pierre ZWAENEPOEL
en vue de l'obtention du diplôme de Master en philosophie,
à finalité didactique

Sous la supervision de : M. Denis SERON
Lecteurs : M. Arnaud DEWALQUE & M. Bruno LECLERCQ

Année scolaire 2021-2022

Remerciements

Je souhaite remercier avant tout Denis Seron pour son enthousiasme et ses remarques encourageantes tout au long de la rédaction de ce travail, ainsi que pour son suivi efficace malgré mon rythme de travail erratique. Merci d'avoir soutenu ce travail en me laissant une totale liberté intellectuelle dans la rédaction de celui-ci.

Merci également aux lecteurs de ce mémoire, Arnaud Dewalque et Bruno Leclercq. J'espère que vous pourrez me pardonner sa longueur.

Merci à François Kammerer d'avoir pris le temps de répondre à mes mails, éclairant ma compréhension de concepts clés de ce travail et infléchissant sa direction de manière décisive. J'aurais souhaité échanger plus, ce qui aurait certainement bénéficié à ce travail, si j'en avais eu la possibilité.

Merci à mes parents d'avoir continué à me supporter envers et contre tout. Merci Maman, pour la relecture.

Merci à mes amis d'avoir rendu douces ces années de master. Merci aux membres du groupe de travail qui m'ont accompagné durant celles-ci, en particulier Kitty, Nico, Philippe et Thomas.

Merci à Héloïse pour son soutien émotionnel, et surtout pour tout le reste.

Table des matières

Introduction	1
Partie I – Le projet illusionniste de Daniel Dennett	
Chapitre 1 – Examen des concepts de Daniel Dennett	7
1.1 Rappel utile concernant les <i>qualia</i>	7
<i>Définition des qualia</i>	7
<i>Caractéristiques des qualia</i>	8
1.2 L'hétérophénoménologie	15
1.3 Le modèle des versions multiples	21
1.4 Conclusion	28
Chapitre 2 – L'illusionnisme	31
2.1 L'illusionnisme selon Keith Frankish	31
2.2 L'illusionnisme selon Daniel Dennett	36
Partie II – La nature de l'illusion de la conscience	
Chapitre 3 – Illusion riche ou légère ?	43
3.1 Introduction au problème de la richesse de l'illusion de la conscience	43
3.2 Défense de l'hypothèse de l'illusion légère	46
<i>Commentaire du syllogisme de Kammerer (1)</i>	49
<i>Commentaire du syllogisme de Kammerer (2)</i>	55
<i>La force d'une illusion dépend de notre compréhension du phénomène illusoire</i>	57
<i>La force d'une illusion dépend de notre proximité avec le phénomène illusoire</i>	62
3.3 Conclusion	65
Chapitre 4 – Vers une théorie doxastique de l'illusion de la conscience	69
4.1 Biais cognitif ou illusion perceptuelle ?	69
4.2 L'erreur de Kammerer	74
4.3 Conclusion	83
Partie III – Hétérophénoménologie de l'expérience psychédélique	
Chapitre 5 – Observer la conscience altérée	87
5.1 États modifiés de conscience	87
<i>La conscience</i>	87
<i>Les états modifiés de conscience</i>	88
<i>Caractérisation des états modifiés de conscience</i>	89
5.2 Expérience psychédélique	101
<i>Aperçu historique des recherches psychédéliques</i>	101
<i>Phénoménologie de l'expérience psychédélique</i>	104
Chapitre 6 – Repenser la conscience grâce aux psychédéliques	111
6.1 Enjeux et méthode des recherches psychédéliques	111
6.2 Renvoyer la balle aux sciences cognitives	113
6.3 Perspective illusionniste sur le travail de Letheby	124
<i>Narrative Self, Minimal Self et Predictive Self-Binding</i>	124
<i>Centre de gravité narrative ou centre de gravité représentationnelle ?</i>	131
<i>Qu'en est-il de la conscience phénoménale ?</i>	134
<i>Argument en faveur de la légèreté de l'illusion</i>	135
Conclusion	143
Annexe	147
Références	149

Introduction

Pendant longtemps, une intuition millénaire érigée en théorie positive par Descartes, est restée l'un des dogmes les plus tenaces de l'histoire de la philosophie. Le dualisme a en effet su fonder sa légende sur le terrain fertile de l'imagination humaine, s'imposant pendant un temps comme un très sérieux paradigme philosophique. La recherche de la nature de la substance mentale, ou encore l'exploration des secrets de son interaction avec la matière, ont dès lors animé philosophes et autres esprits scientifiques, cristallisant le débat au sujet de la conscience.

Or, depuis le tournant du XX^{ème} siècle, l'étude scientifique de l'esprit humain a ouvert la porte à une nouvelle alliance entre conscience et matière, réconciliées dans une approche naturaliste du monde. Des travaux de William James aux recherches récentes dans les domaines des sciences cognitives et des neurosciences, l'écart n'a jamais cessé de se réduire entre ces deux mondes, et avec lui l'horizon des arguments dualistes.

Il semblerait néanmoins que certaines intuitions résistent envers et contre tout à l'influence grandissante du paradigme physicaliste. Les philosophes sont coupables de cultiver celles-ci : ainsi, certains affirment que la conscience échapperait toujours à la science, celle-ci étant incapable de résoudre le Problème Difficile posé par son aspect le plus fondamental, « l'*expérience* » (Chalmers, 1995, pp. 2-3). Celle-ci (ainsi que toutes ses autres manifestations : « conscience phénoménale », « aspect subjectif », « *qualia* » ou encore « effet que cela fait ») s'est ainsi imposée comme un sujet de discussion de première importance en philosophie de l'esprit, propageant l'idée selon laquelle la conscience pourrait échapper à la naturalisation.

Réticents qu'ils étaient à accepter que leur objet d'étude puisse être mieux étudié par la science que par eux-mêmes, ces philosophes manquèrent peut-être l'évolution importante de notre compréhension de la conscience permise par les techniques et approches scientifiques modernes. À moins qu'ils aient préféré affubler leur objet d'étude de propriétés toujours plus extraordinaires afin de l'éloigner toujours plus du giron de la science. Toujours est-il que la croyance à l'irréductibilité de la conscience reste prégnante aujourd'hui, en dépit de notre compréhension toujours plus complète et toujours plus fine du fonctionnement du cerveau et des mécanismes produisant la conscience.

2 • Introduction

En ce qui concerne l'étude scientifique de la conscience, mon intérêt s'est d'abord porté sur l'étude d'un type particulier d'état modifié de conscience, l'expérience mystique, facilitée par l'utilisation de techniques modernes d'imagerie cérébrale. Rapidement, j'ai découvert non seulement qu'il s'agissait d'un champ de recherche plutôt vivace, mais également qu'il était étroitement relié à un autre champ, peut-être plus atypique encore : celui des « recherches psychédéliques ». Le passage de l'un à l'autre est simple : les substances hallucinogènes dites « psychédéliques » ont la capacité étonnante de provoquer des états mystiques avec une grande régularité, comme en atteste leur utilisation, largement documentée, dans les pratiques chamaniques du monde entier. Il est donc possible grâce à celles-ci de provoquer des expériences mystiques chez des sujets afin d'observer ce phénomène « en laboratoire ». En définitive, le sujet de l'expérience psychédélique s'étant révélé plus vaste et plus porteur que prévu, il a pris une place centrale dans la conception de ce travail, allant finalement jusqu'à éclipser les autres états modifiés de conscience.

Brisons tout de suite un *a priori* : associés d'abord à la contre-culture hippie combattue par l'*establishment* de la société occidentale, et aux noms d'intellectuels et scientifiques controversés tels qu'Aldous Huxley ou Timothy Leary (Nutt, 2021, p. 40), les psychédéliques ont rapidement fait les frais d'efforts politiques aveugles aux questions scientifiques et de campagnes de désinformation alimentées par les médias, acquérant au fil des années une aura négative. Pourtant, leur intérêt scientifique et thérapeutique avait été compris très tôt, et l'interdiction de leur usage médical a certainement retardé d'une quarantaine d'années des découvertes qui auraient pu sauver un nombre incalculable de vies. Heureusement, aujourd'hui, les recherches psychédéliques ont repris et se portent relativement bien, comme en témoigne l'augmentation drastique d'études cliniques, portant surtout sur les effets de la psilocybine dans le traitement de troubles psychiatriques, dans les trois dernières années¹.

Il m'est important, en raison de leur histoire houleuse, de souligner dans ce travail la légitimité des psychédéliques en tant qu'objets de science. Comme nous le verrons, ceux-ci constituent un outil précieux pour étudier la conscience. Les recherches psychédéliques se tiennent actuellement à l'avant-garde des sciences de l'esprit et forment un domaine très sérieux d'étude et de questionnements, y compris philosophiques – bien que les approches philosophiques des travaux modernes sur le sujet soient encore rares. Néanmoins, l'étude

¹ Pour un aperçu plus large de ce sujet, voir Tullis, P., How ecstasy and psilocybin are shaking up psychiatry, 2021. En ligne : <https://www.nature.com/articles/d41586-021-00187-9>.

de l'expérience psychédélique n'est elle-même pas épargnée par les *a priori* philosophiques susmentionnés : tout comme il serait impossible, selon certains, d'atteindre une explication objective de la conscience, il serait également impossible d'expliquer l'expérience psychédélique par la voie scientifique, la complexité du vécu associé à celle-ci ne pouvant se laisser réduire à une histoire de neurotransmetteurs et de courants électriques.

Gratter la couche de préjugés et d'intuitions problématiques entourant l'étude de la conscience comme celle des psychédéliques est donc l'objectif qui a guidé l'élaboration de ce travail ; et qui mieux que Daniel Dennett, défenseur de la contre-intuitivité (Dennett, 2012, pp. 209-210) ayant déclaré une guerre de « métaphores » (Dennett, 1993, p. 565) aux chantres des *qualia*, pour servir celui-ci ? C'est donc vers cet auteur que je me suis tourné pour obtenir une base théorique physicaliste au sujet de la conscience qui puisse questionner, et être questionnée, par une approche philosophique de l'expérience psychédélique. La première partie de ce travail portera quasi-exclusivement sur ses travaux et fournira le cadre théorique de ce travail.

Outre l'exposition des concepts de Dennett, cette première partie proposera également un chapitre plus général portant sur l'illusionnisme, qui sera l'occasion de définir le problème de l'illusion. L'illusionnisme concernant la conscience phénoménale, défendu par Dennett, constituera en effet un présupposé important de ce travail. Or, cette notion étant souvent mal comprise, il me semble nécessaire non seulement d'en définir la conception dennettienne, mais également de la replacer dans un contexte théorique plus large.

Une fois la conception dennettienne de l'illusion de la conscience abordée, je proposerai dans une deuxième partie une confrontation de cette conception avec les travaux de Kammerer, dont les arguments en faveur de l'illusionnisme diffèrent considérablement de ceux de Dennett, et permettent à mon sens d'interroger judicieusement ceux-ci.

Troisième et dernière partie de ce travail, l'analyse de cas de l'expérience psychédélique viendra préciser et renforcer la théorie de la première partie, et les prises de position de la deuxième. Après en avoir défini l'objet d'étude, j'exposerai dans cette partie les travaux de Christopher Letheby et questionnerai ceux-ci à l'aide des concepts élaborés plus tôt afin de tenter d'en tirer un argument en faveur de l'illusionnisme de Dennett.

Partie I
LE PROJET ILLUSIONNISTE DE DANIEL
DENNETT

Chapitre 1 Examen des concepts de Daniel Dennett

1.1 Rappel utile concernant les *qualia*

Avant de m'intéresser aux concepts développés par Daniel Dennett, je pense qu'un bref rappel au sujet des *qualia* s'impose afin de mieux cerner les problèmes philosophiques qu'adressent les travaux de celui-ci. Exposer les caractéristiques habituellement attribuées aux *qualia* (telles qu'elles sont relevées par Kammerer (Kammerer, 2018b)) nous permettra en l'occurrence de cibler certains aspects problématiques de cette notion qui seront abordés dans la suite de ce travail.

Définition des *qualia*

Dans son article dédié aux *qualia* dans la *Stanford Encyclopedia of Philosophy*, Michael Tye affirme que « les philosophes se servent souvent du terme « *qualia* » (singulier « *quale* ») pour désigner les aspects phénoménaux, accessibles par introspection, de nos vies mentales », en ajoutant que « sur la base de ce sens large et classique du terme, il est difficile de nier l'existence de *qualia*. »²

Toujours selon Tye, « le statut des *qualia* est vivement débattu en philosophie, surtout parce qu'il est central à une véritable compréhension de la nature de la conscience. » Il est en tout cas certain, *a contrario*, qu'en l'absence de limites clairement établies au concept de conscience, les débats à propos des *qualia* survivront, et les « qualophiles »³ pourront à leur gré repousser toujours plus loin la particularité des *qualia* hors du territoire des sciences de l'esprit (jusqu'à, éventuellement, affirmer leur caractère épiphénoménal (Dennett, 1993, p. 497)).

D'habitude, la façon la plus simple et directe de définir ce que sont les *qualia* consiste à en énumérer quelques exemples facilement identifiables, c'est d'ailleurs par-là que Michael Tye débute l'article susmentionné :

² Traduction de Claude Pichevin, dans (Dennett, 2012, p. 131).

³ Le terme provient de Dennett (Dennett, 1993, p. 478), qui l'utilise pour désigner les défenseurs de la notion de *qualia*. Ce terme possède une connotation légère voire humoristique, mais je prie le lecteur de ne pas y trouver de caractère péjoratif ; je l'utiliserai simplement par facilité pour désigner de façon générale les philosophes adoptant une position réaliste au sujet des *qualia*.

8 • Examen des concepts de Daniel Dennett

« Il existe une très grande diversité de sentiments et d'expériences. Ainsi, par exemple, je passe mes doigts sur du papier de verre, je sens l'odeur d'un putois, je ressens une douleur aiguë au doigt, il me semble voir du rouge vif, j'entre dans une grande colère. Dans chacun de ces cas, je suis porteur d'un état mental marqué par un caractère subjectif très distinctif. Cela me fait quelque chose d'éprouver chacun d'entre eux, il y a une phénoménologie qui lui est propre.

Cette méthode a l'avantage de dessiner un concept qui semble parfaitement intuitif pour une majorité de personnes, y compris pour les non-philosophes (d'expérience, expliquer à un profane ce qu'est un *qualia* me paraît plus aisé en se servant d'exemples saillants qu'en procédant pas à pas par une exposition de concepts tels que « aspect phénoménal » ou « caractère subjectif »). Néanmoins, dès que l'on passe de ces exemples si intuitifs à une tentative de définition exhaustive, l'exercice devient compliqué, la faute sans doute à la pluralité de conceptions pas toujours compatibles avancées par les philosophes, mais aussi, peut-être, à un certain caractère vague inhérent à tout concept de *qualia*.

L'apparente évidence du concept ne doit pas nous empêcher d'en percevoir les limites. Dennett remarque que les discussions au sujet des *qualia* souffrent de confusions découlant de la présence dans les définitions classiques d'un « essaim de présuppositions non examinées et d'élaborations définies de manière circulaire » (Dennett, 2012, p. 132). Survoler la littérature au sujet des *qualia* pourra nous en convaincre, en plus de nous offrir un rappel de l'état actuel des discussions sur ce sujet.

Caractéristiques des qualia

Kammerer décrit, dans un article de l'*Encyclopédie philosophique* dédié aux *qualia*, sept caractéristiques qui ont été attribuées aux *qualia* par les philosophes de l'esprit. Celles-ci ne sont pas indissociables et font l'objet de nombreuses discussions, néanmoins elles renseignent sur l'état actuel de la recherche en philosophie de l'esprit et permettent de délimiter un tant soit peu le concept de *qualia* tel qu'il est conçu aujourd'hui. Dans ce chapitre, j'aimerais aborder ces caractéristiques, ce qui me permettra d'explorer les problèmes posés par les *qualia*, préalablement à l'exposition d'une théorie qui entend les critiquer.

Ineffabilité

L'un des problèmes qui surgit à propos des *qualia* est la difficulté que nous rencontrons lorsque nous tentons d'exprimer par le langage « la nature des *qualia* » (Kammerer, 2018b). Il serait apparemment impossible de se faire comprendre d'un

interlocuteur à qui l'ont décrit un *quale* particulier, si celui-ci n'a pas lui-même déjà fait l'expérience de ce *quale*. Pour savoir « quel effet cela fait » de, par exemple, voir du rouge, il faut Soi-même avoir vu du rouge ; la compréhension de cet effet ne peut être partagée par le langage.

Il semblerait en effet qu'il soit particulièrement difficile de définir certaines sensations de façon satisfaisante (typiquement les sensations de couleur ou de douleur) sans recourir à l'ostension et donc sans recourir au point de vue subjectif (en se référant à notre propre expérience du phénomène) ou à la présomption de la première personne du pluriel (en imaginant que notre interlocuteur pourra lui-même se référer à une expérience du phénomène qui soit sensiblement la même que la nôtre). Pourtant, il existe bien une très large littérature scientifique sur ces sujets, et des phénomènes tels que les couleurs – pour conserver cet exemple paradigmatique – sont plutôt bien compris, mais quelque chose échappe à leur description scientifique : une propriété qui définirait « *la manière dont les choses nous apparaissent à l'œil* (à l'oreille, au nez, etc.) » (Dennett, 1993, p. 462), la façon dont les choses nous apparaissent subjectivement.

En plus de la nature même des couleurs que nous voyons, le « vécu » qui est associé à notre expérience consciente est chargé d'une affectivité particulière qui semble également échapper à la définition. Mais s'il semble que les *qualia* aient une façon particulière de se donner à nous qui soit chargée émotionnellement, il n'est cependant pas impossible d'envisager une description fonctionnaliste de l'expérience consciente qui soit très exacte tout en constituant une réelle explication. D'après Dennett, la « voie "réductionniste" qui consiste à *identifier* "la manière dont cela se présente à moi" avec la somme totale de toutes les dispositions réactives idiosyncrasiques inhérentes à mon système nerveux qui résultent de ma confrontation avec un certain schéma de stimulation » (Dennett, 1993, p. 479) fournit une description adéquate de l'expérience qui, à défaut de la rendre directement accessible (l'expérience restera toujours subjective⁴), nous permet néanmoins de la comprendre (Dennett, 1993, p. 481). Analyser toutes les dimensions fonctionnelles particulières qui composent une expérience particulière devrait ainsi permettre de comprendre celle-ci, à défaut de la vivre nous-même, puisque cela requerrait que nous produisions en nous-même

⁴ Le caractère subjectif des *qualia* sera abordé un peu plus loin.

10 • Examen des concepts de Daniel Dennett

l'ensemble de ces dispositions, ce qui représente un travail énorme⁵ dont le produit ne nous apprendrait finalement rien de plus.

Bien entendu, l'affirmation précitée repose sur une définition réductionniste du *quale* comme étant un « complexe idiosyncrasique de dispositions », et elle ne satisfera pas ceux qui affirmeraient que tout ce travail laisserait toujours un « résidu ineffable » à expliquer. Cette conception particulière sera défendue dans le suite de ce travail. Pour l'instant, contentons-nous de souligner ceci : il n'est pas certain qu'une description très détaillée d'un *qualia* soit incapable d'en épuiser toute la profondeur, et par conséquent il n'est pas certain que la conscience soit en quelque sens que ce soit ineffable.

Caractère directement introspectable ou accointance

Il semblerait que nous ne puissions *pas* nous tromper au sujet de nos expériences, dans le sens où ce qu'il me *semble* se produire en moi et ce qui se produit effectivement en moi semblent être indissociable. Cela rejoint l'affirmation de Searle selon laquelle, en ce qui concerne la conscience, « l'existence d'une apparence est la réalité » (Kammerer, 2019a, p. 338). L'un des exemples souvent apportés à cette affirmation est le cas de la douleur : il serait impossible de se tromper sur le fait que nous ayons une expérience de douleur (Aydede & Murat, 2019). L'une des explications possibles de ce fait serait que nous sommes avec nos *qualia* dans une relation de connaissance directe, ou accointance. Ainsi, l'acte d'introspection par lequel nous découvrons nos expériences serait la condition nécessaire et suffisante d'une connaissance, « infaillible » selon certains (Chalmers, 2003, p. 241), de celles-ci (Kammerer, 2018b).

Mais quel genre de connaissance nous offre cet accès infaillible à nos expériences ? Même si son objet se donne sans filtre, une science qui reposerait sur l'introspection souffrirait de biais qui l'empêcheraient d'affirmer produire une connaissance réelle. En effet, d'une part l'étude de phénomènes exclusivement subjectifs nous laissera toujours avec une science « d'une seule personne », qui ne dit pas grand-chose de la conscience en général, d'autre part l'introspection ne nous donne accès qu'à une représentation schématique du fonctionnement réel de la conscience, qui ne nous apprend rien sur ce qui se passe dans les coulisses de celle-ci (le fonctionnement de la machinerie cérébrale ne

⁵ Énorme mais pas impossible. Ainsi, Dennett propose l'exemple d'un « musicologue qui aurait soigneusement évité tout contact avec la musique postérieure à 1725 et qui se serait profondément familiarisé avec la musique traditionnelle de cette période » comme une première « approximation » de ce qui devrait être mis en place pour recréer aujourd'hui l'expérience particulière qu'a pu vivre un « pratiquant luthérien de Leipzig » entendant une cantate de Bach en 1725 (Dennett, 1993, p. 479).

nous est pas accessible de cette manière). Nous avons un accès direct aux métaphores produites par la conscience, oui, mais pas aux causes de ces métaphores. Pourtant, c'est bien ces causes qui nous permettraient de comprendre le fonctionnement de la conscience.

Dans les deux cas, une « autophénoménologie » qui s'inquiéterait de ce que les *qualia* livrent à notre introspection ne semble pas pouvoir mener l'enquête bien loin. Par conséquent, si l'expérience subjective possède un caractère directement introspectable, celui-ci ne permet d'accéder à aucune information déterminante inaccessible autrement. Ainsi notre relation si privilégiée soit-elle à nos expériences n'apporte rien qui fasse avancer les recherches sur la conscience. Elle n'apporte aucun élément probant en faveur des *qualia*, et ne remet pas en cause la nécessité d'une approche à la troisième personne de la conscience.

Caractère subjectif

La notion de « sujet » est problématique puisqu'elle tend à substantialiser ce qui, pour les illusionnistes, n'est qu'une illusion produite par le fonctionnement en concert d'entités parfaitement idiotes et aveugles. Remplacer le concept de « sujet » par des termes plus neutres (par exemple « agent » ou « observateur ») tendent à éviter les malentendus, mais parler de la « subjectivité » de ces agents ou observateurs risquerait de nous faire retomber dans les pièges que l'illusionnisme s'efforce d'éviter. Il est donc important de préciser quel usage pourrait être fait du terme dans un contexte illusionniste.

Dans la mesure où un *quale* est un ensemble de dispositions particulier survenant dans un système nerveux individuel, il y a au moins un sens à dire que les *qualia* sont toujours « pour » une conscience. Néanmoins, reconnaître le fait qu'il existe un point de vue à la première personne ne doit pas nous pousser à en inférer qu'il existerait une substance mentale (Dennett, 1993, p. 45). Ainsi, reconnaître le caractère subjectif de l'expérience consciente dans le sens donné plus haut, ne nous engage pas à lui reconnaître une quelconque nature intrinsèque ni à la transformer en faire-valoir d'un Sujet qui échapperait à une description en termes physicalistes.

Ce sens particulier de subjectivité ne nous demande pas non plus d'endosser la possibilité d'une relation épistémique d'accointance avec l'expérience, puisqu'il n'empêche pas la possibilité que nous puissions faire erreur au sujet de notre expérience. Enfin, le caractère subjectif (dans le sens donné) de l'expérience consciente ne devrait pas

12 • Examen des concepts de Daniel Dennett

nous amener à conclure que son contenu pourrait échapper ou devrait être mis entre parenthèse par une science à la troisième personne.

Caractère non-représentationnel

D'après Kammerer, « il existe plusieurs manières de comprendre la signification exacte de cette affirmation, mais il est possible de l'introduire dans la version suivante : dire que les *qualia* sont *représentationnels*, c'est dire que tout état mental doté de *qualia* est une *représentation* d'un certain type qui vise un certain objet (on parle d'un objet « intentionnel »), qui est dotée d'un certain contenu, et que ce contenu détermine entièrement les *qualia* de l'état en question » (Kammerer, 2018b). La thèse des *qualia* représentationnels semble intuitive dans certains cas, mais d'autres pourraient la mettre en doute, ainsi pour certains il pourrait exister certaines expériences qui ne s'accompagnent d'aucune représentation : l'expérience d'une « pure douleur » est régulièrement citée en exemple (Kammerer, 2018b).

Mais en quoi exactement devrait consister une telle expérience ? Vers quoi cette douleur serait-elle tournée ? Comment pourrions-nous seulement nous identifier à cette douleur si elle n'était dirigée vers aucune membre, organe ou aucun état interne de notre corps ? Y a-t-il seulement un sens à parler de douleur qui ne soit pas une douleur *quelque part* ou une douleur *d'un certain type* ? L'approfondissement de ce problème dépasserait de loin le cadre de ce travail, je m'arrêterai donc simplement à cette remarque, que le concept de « pure expérience » semble très mal défini et donc peu convainquant en l'état.

Caractère intrinsèque

Les *qualia* sont parfois considérés comme des propriétés « intrinsèques » de l'expérience, c'est-à-dire « des propriétés [d'une] perception qui ne dépendent pas constitutivement d'autre chose que de [cette] perception elle-même, et notamment qui ne dépendent pas constitutivement des relations que cette perception entretient avec l'environnement du sujet qui l'a (bien sûr, ils peuvent en dépendre causalement) » (Kammerer, 2018b).

Le problème avec cette définition est qu'elle isole en quelques sortes causalement les *qualia* au sein du processus perceptuel. Or Dennett note que « si l'on définit les *qualia* comme tout ce qui n'est pas effets du flux aval des expériences (réactions à des couleurs particulières, comptes rendus verbaux, effets sur la mémoire, etc.), ni causes génératrices dans le flux amont des expériences (activité immanente à une région corticale ou à une

autre) », c'est-à-dire s'ils sont isolés « logiquement de toutes les propriétés dispositionnelles », les *qualia* sont alors rendus imperméables à toute explication en termes fonctionnalistes (Dennett, 2012, pp. 243-244). Seulement, postuler la résistance des *qualia* à une interprétation fonctionnaliste pose d'autres problèmes, abordés dans la section ci-dessous.

Caractère non-fonctionnel

Kammerer décrit une propriété fonctionnelle comme « une propriété telle que sa possession par un objet consiste entièrement dans le fait que cet objet joue (ou peut jouer) un certain rôle fonctionnel (ou un certain rôle causal) au sein d'un système, sans que soit fait référence à la composition de cet objet (c'est-à-dire au matériau qui *réalise* la propriété fonctionnelle). » Ainsi, un objet est un cœur en vertu du fait qu'il remplit le rôle fonctionnel attendu d'un cœur (le rôle d'une pompe d'un certain type) (Kammerer, 2018b), qu'il soit fait de tissus organiques, de plastique ou de n'importe quel autre matériau.

Or il serait, selon certains, « impossible d'identifier le fait, pour un état mental, d'avoir un certain *quale*, au fait, pour cet état, de jouer (ou de pouvoir jouer) un certain rôle fonctionnel » (Kammerer, 2018b). En d'autres termes, les *qualia* seraient non-fonctionnels. Notons que cette position est plus forte que celle qui consisterait à identifier les *qualia* à des « complications » qui viendraient mettre en doute une trop grande simplification du fonctionnement des états mentaux (Dennett, 2012, p. 249).

En effet, Dennett remarque qu'il existe un malentendu au sujet de l'approche fonctionnaliste dans les sciences de l'esprit. Il remarque que « le fonctionnalisme marque dans la pratique un penchant en faveur du minimalisme, de l'idée que le moins compte plus qu'on n'aurait pu le croire » (Dennett, 2012, p. 245), découlant du souci de généralisation au cœur de la méthode scientifique. Les sciences cognitives ont pu endosser un « minimalisme fonctionnaliste » « particulièrement radical » (Dennett, 2012, p. 246) à une époque, donnant lieu à certaines hypothèses comme celle de l'IA forte (l'hypothèse selon laquelle n'importe quel système reproduisant les fonctions du cerveau humain serait aussi conscient que ce dernier). Mais, cette approche ayant été remise en question par les découvertes des neurosciences qui ont souligné le rôle joué dans le fonctionnement du cerveau humain par certaines propriétés qui avaient été écartées comme non-fonctionnelles, l'idée selon laquelle le fonctionnalisme aurait été mis en échec a pu gagner certains (Dennett, 2012, p. 248). Or c'est précisément l'inverse, affirme Dennett : la découverte des rôles fonctionnels d'autres propriétés du cerveau ne fait que renforcer l'hypothèse

14 • Examen des concepts de Daniel Dennett

fonctionnaliste, puisqu'elle rajoute de nouvelles propriétés fonctionnelles à prendre en compte dans l'explication d'un phénomène, en même temps qu'elle nous éloigne de la simplification initiale voulue par l'exigence de généralité (Dennett, 2012, pp. 248-249). Le « minimalisme trop optimiste » (Dennett, 2012, p. 249) laisse place à une compréhension plus fine, mais tout aussi respectueuse du fonctionnalisme.

Cette dynamique de redéfinition des caractères fonctionnels d'un phénomène au gré des découvertes ne saurait donc constituer une critique du fonctionnalisme et n'implique aucunement qu'un phénomène qui échappe pour l'instant à une délimitation parfaite en termes fonctionnalistes serait *de facto* non-fonctionnel. Que faire, donc, des arguments supposés constituer une défense de l'absence de rôle fonctionnel des *qualia* ?

Ceux-ci se basent notamment sur la possibilité logique d'une inversion des *qualia* (c'est l'expérience de pensée du « spectre inversé », tirée de l'*Essai sur l'entendement humain* de John Locke) ou d'une absence de *qualia* (dans l'expérience de pensée de la « Nation Chinoise » de Ned Block) (Kammerer, 2018b). Il s'agit de réfuter l'hypothèse fonctionnaliste en proposant des cas où un système identique en tout point (fonctionnellement parlant) à un autre système se voit pourvu différemment du point de vue subjectif, soit qu'il possède des *qualia* différents, soit qu'il ne puisse manifestement pas en posséder.

Mais ces expériences de pensées peuvent-elles réellement remplir cet objectif ? En principe, « si les fonctionnalistes avaient raison, tout état qui joue exactement le même rôle fonctionnel que mon état possesseur d'un *quale* [...] devrait posséder nécessairement exactement le même *quale* ; or la possibilité du spectre inversé contredit précisément ce fait » (Kammerer, 2018b). Cette possibilité « de principe » (Dennett, 1993, p. 485) établie au sujet de l'inversion des *qualia* entraînerait donc qu'il soit possible que la façon dont les choses nous apparaissent change radicalement sans que nos dispositions correspondantes se trouvent changées en conséquence (si nos *qualia* changeaient en même temps que nos dispositions, nous n'aurions aucune confirmation que l'un et l'autre sont dissociables) (Dennett, 1993, pp. 484-485). Par exemple, la victime d'une inversion soudaine des *qualia* se mettrait à voir le ciel jaune, ou l'herbe rouge, sans que se produisent de modifications notamment comportementales dans ses réactions à ces nouvelles expériences. Elle se mettrait alors à agir de façon indiscernable d'un être humain normal, tout en possédant un spectre des couleurs inversé, et même l'hétérophénoménologie ne pourrait rien déceler de

son étrange affection, étant donné l'impossibilité de la comparaison intersubjective des *qualia* (Dennett, 1993, p. 483).

Le phénomène d'inversion des *qualia* semble intuitif, mais s'efforcer de l'imaginer dans toute sa complexité nous montre qu'il s'agit non pas d'un argument mais d'une « pompe à intuition », « une histoire qui [nous] caresse dans le sens du poil en [nous] incitant à mettre en avant [nos] intuitions viscérales sans leur apporter de bonnes raisons » (Dennett, 1993, p. 491). L'efficacité de l'expérience du spectre inversé repose entièrement sur l'intuition que l'inversion des *qualia* entraînerait une différence évidente du point de vue de l'expérience, mais en réalité cette évidence naît d'une conception naïve de la conscience (Dennett, 1993, p. 491).

Affirmer que l'inversion des *qualia* serait « quelque chose de plus que l'inversion de toutes les dispositions réactives que l'on peut avoir » (Dennett, 1993, p. 492) de sorte qu'il serait possible de maintenir cette inversion indépendamment des dispositions est une erreur provenant du mythe du Théâtre Cartésien (dont il sera question dans le chapitre suivant). Si l'on abandonne au contraire l'idée selon laquelle les *qualia* seraient indépendants d'une façon ou d'une autre du complexe de dispositions auxquels ils sont liés, les *qualia* inversés deviennent une impossibilité de fait et ne constituent donc pas un argument en faveur d'un caractère intrinsèque ou non-fonctionnel des *qualia*.

Nature non-physique

Certains auteurs affirment la nature non-physique des *qualia*, et que ceux-ci échapperaient donc aux théories physicalistes de l'esprit. Kammerer note que le sens précis à donner à ce type d'affirmation est encore débattu, puisque la définition d'une « propriété physique » n'est pas évidente (Kammerer, 2018b). Néanmoins, l'ensemble des théories soutenant cette vision des *qualia* ont en commun de ne pas croire au physicalisme concernant l'esprit. Par conséquent, étant donné le présupposé physicaliste de l'approche que je défends, il ne me semble pas intéressant de me pencher sur de telles théories ni d'apporter des éléments qui les contredisent.

1.2 L'hétérophénoménologie

À présent que nous avons mieux cerné quelles sont les caractéristiques problématiques attribuées aux *qualia*, nous pouvons passer à l'examen des concepts de Daniel Dennett, à commencer par celui d'hétérophénoménologie. Ce concept est sans doute

16 • Examen des concepts de Daniel Dennett

la pierre angulaire de la philosophie de Daniel Dennett, or celui-ci semble avoir eu des difficultés à s'imposer parmi les philosophes de l'esprit contemporain, notamment en raison de la similarité de la méthodologie qu'il propose avec celle des sciences. Que propose cette méthode de plus que les sciences de l'esprit qui permette à Dennett d'affirmer prétentieusement qu'elle permet de rendre justice à toute la complexité de l'expérience consciente ? Ne s'agit-il pas que d'une tentative réductionniste d'imposer à la philosophie un agnosticisme qui sied bien aux neuroscientifiques, mais ne fait que masquer le vrai *hard problem* ? Prenons le temps de revenir en détail sur ce que dit Dennett de l'hétérophénoménologie afin de mettre en évidence la nouveauté de cette méthode et les raisons pour lesquelles elle est adéquate pour mener l'enquête qui justifie ce travail.

« L'hétérophénoménologie (« phénoménologie d'un autre que Soi ») est une approche « à la troisième personne de la conscience humaine » qui se présente comme (1)⁶ une extension « simple et conservatrice » de la méthode scientifique qui vise à rendre compte (2) « aussi sérieusement que possible » des phénomènes de la conscience – (3) tous les phénomènes, y compris les plus ineffables. » (Dennett, 2003, p. 19)

(1) Un premier élément apparaît d'emblée dans cette première définition : l'hétérophénoménologie se positionne du côté des sciences de l'esprit, dont elle constitue, si pas une extension, en tout cas une formalisation. Elle n'est, nous dit Dennett, « rien d'autre que la méthode qui a été utilisée par les psychophysiciens, psychologues cognitifs, neuropsychologues cliniques, et à peu près tout qui a jamais prétendu étudier la conscience humaine de façon sérieuse et scientifique » (Dennett, 2003, p. 22). Qui donc questionnerait une telle approche ? Des ennemis de la science ? Pas du tout, en fait. Ceux qui critiquent habituellement l'hétérophénoménologie sont ceux contre qui elle est dirigée originellement : ceux pour qui le *hard problem* est insolvable, les « mystériciens » et plus généralement ceux qui pensent qu'une science à la troisième personne est incapable de s'attaquer à la conscience (Dennett, 2003, p. 23) dans son entièreté – en particulier dans sa dimension subjective.

L'hétérophénoménologie est « simple », dans le sens où elle ne propose que ce qui relève déjà d'une réalité de terrain, une méthodologie qui existe par ailleurs, et « conservatrice », dans le sens où elle n'implique aucun changement de paradigme, ni physique ni métaphysique (Dennett, 1993, pp. 19, 30).

⁶Ma numérotation.

(2) Dennett présente l'hétérophénoménologie comme « la voie solide qui permet d'envisager le point de vue à la *première* personne aussi sérieusement qu'il puisse l'être » (Dennett, 2003, p. 19). Le problème soulevé par cette formulation est que le point de vue à la première personne porte certaines limitations qui l'empêchent de présenter en tant que tel des données pour une science. Ce que Dennett appelle « l'autophénoménologie », dans laquelle « le sujet et l'expérimentateur sont une seule et même personne » peut certes produire des idées, mais celles-ci ne peuvent dépasser le statut d'idées tant qu'elles n'ont pas été confrontées à des « expériences hétérophénoménologiques » (Dennett, 2003, p. 23).

En effet, si certains philosophes ont pu considérer que la validation expérimentale, par exemple par une enquête, de leurs intuitions relevait toujours d'une méthode « à la première personne », Dennett fait remarquer que ce n'est pas le cas puisqu'une fois confronté aux conditions expérimentales, ce qui relevait de l'introspection dans un premier temps entre en interaction avec un environnement de recherche qui rend nécessaire le changement de point de vue préconisé par l'hétérophénoménologie. Pour que les comptes rendus introspectifs des sujets de la situation expérimentale visant à confirmer l'idée du philosophe deviennent des « données » qui permettront d'étayer ce qui est désormais posé comme une hypothèse, ils doivent être *interprétés* (Dennett, 2003, p. 20).

L'hétérophénoménologie met entre parenthèses la question de savoir de quoi exactement sont faits les objets « autophénoménologiques » des sujets pour se concentrer sur l'étude des objets intentionnels du sujet envisagés en tant que « fictions de théoriciens » peuplant leur monde hétérophénoménologique. Les « objets hétérophénoménologiques », nous dit Dennett, sont « des *abstracta*, pas des *concreta* » (Dennett, 1993, p. 127), c'est-à-dire des « fictions utiles » dont l'hétérophénoménologue se sert pour mener à bien son entreprise et dont il n'est pas utile à ce stade de savoir à quelle réalité exacte – s'il en existe une – ceux-ci renvoient. C'est pourquoi on peut dire de l'hétérophénoménologie qu'elle est une méthode « neutre » (Dennett, 1993, p. 130), ou « agnostique » (Dennett, 2003, p. 24), contrairement à la phénoménologie traditionnelle qui donne tous les atours d'une réalité à certains de ses postulats théoriques non démontrés (comme les zombies philosophiques, ou les *qualia*). Ceci nous amène à la troisième et dernière partie de notre définition de l'hétérophénoménologie.

(3) Les données de l'hétérophénoménologie sont les objets qui peuplent le monde hétérophénoménologique des sujets ; ce monde se définit lui-même comme un « catalogue de ce que le sujet croit être vrai à propos de son expérience consciente » (Dennett, 2003,

18 • Examen des concepts de Daniel Dennett

p. 20). C'est donc à des croyances que nous avons affaire. Ces croyances sont les « fictions de théoriciens » citées ci-dessus, c'est-à-dire les abstractions qui servent (aussi bien au sujet lui-même qu'à l'observateur hétérophénoménologue) à décrire différents états rencontrés par le sujet, dont la complexité échappe à son entendement quand elle n'échappe pas carrément à la science (Dennett, 2003, pp. 20, 21).

Certains critiques ont fait remarquer qu'il était vain de tenter d'expliquer la conscience à l'aide de ces abstractions : « les expériences conscientes, et non seulement nos jugements verbaux à propos d'elles, sont les données primaires dont une théorie doit rendre compte »⁷. Dennett contredit ce point. Pour lui, en l'état actuel de notre maîtrise du sujet de la conscience, il nous est impossible de réaliser l'entreprise de catalogage à laquelle est commise l'hétérophénoménologie en agissant directement au niveau des expériences conscientes elles-mêmes. Étant donné que nous ne disposons pas d'une théorie et d'une technique qui nous permettent d'interpréter sans filtre les expériences conscientes, il est plus sûr de s'en remettre aux croyances au sujet de ces expériences.

En effet, l'ensemble des données que nous pouvons relever à propos des croyances au sujet d'expériences conscientes (a) peut être plus petit ou plus grand que l'ensemble des données au sujet d'expériences conscientes (b), selon Dennett. D'un côté on peut poser la possibilité qu'il existe des expériences conscientes sur lesquelles ne portent aucune croyance ($b > a$), auquel cas (a) et (b) auront la même extension puisque les expériences conscientes qui nous échapperaient seraient inaccessibles aussi bien au sujet qu'à l'observateur et ne feraient donc pas partie de notre ensemble de données. D'un autre côté, on peut poser la possibilité qu'il existe des croyances au sujet d'expériences conscientes qui n'existent pas en réalité ($a > b$), et dans ce cas (b) englobera plus de données que (a). Ainsi, s'en remettre aux croyances comme données est un moyen d'éviter « de s'en remettre à des données fausses » tout en épuisant celles qui nous sont accessibles (Dennett, 2003, p. 21).

L'argument cité ci-dessus permet de soutenir l'utilité de l'hétérophénoménologie et sa pertinence pour enquêter sur les expériences dites ineffables. Selon elle, en effet, la croyance qui porte sur une expérience consciente est déjà une donnée qu'il nous est possible de traiter pour en apprendre plus sur le phénomène. Dans le cas d'expériences dites ineffables, cette remarque prend tout son sens. En effet, d'un point de vue

⁷ Levine, J., *Out of the closet : a qualophile confronts qualophobia*, 1994. Cité dans (Dennett, 2003, p. 21).

phénoménologique, il semble logiquement impossible de décrire un phénomène qui se présente comme ineffable (il est de son essence même de ne pas pouvoir l'être, au mieux ne pourra-t-il être décrit que négativement). Le déplacement proposé par l'hétérophénoménologie permet de se saisir efficacement de ce qui semble être la seule donnée incontestable au sujet d'une expérience qualifiée d'ineffable : la croyance du sujet au sujet de l'ineffabilité de son expérience.

Jusque-là, ce changement de point de vue ne semble pas apporter de véritable révolution dans notre façon d'aborder notre enquête. En effet, le phénoménologue peut tout aussi bien s'interroger également sur ses propres croyances au sujet de ses expériences. Cependant, celui-ci postulant un accès privilégié à ses expériences (il possède en effet un accès plus direct à ses propres expériences qu'aux expériences d'autrui, n'en déplaise à l'hétérophénoménologue qui place les deux sur le même plan), il atteindra la même conclusion que précédemment ; pour lui « l'existence d'une apparence est la réalité » quand il s'agit de la conscience⁸. L'hétérophénoménologue pourra amener l'enquête plus avant en abandonnant ce postulat de la transparence des expériences conscientes et en posant en principe la possibilité des croyances fausses au sujet des expériences conscientes. Il s'agit d'une importante possibilité puisqu'elle justifie l'accent mis par l'hétérophénoménologie sur les croyances des sujets au sujet de leurs expériences (et la formulation de ces croyances, que ce soit en termes comportementaux ou sous forme de comptes rendus introspectifs, verbalisés ou non).

Poser une possibilité ne revient pas à affirmer une existence, et ce projet aurait pu se révéler totalement vain, il apparaît cependant que l'hétérophénoménologie est capable d'offrir des résultats intéressants (comprenez des descriptions adéquates vis-à-vis de nos connaissances scientifiques de phénomènes autrement obscures voir inexplicables) en suivant son hypothèse. Le travail de Dennett en offre quelques exemples, ce travail entend modestement en proposer un autre.

Je reviendrai plus précisément sur la possibilité de l'erreur au sujet des expériences conscientes et de leur contenu dans la deuxième partie de mon travail, pour l'instant je souhaite conclure cette présentation de l'hétérophénoménologie en reprenant succinctement les éléments qui la caractérisent :

⁸ Searle, J., *The mystery of consciousness*, 1997. Cité dans (Kammerer, 2019a, p. 338).

20 • Examen des concepts de Daniel Dennett

- 1 L'hétérophénoménologie se constitue en réponse au « mystérianisme » (Dennett, 2003, p. 23), elle se positionne du côté des sciences de l'esprit... (Dennett, 1993, p. 92)
- 2 ...dont elle est une extension de la méthode qui vise à rendre justice aux phénomènes de la conscience humaine, notamment les phénomènes privés (réputés ineffables pour certains). Elle est la seule façon d'appréhender le point de vue à la première personne comme une donnée pour la science (Dennett, 2003, pp. 19, 22).
- 3 Ses données sont les comptes rendus d'expériences conscientes vécues par des êtres humains (Dennett, 2003, pp. 21-22).
- 4 Elle se propose d'envisager ceux-ci, comme des abstractions, des « fictions utiles » (Dennett, 1993, pp. 124-127) dont il ne s'agit de déterminer ni l'essence métaphysique ni le substrat matériel (s'il existe) (Dennett & Akins, Multiple drafts model, 2008 ; Dennett, 1993, pp. 102, 109 ; 2003, p. 20), mais l'adéquation avec les modèles fonctionnels, cognitifs, comportementaux ou neurologiques de la conscience que nous fournissent par ailleurs les sciences, auxquelles elle entend passer le flambeau pour rendre compte des phénomènes d'un point de vue physicaliste.
- 5 À ce titre, l'hétérophénoménologie reconnaît la nécessité pour les comptes rendus d'expériences conscientes d'être interprétés (Dennett, 1993, p. 106), c'est-à-dire ne pas être envisagés comme des réalités, mais comme des objets intentionnels qu'il est possible de décrire en adoptant un point de vue « à la troisième personne » (Dennett, 1993, p. 103).
- 6 En ce sens, l'hétérophénoménologie est une discipline « neutre », qui ne se prononce pas sur la réalité du phénomène en lui-même (Dennett, 1993, p. 102), mais n'exclut pas la possibilité d'erreurs ou d'approximations dans l'accès à celui-ci.

L'hétérophénoménologie constitue, en somme,

« le chemin neutre qui conduit de la science physique objective et du privilège qu'elle accorde au point de vue à la troisième personne, à une description phénoménologique qui peut (en principe) rendre justice aux expériences les plus privées et les ineffables sans jamais abandonner les scrupules méthodologiques de la science. » (Dennett, 1993, p. 98)

Pour cette raison, et puisque proposer des éléments de description phénoménologiques d'expériences privées d'un point de vue physicaliste est l'un des objectifs de ce travail, celui-ci prend acte des propositions de cette méthode et entend mobiliser ses concepts au besoin.

1.3 Le modèle des versions multiples

Le modèle des versions multiples⁹ est un modèle de la conscience proposé par Dennett destiné à remplacer les modèles influencés par le « matérialisme cartésien ». Ce modèle ne propose pas de description ou d'explication de la conscience au niveau physique, et ne correspond pas à un modèle cognitif exhaustif. Il s'agit d'un modèle général, abstrait, répondant à certains critères (destinés à assurer qu'ils répondent bien à l'impératif de lutter contre les erreurs du matérialisme cartésien) pouvant être remplis par une multitude de modèles cognitifs spécifiques (Dennett & Akins, 2008).

Je décrirai plus loin dans ce chapitre les critères en question, mais avant cela je me propose de revenir brièvement sur la notion de « matérialisme cartésien » afin de donner à voir la nouveauté du MVM par rapports aux modèles classiques de la conscience.

Dennett appelle « théâtre cartésien » le lieu hypothétique où se concentrent toutes les données du cerveau pour faire l'objet d'une présentation à l'agent (qu'il s'agisse d'un Soi au sens contemporain du terme ou d'une *res cogitans*) (Dennett & Akins, 2008). Descartes plaçait ce point dans la glande pinéale, minuscule organe du cerveau dont on ignore toujours la fonction, mais qui y occupe une place symbolique : le centre (Dennett, 1993, p. 139). Cette hypothèse est tombée en désuétude, mais d'autres ont suivi qui reposent sur le même postulat : que l'on place le point focal du cerveau dans la glande pinéale, le cortex cingulaire antérieur ou n'importe où ailleurs, on présuppose qu'il existe un centre où « tout se réunit » dans le cerveau (Dennett, 1993, p. 141).

Le « matérialisme cartésien » rassemble les thèses qui, bien qu'ayant abandonné le dualisme, ont conservé du modèle cartésien l'idée qu'il existe une « ligne d'arrivée cruciale ou une frontière quelque part dans le cerveau, qui marque un endroit où l'ordre d'arrivée est identique à l'ordre de " présentation " dans l'expérience¹⁰ parce que *ce qui se passe là*

⁹Abrégé dans ce travail en « MVM ».

¹⁰Une coquille s'est glissée à cet endroit dans la traduction de Pascal Engel, que j'ai ici corrigée sur base de l'édition Back Books de 1992.

22 • Examen des concepts de Daniel Dennett

est ce dont vous êtes conscient. » (Dennett, 1993, p. 141). Cette ligne d'arrivée serait le point de départ de ce que l'on a pu appeler le « flux de conscience » (Dennett, 1993, p. 149) et permettrait de déterminer aussi bien spatialement (il s'agit de l'endroit où un contenu passe du non-conscient à la conscience) que temporellement (le point central permet de distinguer clairement ce qui est *pré-expérientiel* de ce qui est *post-expérientiel* (Dennett, 1993, p. 143)) les contours de la conscience.

Personne n'est encore parvenu à localiser une telle ligne d'arrivée. Pire, présupposer qu'elle existe sans s'atteler d'abord à la fixer définitivement crée la confusion lors de l'interprétation de résultats d'expériences en sciences de l'esprit. Dennett donne plusieurs exemples dans *Consciousness Explained* de faits dont l'explication est rendue difficile par le postulat : ainsi du phénomène *phi* pour les couleurs de Kolars et von Grünau (Dennett, 1993, p. 150), et des expériences de Benjamin Libet (Dennett, 1993, pp. 195, 206).

Pour établir la séquence des événements qui constituent ces phénomènes, il nous faut, si l'on embrasse le matérialisme cartésien, déterminer l'ordre d'arrivée des *stimuli* à l'hypothétique ligne d'arrivée de la conscience. Or pour y parvenir, il nous faut déterminer une seule ligne d'arrivée, fixe et toujours identique. Autrement, une confusion s'installera : un *stimulus* peut arriver premier sur une ligne d'arrivée, mais être battu par un autre *stimulus* sur une autre ligne d'arrivée. Comment déterminer laquelle est prioritaire dans l'ordre que l'on cherche à déterminer (Dennett, 1993, p. 142) ?

On pourrait déterminer, si on le souhaitait, l'ordre dans lequel sont faites les discriminations pertinentes dans le cas qui nous intéresse – ainsi il est raisonnable de supposer que dans un cas tel que celui de « la femme à lunettes » donné en exemple par Dennett, le *stimulus* produit par la vue d'une femme aux cheveux longs dans la rue est traité d'abord, et déclenche le souvenir d'une femme aux cheveux courts et à lunettes, qui vient influencer sur le précédent *stimulus*. Ce qu'il est impossible de connaître, en revanche, c'est la place de ces événements relativement à un point fixe qui correspondrait à « l'expérience consciente effective » (Dennett, 1993, p. 155). Dans l'exemple évoqué plus haut, il existe deux possibilités de réponse à cette question, qui s'accorderont aisément avec ce que nous avons déjà posé (notre cerveau a d'abord reçu les informations visuelles provenant de la femme que nous avons croisée dans la rue, avant de mobiliser le souvenir d'une autre femme), et rendront compte aussi adéquatement que possible du phénomène, tout en s'opposant quant à savoir quand placer le « moment » de la conscience dans la séquence.

Nous pourrions proposer que nous avons eu conscience d'avoir vu une femme aux cheveux longs, avant que notre souvenir de cet événement ne soit modifié par l'image de la femme à lunettes ; nous aurions ainsi d'abord eu une expérience de femme aux cheveux longs, dont le souvenir aurait été effacé et remplacé par autre chose. Mais nous pouvons également imaginer que la perception visuelle de la femme aux cheveux longs a été soumise à révision *avant* son accession à la conscience, de telle manière que ce qui a accédé par la suite à la conscience était une information révisée en amont, de façon pré-expérientielle. Dans le premier cas, nous avons d'abord correctement perçu un événement, dont le souvenir a ensuite été falsifié – c'est ce que Dennett appelle une révision « orwellienne », en référence à l'activité du ministère de la Vérité qui consiste à réécrire l'histoire au profit du Parti dans le 1984 de George Orwell. Dans le second cas, notre perception a été contaminée avant qu'on en ait fait l'expérience et nous n'avons pu accéder qu'au faux souvenir – ce que Dennett appelle une révision « stalinienne », en référence aux pratiques du dirigeant soviétique consistant à monter de toute pièce de faux procès contre ses opposants politiques (Dennett, 1993, p. 151).

À des intervalles de temps très courts, comme ceux qu'on observe lors d'activités cérébrales, il nous est impossible de faire la différence entre révision perceptive (stalinienne, *i.e.* j'ai toujours eu conscience d'une fausse information, celle-ci ayant été falsifiée avant que j'en prenne conscience) ou mémorielle (orwellienne, *i.e.* j'ai eu conscience d'une information vraie mais celle-ci a ensuite été remplacée dans ma mémoire par une information falsifiée) comme on pourrait la faire à l'échelle macroscopique (Dennett, 1993, pp. 141-142). Ainsi, lorsque l'on tente d'expliquer des phénomènes par l'un de ces types de révisions, l'autre offrira toujours une explication également adéquate de ce phénomène, de sorte que l'on aura toujours deux explications opposées mais qui rendent aussi bien compte l'une que l'autre du phénomène et entre lesquelles on pourra choisir, ce qui rend complètement arbitraire le choix de la ligne d'arrivée que nous cherchons (Dennett, 1993, p. 163). Par conséquent, si l'endroit où l'on place la ligne d'arrivée de la conscience est indifférent aux faits mais relève plutôt de nuances de langage, il faut en conclure que c'est une « différence qui ne fait pas de différence », et pas un point crucial de notre théorie : il faut donc chercher ailleurs (que dans l'ordre temporel des discriminations, seulement vérifiable en posant arbitrairement une hypothétique ligne d'arrivée de la conscience) la règle qui fixe l'ordre subjectif des événements dans l'expérience consciente (Dennett, 1993, p. 156).

24 • Examen des concepts de Daniel Dennett

Une fois démenti le mythe de la « ligne d'arrivée de la conscience », il reste à s'attaquer au « théâtre cartésien » à proprement parler, c'est-à-dire au « lieu fonctionnel » où sont projetées les « entités phénoménologiques » (Dennett, 1993, p. 164). Pour cela, Dennett revient d'abord sur la spécificité du phénomène *phi* pour les couleurs.

Le phénomène *phi* pour les couleurs est une illusion perceptuelle décrite pour la première fois par les psychologues Paul Kolars et Michael von Grünau, qui ont investigué le phénomène sur une suggestion du philosophe Nelson Goodman. Voici comment se présente l'expérience : sur un écran, on allume successivement deux points éloignés de couleurs différentes (c'est là la différence avec l'expérience du phénomène *phi* classique tel que décrit par Max Wertheimer). Dans certaines conditions d'espacement et de rythme d'allumage de ces deux points, l'observateur aura l'illusion qu'un seul point se déplace et change abruptement de couleur au milieu de sa course. Or cette conclusion pose un problème : lorsqu'on demande à l'observateur quand le point a changé de couleur, celui-ci répond qu'il lui a semblé que le point changeait de couleur durant sa course, soit *avant* que le point d'une autre couleur ne se soit allumé. Goodman conclut de ces observations que le phénomène « semble nous laisser le choix entre une théorie de la construction rétrospective et une croyance en la voyance »¹¹.

Dennett donne raison à l'hypothèse de la construction rétrospective proposée par Nelson Goodman pour expliquer la phénomène *phi* des couleurs, en y ajoutant qu'il est nécessaire de s'accorder sur une interprétation dans un sens faible de l'expression (Dennett, 1993, p. 166) : la construction rétrospective n'est pas un film passé à rebours dans le temps pour le plaisir d'un auditoire dans un théâtre intérieur ; il n'y a ni auditoire, ni théâtre, ni film d'ailleurs.

Soutenir le sens fort de la construction rétrospective, c'est-à-dire envisager celle-ci comme une construction dans un sens physique, produite par le cerveau et y occupant un lieu et un moment particulier, c'est se ranger du côté du matérialisme cartésien, puisque cette construction a hypothétiquement pour vocation d'être une sorte de film (puisque'il s'agit de re-présenter un mouvement) projeté à rebours dans le temps dans un théâtre cartésien. Pour Dennett cette interprétation est fautive, mais il garde de l'hypothèse de la construction rétrospective l'idée d'une « référence à un moment passé » « incluse dans le contenu » d'un état conscient (Dennett, 1993, pp. 165-166). Une telle référence n'est en

¹¹ Goodman, N., *Ways of worldmaking*, 1978. Cité dans (Dennett, 1993, p. 165).

effet pas incompatible avec nos ambitions physicalistes si on l'envisage comme une propriété ajoutée au contenu en cours de modification et qui débouchera dès lors sur la conclusion qu'un changement s'est produit à tel instant au cours d'un processus : il faut imaginer le cerveau se saisir d'abord de l'image d'un point vert (par exemple), puis de celle d'un point rouge, puis saisir leur clignotement comme une relation, interpréter celle-ci comme un mouvement, et finalement établir qu'un point vert en un lieu A qui termine son mouvement en tant que point rouge en un lieu B a forcément dû changer de couleur quelque part entre les deux, et de conclure que ce changement a eu lieu durant le processus (à peu près en milieu de course, *a priori*). Notons que c'est bien le compte rendu de l'observateur qui constituera la « version finale » qui décrit le changement de couleur ; au niveau neuronal, nulle précognition n'a eu lieu, chaque événement présenté à l'observateur a été reçu et interprété dans un ordre tout à fait naturel : celui-ci a d'abord perçu le point vert, puis la disparition de celui-ci, puis le point rouge à quelques distances du point vert, etc. L'observateur ne peut inférer le mouvement qu'il décrit en fin de compte de la seule perception du point vert.

La conclusion du processus contiendra le renvoi à rebours dans le temps sans que l'événement qu'il suppose ait eu besoin d'être réellement replacé dans une bande repassant la scène dans l'ordre chronologique. De façon imagée, Dennett rapproche la référence dont il est question ici à une affirmation telle que « ce roman nous emmène dans la Rome ancienne » (Dennett, 1993, p. 166) : bien sûr aucun roman ne nous permet de voyager dans le temps, simplement sa construction est telle que nous replaçons ses propositions à une époque passée.

La projection impliquée par l'interprétation au sens fort de la construction rétrospective est une impossibilité physique (il n'existe aucune toile blanche dans le cerveau qui puisse recevoir le film de notre expérience phénoménale), mais l'intuition qu'il existe quelque chose de ce type est néanmoins très prégnante. Si tous peuvent s'accorder aujourd'hui sur le fait que la « projection » qui semble résulter de nos processus sensoriels ne se trouve pas dans l'espace physique, et qu'elle n'y produit aucun effet, la réponse traditionnelle à cet état de fait a été de placer cette projection dans notre « espace phénoménal » (Dennett, 1993, p. 167). Mais il serait trompeur de déplacer simplement le problème en cherchant un théâtre phénoménal plutôt qu'un théâtre physique. Selon Dennett, l'espace phénoménal a « le sens d'un espace logique » (Dennett, 1993, p. 168) : il n'est à chercher nulle part parce qu'il s'agit d'une abstraction permise par l'exploitation des

26 • Examen des concepts de Daniel Dennett

données récoltées au fil des « multiples versions » qui émanent à chaque instant de nos réseaux spécialisés de neurones.

L'enjeu ici est bien de démystifier le fonctionnement du flux de conscience tel qu'on l'envisage en général intuitivement. Dennett se concentre sur un point problématique, encore aujourd'hui, y compris pour les spécialistes, de notre interprétation de la conscience : celui des rapports entre le temps réel et le temps représenté dans notre expérience. Des expériences comme le phénomène *phi* pour les couleurs, mais aussi la célèbre expérience de Benjamin Libet ayant permis la théorisation de son « potentiel de préparation », servent d'arguments aux adeptes du réalisme phénoménal et même aux dualistes parce que leur interprétation repose sur une conception problématique de la conscience, que le modèle proposé par Dennett entend corriger.

L'idée qu'il existe quelque chose comme une expérience phénoménale qui se présente comme une succession d'images ayant lieu en quasi temps réel et dans un ordre chronologique ne résiste pas à l'analyse dès lors qu'on prend soin de distinguer le « véhicule » et son « contenu », c'est-à-dire ce qui représente (un influx nerveux dans le cerveau) et ce qui est représenté (le contenu porté par cet influx), les deux évoluant selon des cadres de références distincts (Dennett, 1993, p. 175) – tout comme l'espace tel qu'il existe extérieurement à nous ne se retrouve pas tel quel dans notre expérience, il ne faut pas s'attendre à trouver le même temps dans le contenu de nos expériences qu'en dehors de celles-ci.

Le contenu dont il est ici question fait partie de cet « espace logique » que l'on se représente fallacieusement comme un théâtre cartésien. En effet, l'on est tenté de donner à ce qui nous apparaît les attributs du réel tout en sachant qu'ils ne le sont pas à proprement parler : nous savons, à nouveau, que notre phénoménalité n'est pas une reproduction à l'échelle du réel qui se trouve tel quelle quelque part dans le cerveau. Pourtant, nous continuons à commettre l'erreur de poser fallacieusement un concept « d'apparence réelle » qui devrait donc être projetée quelque part ; nous pensons que si une perception fautive ne renvoie à rien de réel dans le monde extérieur, les données qui la constituent doivent néanmoins donner lieu à quelque chose de réel : l'apparence, qui serait projetée dans un espace phénoménal (Dennett, 1993, pp. 168-169). Cette intuition est fallacieuse selon le MVM, parce qu'elle postule que ces données doivent être présentées dans le théâtre cartésien, or ce n'est pas nécessaire puisqu'elles ont déjà été traitées dans le jugement (faux dans le cas d'une perception fautive) portant sur les événements réels qui a suivi la

perception. Postuler une « apparence réelle » revient ainsi à « multiplier les entités au-delà de ce qui est nécessaire », et même « au-delà de ce qui est possible » puisque les théories donnant foi à cette intuition ne peuvent éviter de tomber ultimement dans les travers du dualisme cartésien (Dennett, 1993, p. 172).

Nous reviendrons sur l'erreur de catégorie que représente le concept d'apparence réelle. Pour l'instant, tirons une autre conclusion importante de la distinction entre véhicules et contenus des expériences conscientes : les véhicules dont les contenus composent la conscience sont bien localisés dans l'espace et dans le temps, mais « leurs surgissements ne sont pas la marque du surgissement de la conscience de leurs contenus » (Dennett, 1993, p. 148). Il ne suffira donc pas pour « trouver » la conscience, de sonder le système nerveux à la recherche d'événements notoires qui marqueraient le passage de l'inconscient au conscient. Nous ne saurons jamais où et quand regarder, car les questions d'où et quand surgissent la conscience sont trompeuses. Il n'y a pas de lieu ni de moment unique d'apparition de la conscience ; elle est plutôt un processus qui s'étend à l'échelle d'un agent tout entier, et à l'échelle d'un moment qui va du *stimulus* au comportement, de l'input à l'output.

Le MVM est donc une conception parallèle et distribuée de l'activité cérébrale. Ce modèle suppose un changement radical dans la façon d'envisager la conscience, qui sera abordé juste après. Mais avant cela, et en guise de conclusion, rappelons les grandes caractéristiques du MVM, telles qu'elles sont résumées dans l'article de *Scholarpedia* dédié au modèle (Dennett & Akins, 2008). Le modèle des versions multiples, d'après Dennett et Akins, est pensé pour être compatible avec diverses théories au sujet de la conscience, à condition que celles-ci honorent ses propositions fondamentales, qu'ils résument comme suit :

- 1) Là où les théories adhérant au matérialisme cartésien postulent une unité centrale toute puissante responsable de la production d'un flux de conscience, le MVM entend distribuer ce travail entre plusieurs unités moins spécialisées. Dennett utilise l'image de l'homoncule cartésien (l'entité qui, au centre du cerveau, serait la spectatrice du film qui se joue dans le théâtre cartésien) qui devrait être remplacé par une armée d'homoncules plus petits et idiots, pour rendre en métaphore ce passage d'une conception centralisée de la conscience à un processus distribué dans le temps et l'espace.

- 2) Une fois les discriminations adéquates réalisées par les différentes unités spécialisées chacune dans un certain type de tâches, c'est-à-dire une fois que les homoncules ont fait leur travail, celles-ci n'ont plus besoin d'être réitérées à des fins de re-présentation. Le contenu ainsi produit n'a pas à être à nouveau « perçu, discriminé, apprécié ou détesté, ni n'a besoin d'être déplacé quelque part et présenté à nouveau » (à un quelconque théâtre cartésien) pour être « stocké en mémoire ».
- 3) L'action conjointe de ces unités forme un processus « massivement parallèle » dans le cerveau au sein duquel de nombreuses opérations de modification de contenu, souvent incompatibles, ont lieu « simultanément et de manière asynchrone ». Ces opérations sont les « versions multiples » auxquelles le nom du modèle fait référence, et c'est la réalisation de « coups de sonde » par l'agent qui permet d'élever rétrospectivement certaines de ces versions, aux dépens des autres, au rang de « version finale » (celle qui semble être projetée sur le théâtre cartésien). En l'absence d'une telle intervention de l'agent, la question de si un contenu était conscient ou non n'a pas de sens.
- 4) Par conséquent, puisque ce sont ces mécanismes parallèles ayant lieu en divers lieux et moments dans le cerveau qui forment ce que nous appelons ultimement la conscience, il n'y a pas lieu de se demander « quand exactement [avons-nous], par opposition aux diverses parties de [notre] cerveau, été informé, conscient, de certains événements » (Dennett & Akins, 2008). Cette question n'a effectivement pas de sens puisqu'elle repose sur le faux présupposé qu'il existe une « ligne d'arrivée » de la conscience dans le cerveau, ce qui, selon Dennett, comme nous l'avons vu, est faux.

1.4 Conclusion

Daniel Dennett a affirmé n'avoir fait que mettre un mot sur ce qui était déjà, de fait, la méthode appliquée par la plupart des acteurs des sciences de l'esprit. Son concept a été reçu avec hostilité par certains, persuadés que l'approche qu'elle décrit devait « mettre quelque chose de côté », oublier un aspect particulier de l'expérience de sorte que l'hétérophénoménologie ne pouvait pas réellement être la méthode privilégiée d'une science de la conscience, puisqu'elle n'est pas à propos de la conscience du tout ! Une méthode à la troisième personne ne peut aborder la conscience que d'un point de vue

objectif, or celle-ci (ou en tout cas la partie de celle-ci qui représente un défi réel pour les sciences) est un phénomène essentiellement subjectif.

Pourtant, Dennett n'a eu de cesse de rappeler que l'hétérophénoménologie est la méthode dont nous disposons qui soit la plus apte à améliorer notre compréhension de la conscience : elle nous donne accès à l'ensemble des informations significatives au sujet de la conscience (voir l'argument contre Levine cité plus haut, tiré de (Dennett, 2003, p. 21)) par le biais des comptes rendus verbaux des sujets observés, sans en conclure quoi que ce soit *a priori*, puisqu'elle invite à la suspension du jugement à propos de l'exactitude ou de la réalité sous-jacentes de ces comptes rendus.

L'hétérophénoménologie est la méthode que je souhaite suivre dans la dernière partie de ce travail. Aussi considérerai-je que les comptes rendus verbaux qui y seront inclus constituent des clichés suffisamment nets de la conscience des agents qui les rapportent. Je considérerai également que ces données ne sont que les dernières « versions » d'un chaos de « brouillons » produits par les différents systèmes spécialisés du cerveau, qui dessinent un fil narratif, c'est-à-dire une abstraction produite par le langage de ce qui s'est réellement produit dans le corps de l'agent. Ces deux présupposés étant établis, il me reste à explorer la proposition illusionniste concernant la conscience phénoménale, avant de conclure cette première partie en tirant des notions que j'y ai décrites une conception unie de la conscience qui nous aide à rendre compte des phénomènes qui seront décrits dans la dernière partie.

Chapitre 2 L'illusionnisme

2.1 L'illusionnisme selon Keith Frankish

L'illusionnisme est, en philosophie de l'esprit, une position de l'ordre du matérialisme éliminativiste selon laquelle la conscience phénoménale est une illusion. Cette position adresse une solution radicale au problème difficile de la conscience énoncé par David Chalmers : si la conscience phénoménale semble ne pas pouvoir être expliquée en termes physicalistes, c'est parce qu'elle n'existe pas. Ainsi, le problème se trouve déplacé sur un nouveau terrain *a priori* plus porteur ; il conviendra non plus d'expliquer ce qu'est la conscience phénoménale et comment celle-ci fonctionne, mais d'expliquer pourquoi il nous *semble* que nous ayons une telle conscience : ce que Keith Frankish appelle le « problème de l'illusion » (Frankish, *Illusionism as a theory of consciousness*, 2016, p. 1). Notons que le problème de l'illusion nous place bien d'emblée dans une démarche entièrement physicaliste, puisque résoudre ce problème implique d'expliquer ce qui permet l'existence de l'illusion, ce qui peut revenir à expliquer la nature et le fonctionnement du ou des mécanismes qui nous semblent constituer la conscience phénoménale.

Keith Frankish est considéré comme le premier à avoir proposé ce terme pour distinguer son approche de celle d'autres philosophes éliminativistes, par exemple les opposants radicaux à la psychologie naïve (« *folk psychology* »). La portée du projet illusionniste se veut réduite à l'élucidation des conditions d'apparition de l'illusion de conscience phénoménale, ce qui n'exclut pas *a priori* tout ce que l'éliminativisme peut rejeter pour sa part (Frankish, *Illusionism as a theory of consciousness*, 2016, p. 1). Ainsi, l'illusionnisme ne se veut pas à proprement parler irréaliste concernant les propriétés phénoménales, mais celles-ci seront traitées de façon différente que chez les réalistes¹² afin de ne pas tomber dans les mêmes problèmes insolubles que ceux-ci (le gouffre explicatif).

Pour les illusionnistes, l'introspection – outil privilégié d'analyse de nos états mentaux, et le seul en fait qui puisse par définition en déceler les dimensions les plus subjectives – nous fournit des informations trompeuses à propos de ce qui se passe en nous :

¹² Tout au long de ce travail, j'utiliserai les termes « réalisme » et « réaliste » pour désigner les conceptions qui ne sont pas éliminativistes ou illusionnistes.

32 • L'illusionnisme

ce qu'elle nous présente comme des attributs phénoménaux ne sont que des processus physiques complexes, déformés et simplifiés. C'est le manque d'information, ou plutôt la façon trompeuse qu'à notre cerveau de se représenter une information nécessairement parcellaire, qui produit des illusions. Ainsi, dans le cinéma, une série d'images fixes représentant un mouvement, projetées à environ 40 millisecondes d'intervalle, nous font apparaître un mouvement en raison du phénomène de persistance rétinienne. « Le cinéma, dit Frankish, est un artefact des limites de la vision, et, pourraient dire les illusionnistes, la conscience phénoménale est un artefact des limites de l'introspection » (Frankish, *Illusionism as a theory of consciousness*, 2016, p. 6).

Les illusionnistes n'affirment donc pas que rien ne se passe en fait en nous lorsque nous pensons entrer dans des états phénoménaux. Comme Frankish le note, les illusions sont elles-mêmes réelles et peuvent avoir un impact sur le réel (Frankish, *Illusionism as a theory of consciousness*, 2016, p. 1). L'illusion de la conscience phénoménale est comme l'illusion produite par un prestidigitateur : un événement magique en apparence, qui cache en réalité une manipulation complexe et d'ordre purement matériel. De même, les illusionnistes ne nient pas que les propriétés phénoménales, réduites à l'état de simples illusions, soient « causalement actives ». Elles le sont, mais au même titre que les « idées, les histoires, les théories et les mêmes » (Frankish, *Illusionism as a theory of consciousness*, 2016, p. 12), en tant qu'objets de nos états intentionnels. Il n'est donc pas question de nier toute réalité à la conscience phénoménale, ou d'affirmer qu'il n'y a pas lieu de mentionner celles-ci ! La conscience phénoménale est réelle, mais seulement dans le sens où une illusion le serait.

Frankish propose, pour permettre le développement d'un argumentaire sur cette base, le terme de « propriétés quasi-phénoménales » pour faire référence à des propriétés physiques qui semblent phénoménales, et que l'introspection sera amenée à représenter comme telles, sans présenter le caractère irréductible des véritables propriétés phénoménales au sens des réalistes (Frankish, *Illusionism as a theory of consciousness*, 2016, p. 4).

Cette distinction a tout son sens pour distinguer deux approches possibles, apparemment compatibles avec le physicalisme, concernant la conscience phénoménale : on pourrait, en effet, défendre que certaines propriétés habituellement conférées aux propriétés phénoménales, et problématiques voire apparemment incompatibles avec le physicalisme, sont illusoire (par exemple l'ineffabilité, le caractère privé ou infaillible de

celles-ci), mais que nos expériences possèdent néanmoins des propriétés réellement qualitatives (qu'il conviendrait donc d'appeler propriétés phénoménales). Cette position, que Frankish appelle « illusionnisme faible », est à rejeter au profit d'un « illusionnisme fort », qui nie pour sa part tout aspect « qualitatif » aux expériences conscientes, de même que toute autre propriété – physique ou non – apparemment problématiques, celles-ci étant au compte des illusions de la conscience.

Néanmoins, puisqu'il ne s'agit pas de nier les données de l'expérience consciente même, l'illusionnisme fort doit être en mesure de rendre compte de la présence de propriétés *apparemment* phénoménales – des propriétés purement physiques, donc – qui possèdent néanmoins la particularité de toujours s'accompagner, au sein d'une représentation, de l'*illusion* de phénoménalité. C'est pour désigner ces propriétés-là que Frankish propose la notion de propriétés quasi-phénoménales.

Pour sa part, l'illusionnisme faible devra être en mesure de proposer un concept de phénoménalité qui se situe entre le concept traditionnel (réaliste, conférant aux propriétés phénoménales un ensemble de propriétés problématiques pour le physicalisme) et le concept de quasi-phénoménal (la voie de l'illusion totale), car le premier ne soutiendra pas un traitement physicaliste conservateur (laissant le choix entre renoncer au physicalisme, ou rejoindre les radicaux en attente d'une révolution théorique qui permettra enfin de combler le gouffre explicatif), tandis qu'endosser le second revient à accepter l'illusionnisme fort. Frankish considère qu'un tel projet semble impossible¹³, ce qui ne laisse pas d'alternative entre le réalisme radical (la conscience phénoménale existe réellement, mais reste inexplicable dans l'état actuel de nos connaissances) et l'illusionnisme fort, qui semble alors être la seule voie de recherche capable de nous approcher de façon pragmatique d'une explication complète de la conscience (si l'on choisit de laisser de côté toute tentative d'explication non-physicaliste, ce qui semble raisonnable) (Frankish, *Illusionism as a theory of consciousness*, 2016, pp. 3-4).

David Chalmers reconnaît également la faiblesse de l'illusionnisme faible, ajoutant que la discussion d'un point de vue illusionniste au sujet de propriétés problématiques telles

¹³Dans un autre texte, Frankish détaille pourquoi cela lui semble impossible : selon lui, il est impossible de forger un concept satisfaisant de *qualia* qui soit débarrassé des propriétés problématiques habituelles des *qualia* tout en en conservant suffisamment d'attributs pour ne pas devenir de simples propriétés quasi-phénoménales (Frankish, *Quining diet qualia*, 2012). Le détail de cette démonstration dépasse le cadre de ce texte, je me contenterai d'en retenir la conclusion : la voie moyenne entre le réalisme radical et l'illusionnisme est bien plus problématique qu'elle ne semble l'être au premier abord, et l'est même bien plus que l'illusionnisme lui-même.

34 • L'illusionnisme

que le caractère intrinsèque, l'ineffabilité, etc. de la conscience phénoménale ne permettra pas de répondre au problème difficile de la conscience. Que l'on croie à ces propriétés ou pas, le problème difficile reste entier du moment que l'on reconnaît qu'il existe un « effet que cela fait d'être nous », puisque c'est uniquement ce phénomène que le problème difficile cherche à expliquer. La seule façon de répondre d'un point de vue illusionniste au problème semble donc être de nier entièrement l'existence d'un effet que cela ferait d'être nous, et donc de suivre un illusionnisme fort (Chalmers, *The meta-problem of consciousness*, 2018)¹⁴.

Il semble que la seule alternative crédible et compatible avec le physicalisme à l'illusionnisme est une position radicale qui fait dépendre l'explication de la conscience d'une révolution scientifique. Le coût de cette hypothèse est donc énorme, et on comprendra aisément que beaucoup de théoriciens préfèrent s'en tenir à une approche conservatrice. À défaut d'être intuitive, la position illusionniste semble du moins être la seule solution entièrement compatible avec le physicalisme qui laisse encore croire en une réelle avancée dans l'étude de la conscience. Entre les mystères du gouffre explicatif et la folie illusionniste, la plupart ont choisi les premiers. Mais d'autres préféreront, pour reprendre la belle formule de Keith Frankish, « une illusion explicable à une réalité inexplicable » (Frankish, 2016, p. 12).

Reste que la proposition illusionniste est très contre-intuitive, ce qui est certainement la raison principale de la résistance à son encontre. À ce sujet, Frankish a également quelques arguments frappants : d'abord, pour Frankish « la question n'est pas de savoir si l'illusionnisme est intuitivement plausible, mais plutôt s'il est rationnellement attirant » (Frankish, 2016, p. 20), puisque si l'illusion existe réellement, elle a probablement émergé pour des raisons adaptatives. En effet, comme pour les illusions perceptives, il semble que la croyance théorique que notre expérience phénoménale est une illusion n'a pas d'influence sur celle-ci, ce qui pointe vers une origine adaptative de l'illusion (Frankish, 2016, p. 6). Elle serait « câblée » de façon innée dans notre cerveau, et il faut donc s'attendre à ce que son fonctionnement soit relativement impénétrable cognitivement à l'introspection (Frankish, 2016, p. 20).

¹⁴ Chalmers note qu'il existe un type d'illusionnisme faible qui échappe à la critique, celui qui affirmerait que la conscience phénoménale est un concept fonctionnel. Mais, comme il l'explique, cette position devrait en fait être comprise comme un illusionnisme fort, puisqu'il revient à nier l'existence de la conscience au sens où elle pose un problème difficile (si la conscience phénoménale est fonctionnelle, le gouffre explicatif disparaît) (Chalmers, *The meta-problem of consciousness*, 2018, p. 50).

Un second argument repose sur un raisonnement abductif et un principe de parcimonie (Frankish, 2016, p. 13), il pourrait se résumer comme suit :

1. *Une anomalie évidente est généralement le signe d'une illusion.*
2. *Beaucoup de raisonnements en faveur d'une approche non-illusionniste de la conscience présentent des anomalies manifestes (le gouffre explicatif, les zombies philosophiques, l'aspect purement subjectif de la conscience phénoménale). Soit ces anomalies sont effectivement le signe d'une illusion, soit elles ne le sont pas.*
3. *S'il existe une possibilité que notre conscience phénoménale soit une illusion, ces anomalies supporteront le fait qu'elle en soit effectivement une (en raison de (1)).*
4. *Or il semble qu'il soit possible que la conscience phénoménale soit une illusion :*
 - 4.1. *Si nos cerveaux sont des systèmes représentationnels, alors nous devons représenter des propriétés pour en tirer des états cognitifs, affectifs ou motivationnels signifiants (« in order to have cognitive, affective, or motivational significance ») - nous devons posséder un accès cognitif à nos contenus mentaux et être en mesure de former des croyances à propos de ceux-ci pour avoir un « effet que cela fait » d'être nous.*
 - 4.2 *Or nous n'avons aucune façon de vérifier introspectivement la véracité de ces représentations.*
 - 4.3 *Nous ne pouvons donc pas exclure la possibilité que celles-ci ne soient pas réelles.*
5. *De la possibilité que les propriétés phénoménales que nous nous représentons soient le fruit d'une illusion de la conscience (4.3), nous pouvons tirer qu'il est possible que les anomalies vers lesquelles pointent les raisonnements au sujet des propriétés phénoménales soient le signe que la conscience phénoménale est bien une illusion (2 et 3).*
6. *Nous pouvons inférer abductivement (en raison de (1)), qu'il est probable que les propriétés phénoménales soient une illusion.*

Posé succinctement, l'argument peut se résumer ainsi : « Si une propriété résiste à l'explication en termes physiques, ou n'est détectable que selon une certaine perspective, l'explication la plus simple [de cette propriété] est qu'elle est une illusion » (Frankish, 2016, p. 13). Cet argument souligne à nouveau le scrupule méthodologique illusionniste : il s'agit bien d'un projet conservateur et pragmatique (au sens commun du terme), en ce qu'il entend donner des réponses claires à des problèmes posés clairement, sans que ceci n'implique de révolution dans nos théories actuelles.

Les propriétés phénoménales sont encore trop mystérieuses en l'état pour faire partie d'une théorie scientifique. Les écarter au profit de l'illusion permet de relancer la recherche d'une explication réelle de la conscience, en contournant le problème difficile qui se présente comme un horizon théorique indépassable. Ce que l'illusionnisme propose,

c'est une voie de chantier certes peu attractive au premier regard, mais qui possède l'avantage de rendre possible les avancées en philosophie de l'esprit en ne fermant pas *a priori* certains terrains pourtant inexplorés. En outre, l'illusionnisme a l'avantage de présenter peu de risques en raison de son postulat de départ plus faible que celui du réalisme : si, au fil de la recherche sur le problème de l'illusion, l'on parvenait à la conclusion que la raison pour laquelle il nous semble que nous avons des expériences phénoménales est que nous en avons réellement, l'illusionniste n'en serait pas troublé. Son problème se serait réduit à nouveau au problème difficile, auquel il pourrait à nouveau s'attaquer, fort de ses nouvelles découvertes.

Enfin, Frankish ajoute également que nier la réalité de la conscience phénoménale permet aux illusionnistes de conserver le caractère « magique, non-physique » de celle-ci, rendant ainsi mieux compte, en un sens, de nos intuitions premières à son sujet que ne le feraient les réalistes (en accord avec le physicalisme). Ces derniers, considérant les propriétés phénoménales comme des causes physiques, sont obligés de reconnaître qu'elles ne sont pas « aussi magiques qu'elles en ont l'air » (Frankish, 2016, p. 14). Au contraire, rejeter les propriétés phénoménales dans l'intentionnalité nous permet de leur reconnaître l'aspect magique qu'elles semblent avoir, les traitant comme des fictions susceptibles de nous affecter d'autant mieux qu'elles semblent posséder un statut particulier. Ces fictions reflètent bien entendu ce qu'il se passe réellement dans le cerveau de ceux qui les imaginent (Frankish, 2016, p. 9), mais il n'y a pas de raison de croire qu'elles cachent une réalité particulièrement mystérieuse ou problématique pour le physicalisme. Daniel Dennett, en plus d'être l'un des défenseurs les plus importants de l'illusionnisme, a beaucoup exploré ce lien entre la conscience phénoménale et la fiction, raison pour laquelle il occupe une place centrale dans ce travail. Intéressons-nous donc à présent à son projet illusionniste.

2.2 L'illusionnisme selon Daniel Dennett

Daniel Dennett est un défenseur de l'illusionnisme, et ses concepts présentés plus haut vont dans le sens du projet illusionniste en permettant une remise en cause sur des bases sérieuses des conceptions traditionnelles de la conscience. Néanmoins l'illusionnisme de Dennett n'est pas celui d'autres penseurs, aussi est-il intéressant de comparer plusieurs positions illusionnistes pour mieux se positionner dans le débat.

Dennett a été très élogieux à propos de l'article programmatique de Frankish (Dennett, 2016) et a professé avoir toujours été un défenseur de l'illusionnisme fort

(Dennett, 2019). Il est donc tout à fait consciemment en accord avec la thèse selon laquelle la conscience phénoménale, dans sa totalité, est une illusion : nos expériences ne possèdent en réalité aucune propriété qualitative, et encore moins d'autres propriétés problématiques telles que l'ineffabilité. Il décrit cette illusion comme proche dans le principe des tours de magie d'un illusionniste – dans le sens courant du terme, un « prestidigitateur » (remarquant par-là que le terme de Frankish se trouve particulièrement bien choisi). On dit parfois de l'illusionniste qu'il pratique une forme de « magie », mais on entend par là tout autre chose que la magie des sorciers et autres thaumaturges ; ce que l'illusionniste réalise, ce sont des tours de passe-passe, des manœuvres complexes qui font naître une illusion de magie, sans avoir recours à aucun pouvoir occulte (Dennett, 2016, pp. 65-66).

L'illusion ainsi produite est tout à fait réelle : elle provient de causes matérielles, et produit des effets observables dans le monde. Il n'y a rien de paranormal dans celle-ci, et nous nous mettrons aisément d'accord sur le fait qu'il y ait « un truc », et qu'expliquer le tour de l'illusionniste ne nécessitera pas de révolution dans les lois de la physique (Dennett, 2016, p. 66). Néanmoins, et bien qu'il n'y ait en réalité rien d'anormal dans la situation, il nous *semble* que ce soit le cas. À condition que notre illusionniste réalise son tour avec la plus grande dextérité, nous verrons bien se produire quelque chose qui *ressemble assez* à de la magie : une femme coupée en deux se trouve « recollée », un lapin sort d'un chapeau, un objet disparaît pour réapparaître autre part...

C'est peut-être ce qui nous pousse à faire confiance à nos intuitions concernant la conscience phénoménale : tout nous *semble* réel, et nous ne pensons pas avoir de raisons de trouver cela suspect. Après tout, n'est-ce pas plus parcimonieux de supposer que tout est vrai, plutôt que de poser la possibilité d'une mécanique d'une complexité affolante destinée uniquement à nous tromper ? À nouveau, le parallèle avec l'illusionnisme (le « vrai ») se fait pertinent : tout le travail du prestidigitateur est de construire des tours dont les spectateurs ne pourraient pas imaginer les ficelles, afin de faire vaciller leurs certitudes. C'est ce que Dennett appelle « *suffering a failure of imagination and mistaking it for an insight into necessity* » : si quelque chose nous semble trop difficile à imaginer, nous faisons l'erreur d'en inférer que cela est nécessairement impossible (Dennett, 2016, p. 67). Pour autant, nous ne prendrions jamais au sérieux un prestidigitateur qui affirmerait que ses tours restent inexplicables par des causes ordinaires, qu'ils sont de la « vraie magie » (Dennett, 2016, p. 67). Aussi est-il nécessaire de ne pas se rendre à cette intuition d'inexplicabilité, en tout cas pas avant d'avoir épuisé toutes les possibilités offertes par le

38 • L'illusionnisme

physicalisme (Dennett, 2016, p. 67). L'illusionnisme est l'une de ces possibilités, comme nous l'avons vu plus haut avec l'argument de Frankish, il semble donc important de lui laisser sa chance ; s'arrêter dans la recherche sur la conscience aux « intuitions inébranlables » (Dennett, 2016, p. 67) de certains philosophes revient à faire une pétition de principe contre toute autre approche avant même d'avoir réellement avancé un seul argument.

Dennett va plus loin en affirmant même que l'illusionnisme devrait être vu, en dépit de son aspect contre-intuitif, comme vrai « jusqu'à preuve du contraire » et donc comme le paradigme dans le domaine, étant donné qu'il s'agit d'une théorie conservatrice (qui n'implique pas de révolution dans nos paradigmes scientifiques), qui répond néanmoins du caractère anormal de la conscience phénoménale. Adopter le point de vue illusionniste permettrait aux philosophes, selon Dennett, de recentrer le débat sur de « vraies » questions et de laisser place à une exploration conservatrice et entièrement physicaliste des problèmes posés par la conscience (Dennett, 2016, p. 68).

La plus importante de ces questions est sans doute la « question difficile » de Dennett : « et ensuite, que se passe-t-il ? » Si l'on s'en tient à une explication au niveau personnel de nos expériences, c'est-à-dire à une description susceptible d'être produite par un examen introspectif de celles-ci, nous n'en obtenons que des descriptions sommaires ; celles-ci renvoient certes à une facette de notre vie mentale, mais n'en constituent en aucun cas une explication. Car une fois décrit le « contenu phénoménal » de nos expériences, il reste encore à expliquer comment celui-ci est « implémenté dans des processus subpersonnels ». Comment le cerveau réalise-t-il en pratique ce qui définit notre phénoménalité ? Représentant et contenu représentés doivent pouvoir se rencontrer sans que l'on ait à faire appel à un miracle ou à l'action d'un homonculus dans un théâtre cartésien... Or aussi bien les neuroscientifiques que les philosophes ou les psychologues cognitifs semblent s'arrêter de l'un ou de l'autre côté du problème sans chercher à réellement combler le vide, ce qui les empêche nécessairement d'avancer dans la résolution du problème difficile de la conscience (Dennett, 2019, pp. 51-52).

Ainsi de Crick et Koch qui, présentant leur théorie d'un mécanisme lié à la conscience visuelle provoquant la synchronisation des neurones à une fréquence oscillatoire de 40 à 70 Hz (Crick & Koch, 1990), s'arrêtent dans leur analyse au constat que ce mécanisme rend possible une fonction de présentation que posséderait la conscience, sans préciser à qui ou à quoi serait destinée cette présentation. À l'inverse, la psychologie

cognitive offre des modèles, tels que celui de « l'espace de travail global », avec lesquels ils représentent le chemin allant de la fonction cérébrale au comportement, en laissant à d'autres le soin d'expliquer comment cette fonction se trouve réalisée dans le cerveau (Dennett, 1993, pp. 320-321).

Le problème difficile de la conscience ne trouvera pas de solution dans une approche au niveau personnel de la conscience (Dennett, 2019, p. 56). En posant la question difficile et en tentant d'y répondre, nous évitons de tomber dans le piège consistant à arrêter notre recherche à un élément à première vue inexplicable. La description des phénomènes au niveau personnel, telle que la pratique la phénoménologie, possède son importance, mais elle doit nous amener à dialoguer avec les sciences de l'esprit et non nous arrêter à des affirmations péremptoires basées uniquement sur une intuition, fût-elle particulièrement forte.

Tout au long de sa carrière, Dennett s'est évertué à remettre en question les faiblesses des arguments des philosophes qui défendent une conception réaliste de la conscience phénoménale en attaquant dans une perspective illusionniste les limites de l'autophénoménologie telle qu'elle se trouve encore pratiquée aujourd'hui. C'est là un de ses chevaux de bataille, et l'aspect de sa philosophie qui nous intéressera particulièrement. La critique offerte par Dennett de concepts et expériences de pensée liés aux *qualia* comprend deux facettes qui la rendent à mon sens particulièrement percutante : d'une part Dennett fonde son argumentation sur les données de la science, en naviguant entre sa grande culture scientifique classique et les derniers travaux en date, et d'autre part il pointe du doigt les pièges du langage dont sont truffés les débats en philosophie de l'esprit. Ce scrupule à rester au plus près des sciences, ainsi que cette approche héritée de son maître Gilbert Ryle et de Wittgenstein, sont les deux outils qui permettent à Dennett d'armer notre imagination pour nous aider à dépasser les intuitions phénoménologiques qui nous retiennent de poser les bonnes questions au sujet de la conscience (Dennett, 1993, p. 30).

Il semblerait en effet que Dennett soit optimiste sur notre capacité à nous débarrasser d'intuitions fallacieuses au sujet de la conscience. C'est là une particularité de son illusionnisme, partagée par d'autres mais également critiquée par certains. En effet, si, au sein de l'illusionnisme, le réductionnisme au sujet de la conscience est accepté, et par conséquent l'idée que résoudre le problème de l'illusion dissoudra le problème difficile, il n'en reste pas moins des dissensions, notamment au sujet de la nature de l'illusion, dont on cherche à comprendre l'origine et le fonctionnement.

Partie II
LA NATURE DE L'ILLUSION DE LA
CONSCIENCE

Chapitre 3 Illusion riche ou légère ?

3.1 Introduction au problème de la richesse de l'illusion de la conscience

François Kammerer propose de distinguer les théories illusionnistes selon qu'elles envisagent l'illusion de la conscience comme « cognitivement pénétrable »¹⁵ ou non. Cette notion de pénétrabilité cognitive possède une certaine importance pour l'argumentation illusionniste, puisqu'elle apparaît comme une parade raisonnable contre l'argument réaliste consistant à pointer simplement du doigt la force de l'intuition qu'il existe quelque chose comme des propriétés phénoménales : si la conscience phénoménale est une illusion, pourquoi me semble-t-il, de manière persistante et résistante à l'acquisition de nouvelles informations, qu'elle n'en est pas une ? Pour les illusionnistes, la force de cette intuition peut provenir de ce que, bien qu'elle ne soit qu'une illusion, celle-ci est rendue particulièrement tenace par le fait qu'il nous est impossible de réellement changer le contenu de nos croyances à son sujet, elle est en quelque sorte encapsulée et rendue inaccessible à l'appareil critique de la conscience de niveau supérieur (Kammerer, 2022, p. 4).

Néanmoins, il est possible d'envisager l'illusion comme possédant un degré plus ou moins important « d'impénétrabilité cognitive ». Kammerer propose de distinguer les théories illusionnistes selon qu'elles postulent un haut degré d'impénétrabilité cognitive ou non (Kammerer, 2022).

Il existerait en effet deux types d'illusions bien différentes, que Kammerer illustre par deux exemples simples : d'un côté il y a les illusions perceptives telles que les deux lignes de Müller-Lyer, d'un autre les illusions produites par les magiciens tels que le tour de la « femme sans tête » (Kammerer, 2022, pp. 4-5). Dans le cas des lignes de Müller-

¹⁵ D'après Kammerer, l'impénétrabilité cognitive « peut être comprise comme l'encapsulation informationnelle par rapport à l'information stockée dans la mémoire centrale – souvent sous forme de croyances » (Kammerer, 2022, p. 4), il faut donc comprendre le terme inverse (pénétrabilité cognitive) comme désignant une *non-encapsulation* d'un processus par rapport aux composantes rationnelles de la conscience, telles que les croyances. Le terme est issu de Pylyshyn, qui désigne par « pénétrabilité cognitive » « l'altérabilité explicable rationnellement du comportement d'un composant en réponse aux changements de buts et de croyances » (Pylyshyn, 1985, p. 133). En d'autres termes, est pénétrable cognitivement tout processus dont la modulation provient d'un changement explicable au niveau sémantique (au niveau des croyances, buts...) et non en termes d'architecture fonctionnelle (Pylyshyn, 1985, pp. 132-135).

44 • Illusion riche ou légère ?

Lyer, nous voyons deux lignes parallèles de tailles différentes. Lorsque nous apprenons que les deux lignes sont en fait de même taille, nous ne cessons pas de voir les lignes de tailles différentes. Bien que notre croyance que les lignes sont de tailles différentes ait été détruite et remplacée par la croyance que les lignes sont de taille égale, nous n'en continuons pas moins à les voir différentes. Les illusions de ce type sont dites « riches » car elles possèdent un haut niveau d'impénétrabilité cognitive, c'est-à-dire qu'elles se distinguent d'une fausse croyance par le fait qu'elles résistent à l'analyse et à l'apport de nouvelles connaissances.

À l'inverse, les tours de magie présentent des illusions dites « légères » parce que leur force repose sur le fait que nous les percevons de façon incomplète. Si le tour de magie semble présenter un événement impossible, c'est en raison du fait que certaines informations cruciales pour que nous puissions comprendre exactement ce qu'il se passe nous sont cachées. Dans le tour de la « femme sans tête », une femme dont la tête est couverte d'un tissu noir est placée dos à un arrière-plan noir sur une scène illuminée de façon précise. Le spectateur est incapable de voir sa tête, et pourra donc en inférer qu'elle n'en a pas. Pourtant, s'il connaissait « le truc », s'il prenait conscience des artifices responsables de l'illusion, il serait en mesure de briser l'illusion. En effet, si sa représentation de la femme dont la tête n'est pas visible est, elle, cognitivement impénétrable, sa représentation d'une femme sans tête, découlant d'un processus inférentiel, est quant à elle cognitivement pénétrable et peut être corrigée par un nouvel apport d'information ; ainsi un spectateur naïf verra simplement une femme sans tête, là où le spectateur ayant compris « le truc » ne verra plus qu'une femme qui *semble* ne pas avoir de tête en raison d'une illusion. Il sera en mesure de surmonter l'illusion, ce qui était impossible dans le cas de l'illusion de Müller-Lyer dans laquelle l'apport de nouvelles informations ne permet pas de voir autre chose que deux lignes de longueurs inégales.

Les illusions légères sont donc des représentations correctes mais incomplètes, menant à un jugement incorrect à propos de la situation. Dans leur cas, même si la représentation en elle-même est impénétrable (l'activation de cette représentation se fait d'une façon impénétrable cognitivement), le processus inférentiel permettant de passer de la représentation incomplète au jugement faux est, quant à lui, parfaitement pénétrable cognitivement : si le sujet est suffisamment prudent, il peut éviter de glisser de l'une à l'autre (Kammerer, 2022, p. 5). Face à une illusion légère, nous pouvons faire confiance à nos capacités de perception, mais c'est des jugements basés sur celles-ci qu'il faudra se méfier.

Les illusions riches sont des représentations incorrectes activées à travers un processus entièrement impénétrable cognitivement, ce qui implique que l'acquisition de nouvelles informations ne soit pas en mesure d'entraver cette activation (Kammerer, 2022, p. 4). Dans ce cas, ce ne sont pas nos jugements qui produisent l'illusion, mais nos systèmes perceptuels eux-mêmes, et c'est donc de leur verdict que nous devons nous méfier (Kammerer, 2022, p. 5).

Toutes les illusions sont résistantes jusqu'à un certain point à l'acquisition de nouvelles informations « encodées dans un format semblable à celui des croyances à propos de la situation qui fait l'objet de l'illusion » (Kammerer, 2022, p. 4). Néanmoins les illusions riches de Kammerer le sont plus que ses illusions légères, puisque dans les premières nous recevons une information fausse que nous n'avons pas de moyen de contester, tandis que dans les secondes nous recevons une information incomplète – que nous n'avons pas plus de moyen de contester – qui nous pousse vers des inférences fausses mais, pour leur part, contestables.

Pour Kammerer, Dennett présente plutôt l'illusion de la conscience comme une illusion légère. C'est en tout cas ce que laisse penser la métaphore de l'illusion de l'utilisateur que celui-ci utilise pour décrire la conscience phénoménale (Kammerer, 2022, p. 7). D'après lui, notre introspection ne nous donne que des informations très simplifiées, sous forme de représentations, que nous pouvons interpréter à tort comme étant des états phénoménaux simples, de la même façon que nous pourrions être tentés d'interpréter les icônes et les dossiers présents sur notre écran d'ordinateur comme des réalités intrinsèques, alors qu'il ne s'agit que de représentations schématiques d'états structurels complexes produits par la machine. En d'autres termes, nous sommes face à nos « *qualia* » comme un utilisateur naïf est face aux icônes de son écran : nous avons tendance à croire qu'il s'agit d'entités simples existant réellement parce que nous ignorons le fait qu'il s'agit en fait d'une représentation erronée, en raison du fonctionnement de notre cerveau, d'un état complexe de celui-ci dont les détails restent inaccessibles à l'introspection (Kammerer, 2022, p. 8).

L'illusion de l'utilisateur est légère, puisqu'elle provient d'une représentation incomplète du fonctionnement de l'ordinateur (c'est-à-dire une représentation qui ne nous donne qu'une partie seulement de l'information sur le fonctionnement de celui-ci, sans pour autant être *stricto sensu* trompeuse à son sujet – en nous donnant pour vraie une information fausse, ou inversement), représentation donnée de façon impénétrable cognitivement, et qui

46 • Illusion riche ou légère ?

amène l'utilisateur à produire des inférences fausses au sujet du système (le fait que les icônes qu'il voit possèdent une réalité intrinsèque) de façon pénétrable cognitivement (si nous ne pouvons pas remettre en question le contenu de la représentation incomplète, il n'en va pas de même pour les inférences subséquentes). Même si de façon « systématique, habituelle et non-réflexive » (Kammerer, 2022, p. 8) nous aurons tendance à considérer l'interface comme une réalité en soi, la pénétrabilité cognitive de telles inférences assure que nous pourrions revenir réflexivement sur ce constat, moyennant un travail sur nos présupposés au sujet de la machine.

En va-t-il de même dans le cas de l'illusion de la conscience phénoménale ? Il me semble que, dans ce cas-ci, elle ne soit pas parfaitement comparable avec l'illusion de l'utilisateur. Plus exactement, la première entre bien dans un certain sens dans la catégorie des illusions légères (ou peut-être dans celle des illusions « ni-riche-ni-légère », comme nous le verrons ci-après), mais sa spécificité lui permet d'échapper aux critiques de Kammerer à l'encontre de la position qui tiendrait l'illusion de la conscience pour une illusion légère. Car en effet, l'illusion de l'utilisateur n'est qu'une analogie imparfaite de ce qui se produit dans la conscience (Kammerer, 2022, p. 13) (Dennett, 1993, p. 565), et je souhaiterais proposer une hypothèse quant à cette spécificité qui ferait de l'illusion de la conscience une illusion légère (suffisamment, en tout cas, pour être dépassée) tout en rendant compte de la ténacité de notre intuition anti-physicaliste au sujet de la conscience.

L'illusion de l'utilisateur est une illusion légère, qu'on considérera peut-être intuitivement comme telle étant donné que nous sommes plutôt familiers des ordinateurs. Intuitivement à nouveau, il nous semble que l'illusion de la conscience est beaucoup plus tenace. Kammerer en déduit qu'il est par conséquent plus raisonnable de considérer celle-ci comme une illusion riche (Kammerer, 2022, p. 12). Je pense que ses conclusions au sujet de la richesse de l'illusion ne sont pas les bonnes, notamment parce qu'il sous-estime le rôle de la croyance dans la formation et le renforcement de celle-ci. Pour soutenir mon propos, je commencerai par résumer ici certains arguments pour et à l'encontre de l'hypothèse de l'illusion légère donnés par Kammerer.

3.2 Défense de l'hypothèse de l'illusion légère

Kammerer souligne à juste titre deux avantages majeurs des théories de l'illusion légère : d'abord, étant donné qu'elles ne considèrent pas l'introspection comme une

machinerie délivrant des informations systématiquement fausses, elles n'ont pas à expliquer la raison de cette tromperie systématique. En effet, si considérer que l'introspection délivre des informations parcellaires au sujet de ce qui se passe en nous semble raisonnable, imaginer qu'elle produit systématiquement un mensonge à ce sujet (en présentant ces informations comme possédant des qualités extraordinaires) semble théoriquement plus coûteux : il s'agirait alors d'expliquer non seulement comment l'introspection parvient à produire ces mensonges, mais également quelle serait la fonction de ce mécanisme. Qu'est-ce qui, évolutivement, pourrait expliquer l'apparition d'un « système complexe de représentations introspectives qui délivre systématiquement des représentations incorrectes » (Kammerer, 2022, p. 9)^{16 17} ?

Cela est d'autant plus compliqué que les propriétés phénoménales semblent posséder des qualités étranges et particulières. Le contenu de représentations fausses non phénoménales possède en principe un lien avec des propriétés existantes, soit qu'il s'agisse d'une représentation d'une propriété parfois instanciée dans la réalité, mais qui ne soit pas le cas ici (par exemple dans l'illusion de Müller-Lyer, les deux lignes semblent de tailles différentes, la différence de taille étant une propriété qui ne pose aucun problème, simplement elle est ici mobilisée à tort dans notre représentation), soit qu'il s'agisse d'une représentation d'une propriété complexe jamais instanciée dans la réalité, mais dont les composants peuvent l'être (Kammerer donne l'exemple d'une montagne dorée : celle-ci n'existe pas, mais les montagnes et les choses dorées existent bien) (Kammerer, 2022, p. 10). Les théories de l'illusion riche devront donner les raisons de ce fait mystérieux, que l'introspection nous présente les propriétés phénoménales comme étant différentes de ou extérieures à la réalité, et comme possédant des qualités subjectives qui les distinguent de tout phénomène physique.

Les théories de l'illusion légère déplacent ce problème à un autre niveau, puisqu'elles considèrent que ce n'est pas l'introspection qui donne aux représentations les

¹⁶ Le besoin d'explication du mécanisme d'un point de vue évolutif est à relier au postulat innéiste selon lequel ledit mécanisme serait une fonction naturelle du cerveau humain, partagé par l'espèce entière. D'après Kammerer, la majorité des théories de l'illusion riche soutiennent ce type d'innéisme. Il n'est pas impossible, toujours selon Kammerer d'imaginer une théorie de l'illusion riche qui ne soit pas innéiste, et ne rencontre donc pas le même problème, néanmoins l'exemple proposé, une théorie qui postulerait que le mécanisme introspectif cognitivement impénétrable ne serait pas inné mais construit durant l'enfance, ne me semble pas suffisant puisqu'il n'explique pas pour quelles raisons un tel mécanisme serait forcément impénétrable cognitivement (Kammerer, 2022, p. 9n9).

¹⁷ Pour une discussion sur l'intérêt évolutif de la conscience, voir l'ouvrage de Nicholas Humphrey *Soul Dust : The Magic of Consciousness* (Princeton University Press, 2011).

48 • Illusion riche ou légère ?

propriétés problématiques, mais le processus inférentiel qui en résulte. Le problème devient donc d'expliquer comment des états tels que des croyances peuvent posséder des propriétés apparemment impossibles. Or il est plus facile d'imaginer que cela puisse être le cas, puisque nos théories sémantiques naturalistes permettent d'envisager qu'il existe dans le cas des représentations conceptuelles une absence de liens causaux réels entre la représentation et les propriétés représentées, et que par conséquent ces propriétés ne sont pas nécessaires et pourraient tout aussi bien ne pas être instanciées dans la réalité (Kammerer, 2022, pp. 10-11). Par exemple, après avoir vu le tour de la « femme sans tête », je pourrais croire que j'ai effectivement vu une femme sans tête, mais je peux aussi croire que j'ai été témoin d'une illusion ; ma représentation d'une femme sans tête est en réalité une représentation d'une femme dont la tête ne m'est pas visible, la propriété « sans tête » n'est pas instanciée dans cette situation, mais elle peut être ajoutée par inférence à ma représentation, par exemple si je crois en l'existence de la « vraie » magie¹⁸. Ainsi les théories de l'illusion légère seraient plus aptes que les théories de l'illusion riche à expliquer l'existence des propriétés phénoménales, puisque celles-ci seront considérées comme existant uniquement au niveau du contenu des croyances, et non comme étant données par quelques mécanismes du cerveau.

En dépit de ces arguments, Kammerer souhaite continuer à défendre une théorie de l'illusion riche. Voyons à présent l'argument qu'il propose en faveur d'une telle théorie, et contre les théories de l'illusion légère. Celui-ci prend la forme suivante :

1. Si les théories de l'illusion légère sont vraies, alors « quand nous sommes informés du fait que nos représentations introspectives de nos états internes sont en fait schématiques et incomplètes, de sorte que nos états internes ont une nature intrinsèque qui ne nous est pas présentée à travers l'introspection (à savoir une nature physico-fonctionnelle), notre réticence à accepter cette nouvelle information ne devrait pas être substantiellement plus forte et plus résiliente que la réticence dont nous faisons l'expérience quand nous apprenons que n'importe laquelle de nos autres représentations incomplètes "standards" sont réellement incomplètes (Kammerer, 2022, p. 11). »

2. Or « nous sommes plus réticents à accepter la nature physico-fonctionnelle de nos états internes que nous le sommes à accepter que n'importe

¹⁸ Si la « femme sans tête » avait été une illusion riche, il aurait fallu supposer que j'aie reçu une représentation d'une femme sans tête (de façon cognitivement impénétrable) – et non une représentation d'une femme dont la tête n'est pas visible – et donc il aurait été nécessaire que cette représentation ait été causée par l'instanciation des propriétés « femme » et « sans tête », puisque celles-ci étaient par hypothèse bien présentes dans ma représentation sans que des processus inférentiels aient eu besoin d'intervenir. Il aurait alors fallu expliquer d'où vient le fait que je puisse apposer la propriété « sans tête » à une représentation, ce qui dans ce cas ne pose pas de problème, puisque cette propriété peut effectivement se trouver instanciée dans le monde, mais pose question dans le cas des propriétés phénoménales qui ne sont par définition jamais instanciées dans le monde.

laquelle de nos autres représentations incomplètes "standards" soit incomplète (Kammerer, 2022, p. 12). »

3. Donc les théories de l'illusion légère sont fausses.

Il ne me semble pas nécessaire de commenter d'emblée ce syllogisme, qui apparaît plutôt clair. En revanche, je souhaite émettre des remarques au sujet de ses différentes parties.

Commentaire du syllogisme de Kammerer (1)

La première prémisse affirme que si l'illusion de la conscience est effectivement une illusion légère, alors quand nous apprenons l'existence de cette illusion (c'est-à-dire quand « nous sommes informés du fait que nos représentations introspectives de nos états internes sont en fait schématiques et incomplètes » (Kammerer, 2022, p. 11)), nous ne devrions pas avoir plus de mal à accepter cette nouvelle information que nous n'en avons eu en apprenant la réalité sous-jacente à n'importe quelle autre illusion légère. Cette prémisse sous-entend qu'il ne pourrait pas y avoir de différence notable de résistance à des illusions légères pourtant très différentes. Or cela est manifestement faux, puisqu'il me semble au contraire que certains facteurs (par exemple la proximité de l'illusion, au sujet de laquelle je reviendrai dans mon développement au sujet de la seconde prémisse de l'argument de Kammerer) peuvent influencer notre résistance à une illusion, sans remettre en question la légèreté de celle-ci (c'est-à-dire le fait que cette illusion provienne bien d'une inférence trompeuse entraînée par une représentation schématique et incomplète).

Kammerer défend d'abord cette prémisse en en offrant une version plus faible : si l'illusion de la conscience était conçue purement sur le modèle de l'illusion de l'utilisateur [...] alors en effet nous ne devrions pas être plus réticents à apprendre la nature physique de la conscience que nous ne le sommes à apprendre la vraie nature cachée de fichiers d'ordinateur (Kammerer, 2022, p. 13). Cette version de la prémisse instruit le problème principal de l'argument de Kammerer : il repose sur un parallèle établi entre des cas très différents en gommant leurs spécificités pour se concentrer uniquement sur un aspect limité du problème (ici, la richesse de l'illusion, tel que Kammerer la définit). Or, comme Kammerer le souligne également, un partisan de l'illusion faible rappellerait justement que l'illusion de l'utilisateur n'est qu'une « analogie imparfaite » (Kammerer, 2022, p. 13) de l'illusion de la conscience, il ne faut donc pas s'attendre à ce que tout se passe dans le cas de l'illusion de la conscience exactement comme cela se passe avec l'illusion de l'utilisateur. Ainsi, la condition de l'implication proposée est manifestement fautive en l'état,

50 • Illusion riche ou légère ?

puisque aucune théorie n'affirme que l'illusion de la conscience devrait être « conçue purement sur le même modèle » que l'illusion de l'utilisateur. Les deux illusions ne sont pas suffisamment proches dans leur fonctionnement pour que nous nous attendions à pouvoir tirer des conclusions à propos de la conscience sur base d'une recherche à propos de l'illusion de l'utilisateur.

Kammerer fait peser la charge de la preuve sur les théories de l'illusion légère : si l'on souhaite outrepasser sa première prémisse, il faudrait d'abord définir les différences déterminantes entre les deux illusions qui justifient la fausseté de la condition de la prémisse. Or justement personne n'a affirmé qu'il y avait plus de ressemblances entre elles que ce que décrit l'analogie à l'origine, à savoir que l'illusion de l'utilisateur fournit une illustration simplifiée de ce qui se produit dans le cas de la conscience – un accès insuffisant au fonctionnement interne nous pousse à inférer l'existence d'entités qui ne sont en réalité que des caricatures de ce qui se produit réellement dans la machinerie.

Il faudrait donc, en tant que théoricien de l'illusion légère, commencer par s'interroger sur les ressemblances et différences entre nos deux illusions. Kammerer, pour sa part, propose deux exemples de différences qui, d'après lui, répondraient au problème mais feraient alors de notre théorie une théorie de l'illusion riche (Kammerer, 2022, p. 13), mais discute également d'un exemple d'explication de la spécificité de l'illusion de conscience qui nous amènerait vers une théorie de l'illusion « ni-riche-ni-légère » (Kammerer, 2022, p. 14). Les deux premiers exemples ne m'intéresseront pas ici, puisqu'ils proposent une différence entre les cas de la conscience et de l'illusion de l'utilisateur uniquement au niveau cognitif. Or pour ma part je souhaite m'interroger avant cela sur les similarités et différences entre les deux du point de vue comportemental et hétérophénoménologique, puisqu'à ce niveau les différences deviennent manifestes. La troisième possibilité ouverte par Kammerer, qui amène selon lui à une théorie de l'illusion « ni-riche-ni-légère », fait intervenir la force de la croyance comme facteur explicatif de la différence entre le cas de la conscience et l'illusion de l'utilisateur. Interroger le rôle de la croyance dans l'illusion me semble pertinent, mais celui-ci doit être envisagé d'un point de vue plus large que le simple point de vue cognitif afin de comprendre en quoi la croyance peut expliquer notre réticence à accepter l'illusionnisme sans pour autant que nous ayons à recourir à une explication en termes d'impénétrabilité cognitive.

Étant donné que la théorie de l'illusion légère défendue par Dennett fait reposer sur la ténacité des croyances la force de l'illusion de la conscience, plutôt que sur le

fonctionnement en soi de la machinerie introspective, et étant donné que ce qui donne à Kammerer l'intuition d'une différence entre illusions riches et légères est la résistance des sujets à l'acquisition de nouvelles croyances au sujet de ces illusions, il serait intéressant de s'intéresser aux différentes croyances qui sont en jeu ici. Car en effet, si l'illusion de l'utilisateur est une bonne analogie pour comprendre l'illusion d'un point de vue fonctionnel, elle l'est moins si l'on souhaite établir une comparaison du point de vue de la richesse étant donné qu'*a priori* les croyances entrant en jeu dans les deux cas sont de nature totalement différente (elles ne font pas du tout appel aux mêmes répertoires conceptuels, aux mêmes domaines de la connaissance, aux mêmes usages et habitudes comportementales, etc.). Comment tirer du fait que la croyance (ou non) à la réalité intrinsèque des icônes d'un ordinateur semble facilement pénétrable la conclusion que la croyance en la réalité des *qualia* ne l'est pas, alors que l'une et l'autre croyance n'ont que bien peu en commun ?

Il serait pertinent d'envisager un autre exemple que celui de l'illusion de l'utilisateur, qui mobilise des croyances notablement plus solides et susceptibles de résister à l'apport de nouvelles informations en dépit de la légèreté manifeste de l'illusion sur laquelle elles reposent. Kammerer proposait l'exemple du tour de magie de la « femme sans tête » comme exemple d'illusion légère, je souhaiterais envisager un exemple proche, tiré également du monde de la prestidigitation – un exemple qu'il sera aisément possible de classer parmi les cas d'illusions légères (telles que définies par Kammerer) : les tours d'apparitions fantomatiques.

Il fut une époque où ces tours étaient donnés en représentation par des charlatans qui prétendaient être capables d'évoquer de vrais esprits. À la fin du XIX^{ème} siècle et au début du XX^{ème}, à une époque de regain d'intérêt pour l'occultisme, ce type de manifestation avait un réel succès et attirait aussi bien les convaincus qui y assistaient dans l'espoir d'y trouver un contact avec un proche décédé, que des sceptiques présents pour tenter de percer le mystère du tour (Christine Bergé donne un aperçu de cette époque et de son contexte épistémologique dans son ouvrage *La voix des esprits* (Bergé, 1990), qui retrace l'histoire du mouvement spirite). Or si l'on devait interroger un croyant et un sceptique au sortir d'une représentation, l'un et l'autre nous donneraient un compte rendu semblable mais différent : « il y avait *réellement* un fantôme ! » affirmerait le Croyant, « non, répondrait le Sceptique, il *semble* qu'il y avait un fantôme, mais il n'y en avait pas

52 • Illusion riche ou légère ?

en réalité ». Le second aborderait les faits comme provenant d'une illusion, alors que pour le premier il devait s'agir d'une réalité.

Imaginons que, chacun pour convaincre l'autre, le Sceptique et le Croyant se soient donné rendez-vous pour une autre représentation. Soucieux d'analyser le moindre détail de la situation, ils s'assoient au premier rang et portent une attention particulière à l'apparition : ils voient ainsi avec autant de détails que possible la forme diaphane d'une personne se matérialiser sur scène. Au sortir de la représentation, le Croyant assène au Sceptique : « Alors ? Tu vois, il y avait *réellement* quelque chose. Tu ne vas pas me dire que tu n'as rien vu ? ». « J'ai bien vu quelque chose *qui ressemble assez* à l'image que je me fais d'un fantôme, répond le Sceptique, pas le moins du monde convaincu. Mais ce n'en était pas un, car les fantômes n'existent pas... Il doit y avoir un truc. »

De la deuxième représentation, ni l'un ni l'autre n'a appris quoi que ce soit qui soit susceptible de le faire changer d'avis : à peine le Croyant a-t-il vu ses croyances dans le paranormal renforcées, et peut-être le Sceptique a-t-il profité de l'occasion pour déjà explorer certaines pistes : « étant donné la transparence de la forme et la façon dont elle apparaît, je pense qu'on peut éliminer l'hypothèse d'un complice déguisé... » Mais le Sceptique entend bien avoir raison, et il possède – pense-t-il – un avantage sur le Croyant pour cela : « Il suffirait que je parvienne à reproduire le tour, en dévoilant les ficelles, pour convaincre le Croyant qu'il s'agit d'une supercherie, réfléchit-il ! Si je parviens à montrer qu'il est possible de reproduire l'apparition du fantôme grâce à un trucage, alors peut-être que le Croyant acceptera qu'il soit plus probable que ce “fantôme” que nous avons vu n'était qu'une illusion. »

Et voici qu'il entame des recherches pour percer le mystère du tour. Il s'entoure d'illusionnistes (qui l'aident avec plaisir à démasquer le charlatan dont les mensonges entachent la profession), de professionnels du théâtre (qui apportent leurs connaissances des artifices de la scène), et de scientifiques (physiciens et chimistes qui procurent le matériel et le cadre nécessaire aux recherches), avec qui il travaille à reproduire le plus fidèlement possible le tour du charlatan. Après quelques semaines, ils parviennent à un résultat satisfaisant, et le Sceptique propose donc à son ami Croyant d'y assister.

Ils se retrouvent donc dans une salle de spectacle similaire à celle où ils ont assisté aux représentations du charlatan. Le Sceptique invite ses complices à effectuer le tour. À la fin de celui-ci, le Sceptique se tourne vers le Croyant :

« Alors, comment trouves-tu mon fantôme ?

– C'était très troublant, répond le Croyant, hésitant. Il m'a bien *semblé* voir une apparition semblable à celle de la dernière fois... »

– Et bien cher ami, vous serez peut-être déçu d'apprendre que tout ce que vous avez vu n'était qu'un tour de magie élaboré ! Je me suis entouré d'experts pour monter cette petite expérience pour vous montrer, précisément, que l'apparition de la dernière fois n'était rien de plus qu'une illusion ! »

– Ah ! Je ne suis pas étonné, car en réalité il me semblait bien que *quelque chose* n'allait pas. Malheureusement, votre apparition n'était qu'une pâle copie des *vrais* esprits que les médiums sont capables d'évoquer. Certes il m'a *semblé* voir, à l'instant, une apparition fantomatique, mais quelque chose clochait. J'ai bien senti que celle-ci, contrairement à la *vraie* apparition, n'était qu'une illusion. »

Le Sceptique reste sans voix : de son point de vue, l'exécution du tour était sans défaut – l'illusion était parfaite – et aurait dû produire exactement le même effet sur son ami que le tour du Charlatan. Mais d'une façon ou d'une autre, le Croyant a vu clair dans son jeu et a perçu l'artifice dans ce cas-ci alors qu'il y était resté aveugle dans les cas précédents. À moins que, ses croyances ayant été ici mises à rude épreuve, il ait préféré renier son impression première (celle d'avoir effectivement vu un fantôme ici aussi) afin de ne pas être mis face à une contradiction...

Par hypothèse, le Croyant a bien vu un fantôme, semblable en tout point à celui qu'il avait vu précédemment, il n'y a pas de différence *a priori* entre son expérience à ce moment et lors de la représentation organisée par le Sceptique (c'est là un présupposé important pour la conclusion de mon expérience de pensée, mais je le pense à la fois intuitif et crédible étant donné les faits dont celle-ci s'inspire). Le Croyant aurait donc bien dû être, lors de la révélation de la supercherie, mis face à un certain malaise. Plus précisément, il a dû opter pour l'une des deux solutions suivantes : soit accepter, puisque le Sceptique est parvenu à le duper, qu'il est possible que le médium-charlatan y soit parvenu également et qu'il n'y ait jamais eu de fantôme dans toute cette histoire mais seulement des illusions – cela revient à remettre en question une précédente certitude en acceptant la possibilité de la fausseté d'une croyance (« un certain médium parvient à évoquer de vrais fantômes ») qui participe elle-même de la croyance plus large qu'il existe des fantômes –, soit imaginer une différence entre les deux situations (fonctionnellement, elles sont identiques – j'exclus

54 • Illusion riche ou légère ?

ici la possibilité que le médium dont il a été question soit réellement doté de pouvoirs magiques, évidemment) qui permette d'ôter tout soupçon du médium – ce qui revient à éviter la remise en question de la précédente croyance par un petit tour de passe-passe intellectuel.

Si le Croyant est conscient de ce dilemme, ce qui n'est pas nécessairement le cas (le raisonnement décrit ci-dessus peut avoir lieu sous forme d'inférences inconscientes), il choisira peut-être par pure honnêteté intellectuelle la première solution et reconnaîtra son erreur. Si ce n'est pas le cas en revanche, c'est très certainement la seconde solution qui s'imposera. Mais dans les deux cas, ce qui n'est pas certain, c'est que le Croyant verra sa croyance évoluer : même s'il reconnaît s'être fait avoir, il faudra un travail intellectuel supplémentaire pour que la croyance que le Sceptique est bien parvenu à faire apparaître un fantôme sans recourir à aucune « vraie magie »¹⁹ donne lieu à une recalibration de son système de croyance telle qu'il devienne désormais plus sceptique à l'égard des démonstrations et des récits d'apparitions de fantômes. Tant que cette modification importante n'aura pas eu lieu, le Croyant sera toujours susceptible de se faire bernier par ces mêmes illusions dans le futur.

Ce genre de situation a tout à fait pu avoir lieu à une époque où les croyances occultes étaient très prégnantes. De nombreux sceptiques, scientifiques ou non, sont parvenus à démasquer des charlatans en exposant leurs « trucs » : Faraday démasqua le charlatan Douglas Home en 1853, tandis que certains illusionnistes comme le comte Paul-Alfred de Saint-Geoirs ou l'illustre Harry Houdini s'étaient fait une spécialité de dévoiler les « trucs » de prestidigitateurs utilisés par les médiums (Bergé, 1990, pp. 96-97). A-t-on abandonné la croyance au paranormal pour autant ? Aucunement, elle existe toujours et existera toujours. Soit que l'on repousse les événements paranormaux toujours plus loin hors de l'emprise de la science (jusqu'à l'épiphénoménisme), soit que l'on choisisse d'ignorer ou de nier les preuves, il y a toujours une échappatoire pour esquiver la remise en question et dissoudre la dissonance cognitive. Il ne suffit pas d'être mis face à la vérité pour l'accepter. Même la compréhension profonde d'un phénomène ne suffit pas toujours.

¹⁹ Dennett a souvent cité une anecdote rapportée par Lee Siegel dans son ouvrage *Net of Magic. Wonders and Deceptions in India* (1991), dans laquelle est qualifiée de « vraie magie » la magie qui n'existe pas, par opposition à l'illusionnisme, la prestidigitation qui existe bel et bien mais ne serait pas de la « vraie magie » ... (Dennett, 2017, p. 318)

Faut-il conclure que pour les croyants, les tours qui montrent des phénomènes apparemment paranormaux provoquent en eux une illusion riche – c'est-à-dire que pour eux leur introspection leur fournit une représentation fautive au sujet de la situation, alors que les sceptiques voient quant à eux une illusion légère – c'est-à-dire une représentation incomplète qui les amènent vers un jugement faux à propos de la situation, jugement qu'ils s'empressent de corriger ? Je doute que la croyance modèle à ce point la machinerie cérébrale qu'elle soit capable de provoquer de telles disparités de fonctionnement. En revanche, il est tout à fait plausible que des systèmes de croyance différents poussent deux personnes à faire des inférences opposées à propos d'une situation identique.

Ceux qui croient aux fantômes succomberont à l'illusion et seront susceptibles de refuser l'idée que c'était juste un tour, affirmant qu'ils ont vraiment vu un fantôme. La façon dont certains parviennent à surmonter les dissonances cognitives régulières auxquelles ils font face lorsqu'on leur met sous le nez des preuves devrait suffire à montrer comment une illusion légère (légère pour tout individu sceptique) peut soudainement devenir bien plus forte chez quelqu'un dont les croyances entrent en résonance avec le contenu de l'illusion.

Il est maintenant temps de conclure. Nous avons vu qu'en considérant d'autres exemples que celui de l'illusion de l'utilisateur dans notre réflexion, nous pouvons arriver à des conclusions bien différentes de celles de Kammerer. En prenant un exemple très similaire à celui qu'il avait fourni en premier lieu, je pense avoir donné une idée de la ténacité que peut avoir une illusion même quand celle-ci est envisagée comme étant légère. Nous devons en conclure soit que l'exemple du tour de la « femme sans tête » est un mauvais exemple, soit que les réflexions de Kammerer au sujet de l'illusion légère doivent être remises en question. Je présume que Kammerer a mûrement réfléchi ses exemples durant son travail, je pencherais donc pour la seconde conclusion. Étant donné qu'il nous est possible de décrire un cas d'illusion légère (manifestement) dans lequel certaines personnes sont particulièrement réticentes à accepter la réalité sous-jacente à l'illusion, je conclurai que la première prémisse de l'argument de Kammerer est incorrecte.

Commentaire du syllogisme de Kammerer (2)

Avant de passer à la critique que je souhaiterais émettre au sujet de l'argument même, j'aimerais souligner la présence dans cette seconde prémisse d'une pétition de principe qui, si elle ne me semble pas remettre en cause l'argument par elle-même, permet

56 • Illusion riche ou légère ?

d'entrevoir un élément important de ma critique plus globale de l'argument. La seconde prémisse affirme sur base d'une intuition ainsi que sur base des données « sociologiques » qu'il est plus facile d'accepter l'existence d'une illusion au sujet de la nature d'une représentation dans un cas « standard » (c'est-à-dire, supposément, dans n'importe quel cas qui n'a pas trait à la phénoménologie) que dans le cas d'un état interne (c'est-à-dire, supposément, dans n'importe quel cas qui a trait à la phénoménologie). Dans les cas « standards » proposés en exemples d'illusions légères, l'illusion serait donc plus facile à accepter, étant donné qu'apprendre « que nos représentations sont incomplètes et schématiques » et nous « habituer à cette idée » (Kammerer, 2022, p. 12) nous permettrait de ne plus tomber dans le piège tendu de l'illusion. Or il semble à première vue que cela ne soit pas vrai dans le cas de la conscience.

Cette intuition semble également partagée par beaucoup, si l'on en croit les chiffres des études menées par Chalmers au sujet des opinions de philosophes et étudiants au sujet du physicalisme ou des zombies (Kammerer, 2022, pp. 12-13). Notons toutefois qu'en dépit de l'optimisme de Kammerer au sujet des données évoquées par Chalmers, ce dernier affirme lui-même que rien de concluant au sujet de la conscience ne peut pour l'instant être affirmé sur leur base ; d'autant plus que la (supposée) large distribution d'intuitions anti-physicalistes n'occulte pas l'existence d'intuitions physicalistes possiblement aussi fortes et qui mitigeront toujours la force des affirmations basées sur l'assertion de la prééminence des premières (Chalmers, *The meta-problem of consciousness*, 2018, pp. 13-15). Kammerer travaillera donc également sur cette assertion assez vague, affirmant qu'aucun adulte « éduqué et raisonnablement intelligent » ne resterait persuadé de la réalité suggérée par l'illusion de l'utilisateur, contrairement à l'illusion de la conscience (Kammerer, 2022, p. 13).

Je tenais à souligner rapidement que cet *argumentum ad populum* ne repose pour l'instant que sur des données très maigres, et surtout probablement trop rapidement interprétées. Il est donc questionnable de la part de Kammerer d'asséner, sur cette base, avec tant d'assurance que l'illusion de la conscience est notablement plus difficile à accepter que d'autres illusions telles que l'illusion de l'utilisateur.

Venons-en maintenant à la seconde prémisse de l'argument de Kammerer : celle-ci affirme qu'il existe une différence entre nos intuitions au sujet d'illusions légères telles que

l'illusion de l'utilisateur et nos intuitions au sujet d'illusions riches telles que l'illusion de la conscience²⁰, à savoir que les premières sont plus faciles à accepter que les secondes.

Kammerer impute cette différence à la « richesse » de l'illusion de la conscience : si celle-ci est plus difficile à accepter, c'est qu'elle provient d'une représentation fautive qui nous est donnée par un processus entièrement impénétrable cognitivement. Mais ne pourrait-on pas attribuer la différence entre ces intuitions à autre chose que la richesse de l'illusion ? Par exemple, certains facteurs extérieurs aux mécanismes représentationnels plus ou moins trompeurs dont Kammerer parle ne pourraient-ils pas expliquer aussi bien cette différence sans en appeler à une machinerie introspective ouvertement trompeuse ?

Il existe des différences importantes entre l'illusion de l'utilisateur et l'illusion de la conscience (Dennett n'utilisant la première que comme une métaphore ou une « béquille pour l'imagination », pour reprendre l'une de ses expressions (Dennett, 1993, p. 398)), ce dont Kammerer est très conscient : aussi invite-t-il à clarifier ces différences si l'on devait l'attaquer sur ce point (Kammerer, 2022, p. 13). J'ai envisagé plus haut l'une des différences majeures – la nature de la croyance – qui séparent les deux, je souhaite à présent aborder un autre aspect de cette différence qui me semble permettre de contredire l'affirmation de Kammerer au sujet de notre réticence à accepter l'illusion de l'utilisateur.

Je commencerai par poser deux questions : 1) Pourrait-on aussi facilement accepter l'illusion de l'utilisateur si l'on connaissait aussi peu le fonctionnement et la réalité physique de notre ordinateur que celle de notre cerveau ? 2) Pourrait-on aussi facilement accepter l'illusion de l'utilisateur si l'on était aussi « proche » de l'interface de notre ordinateur que nous le sommes de notre conscience ?

La force d'une illusion dépend de notre compréhension du phénomène illusoire

Concernant la première question, nous pourrions commencer à y répondre par une récolte de données empiriques, non parmi les philosophes mais parmi les chercheurs des domaines des sciences de l'esprit. Les données récoltées par Chalmers et évoquées plus haut (Chalmers, *The meta-problem of consciousness*, 2018, pp. 13-15) montrent avant tout les idées partagées dans le milieu philosophique, et n'en déplaie aux philosophes de

²⁰ Dans la suite de ce développement, je ne mettrai en comparaison que l'illusion de l'utilisateur (comme exemple d'illusion dite « légère ») et l'illusion de la conscience (comme exemple d'illusion dite « riche »), mais le propos devrait être valable dans tous les autres cas, en tout cas dans ceux que Kammerer évoque dans (Kammerer, 2022).

58 • Illusion riche ou légère ?

l'esprit et philosophes des sciences, celles-ci sont parfois bien éloignées des idées ayant cours dans les milieux scientifiques ; Dennett fait remarquer à ce sujet que le concept de zombie philosophique, bien que parfois utilisé par les scientifiques, n'est pas conçu de la même façon par ces derniers – ainsi même sur un terrain qu'on pourrait croire partagé, les intuitions des scientifiques peuvent être très différentes de celles des philosophes (Dennett, 2012, p. 46). Ainsi nous pourrions penser que les chercheurs en neurosciences auront une autre idée sur les problèmes philosophiques de la conscience. Mais ne nous égarons pas à spéculer sur ce sujet. Plutôt, nous pouvons approcher la question par l'autre côté en nous interrogeant sur le caractère intuitif ou non de l'illusion d'utilisateur : à défaut de tirer des conclusions de l'avis de ceux qui en connaissent le plus sur le cerveau à propos de l'illusion de conscience, interrogeons-nous donc sur l'avis de ceux qui en connaissent le plus sur les ordinateurs à propos de l'illusion de l'utilisateur ! À défaut, à nouveau, de données statistiques, les intuitions de ces derniers nous sont sans doute plus accessibles que celles des spécialistes du cerveau, étant donné que nous sommes sans doute bien plus à maîtriser l'informatique que les neurosciences²¹.

Quel serait l'avis d'un expert en informatique (soit qu'il s'agisse d'un professionnel du *hardware* ou d'un chercheur dans un domaine lié aux sciences informatiques) sur l'illusion de l'utilisateur ? Quel serait l'avis, par contraste, d'une personne tout à fait inculte au sujet des ordinateurs (aussi bien en ce qui concerne le *hard* que le *software*) ? Y aurait-il une différence entre ces deux personnes quant à leur façon de saisir le fonctionnement de l'illusion de l'utilisateur, ou le problème posé par celle-ci ?

Nos personnages entrent tous deux dans la catégorie assez vague donnée par Kammerer des adultes « éduqués et raisonnablement intelligents » (Kammerer, 2022, pp. 12-13), ainsi on peut s'attendre de leur part à ce qu'ils ne soient pas dupes au sujet de cette illusion : même s'ils possèdent tous deux une intuition plus ou moins forte que les entités présentes à leur écran semblent posséder une existence propre, aucun des deux ne sera entièrement dupe de l'illusion de l'utilisateur. Aussi bien l'expert en informatique que l'inculte sont conscients que ces entités ne sont que des représentations schématiques d'un état complexe du système dont le substrat physique se trouve dans le boîtier posé à côté de

²¹ Précisons que par hypothèse, il suffit d'être « éduqué » et « raisonnablement intelligent » (Kammerer, 2019b, pp. 12-13) pour percer à jour l'illusion de l'utilisateur (Kammerer ne précise pas à quelle éducation et à quel degré d'intelligence raisonnable il fait référence), alors qu'il est extrêmement difficile (voire impossible) pour n'importe qui (mais peut-être moins pour les chercheurs en neurosciences ?) d'accepter l'illusion de la conscience.

leur bureau – en fait, ils sont probablement plus conscients de ce fait que du fait qu'ils pourraient être victimes d'une illusion au sujet de ces entités lorsqu'ils travaillent sur un ordinateur.

Il existe néanmoins une différence entre eux, qui se montre à l'échelle du comportement. En ce qui concerne les intuitions et autres produits de l'introspection, le comportement « *in situ* » est parfois plus révélateur des croyances profondes d'une personne que ses comptes rendus réflexifs. Ainsi, il suffit d'imaginer les réactions de nos personnages face à un problème informatique pour se rendre compte de l'écart qui peut exister entre l'avis qu'ils professent au sujet de l'illusion de l'utilisateur et l'effet réel de celle-ci sur leur fonctionnement interne. Confrontés, par exemple, à une perte soudaine de l'image sur leur moniteur (un « écran noir »), l'expert en informatique ne manifestera pas la même perplexité que le non-initié. Fort de ses connaissances, il voit dans ce phénomène un panel de problèmes possibles étant survenus au niveau logiciel ou matériel.

J'ai récemment rencontré un tel problème et, sans être un expert, mais un connaisseur assez fin de ma machine, voici ma réaction : en voyant l'image disparaître, j'ai d'abord pensé à un problème de *driver* de la carte graphique, ceux-ci provoquant de temps à autre une saute de l'image suivie d'un message d'erreur. L'image ne revenant pas, j'ai d'abord redémarré la machine – si le problème persistait après le redémarrage, et étant donné que je n'avais rien fait durant ma session qui aurait pu provoquer un problème important, le problème logiciel aurait été exclu. L'écran n'affichant toujours aucune image après plusieurs redémarrages, je passai à l'hypothèse du problème matériel provenant soit de l'écran, soit de l'ordinateur. Après avoir testé l'écran en le branchant à un autre appareil, j'en vins à la conclusion que le problème se trouvait du côté du PC. Un problème d'image pouvant *a priori* provenir d'un problème de carte graphique, je m'assurai du fonctionnement apparent de celle-ci (clignotement des lampes-témoins, fonctionnement normal des ventilateurs), avant de tourner mon attention vers la carte mère. L'écran LCD de celle-ci m'informa alors que la séquence de démarrage ne s'effectuait pas correctement, sans plus d'informations. C'est à ce stade que, le diagnostic précis dépassant mes compétences, j'apportai mon PC chez un réparateur qui me confirma que la carte mère était bien tombée en panne. Je pus remplacer la pièce, réinstaller ce qui devait l'être et redémarrer mon ordinateur.

En comparaison, pour avoir déjà observé des non-initiés tenter de faire face à un problème informatique, je pense pouvoir affirmer que leur réaction est généralement tout

60 • Illusion riche ou légère ?

autre. Souvent, celle-ci implique un recours à la posture intentionnelle pour qualifier le problème : « l'ordinateur ne *veut plus* démarrer ! », « *il* me fait un caprice ou quoi ? », et de conclure qu'il « est mort ». Bien sûr, ces personnes ne pensent pas *réellement* que la machine leur joue un tour, et ils ne s'endeuilleront pas de devoir la porter chez le réparateur. Plutôt, elles usent d'un vocabulaire commode qui pallie à leur manque de connaissances au sujet de la machine.

Il en va de même en ce qui concerne le comportement de chacun face aux symboles à l'écran, qui sont au cœur de l'illusion de l'utilisateur : l'expert mis face à ces symboles a tôt fait de mobiliser ses connaissances (éventuellement de façon préconsciente) au sujet, par exemple, du format de ses fichiers, des procédés d'indexation de la machine, plus généralement du processus qui permet l'affichage grâce aux pixels de l'écran de données stockées sous forme binaires sur les plateaux magnétiques de ses disques-durs. De son côté, le non-initié ne verrait que les icônes comme une ressource manipulable sur l'écran, ou *dans* l'écran, et non au niveau du disque dur (alors que c'est bien là que la majorité des opérations d'un ordinateur commencent et se terminent).

Le non-initié comme le connaisseur affirmeraient sans doute ne pas se laisser berner par l'illusion de l'utilisateur. Mais les faits contrediraient alors l'affirmation du premier : celui-ci agit comme s'il était victime de l'illusion, quoi qu'il affirme par ailleurs à ce sujet. Même s'il peut réflexivement dépasser l'illusion en mobilisant des croyances physicalistes au sujet de la machine, de façon instinctive, il semble plutôt agir comme une sorte de « dualiste dispositionnel ». S'il avait réellement dépassé l'illusion, il devrait être en mesure d'adopter face à la machine un comportement en adéquation avec des croyances entièrement physicalistes à son sujet.

Cet exemple (quelque peu caricatural mais néanmoins, je l'espère, pertinent) montre qu'en dépit de ce que nous pourrions affirmer au sujet de notre résistance à l'illusion de l'utilisateur, nous manifestons un rapport très différent à la machine en fonction de notre degré de connaissances et de compétences en informatique. Dans le cas du non-initié, il est difficile de voir en quoi celui-ci peut affirmer avec certitude ne pas être affecté par l'illusion de l'utilisateur, alors qu'il semble par ailleurs attribuer des états intentionnels à la machine et est incapable de concevoir ses opérations à l'écran autrement que comme des manipulations d'objets. Même si, réflexivement, il en viendrait à nier que ces propos et dispositions renvoient à une quelconque réalité, il n'en reste pas moins qu'intuitivement, il

semble bien réagir face à la machine comme s'il ne concevait pas la réalité physique sous-jacente au fonctionnement de celle-ci

Il semble raisonnable d'en conclure qu'il existe un décalage entre ce que nous pourrions affirmer de nos croyances au sujet de l'illusion de l'utilisateur, et la façon dont elle nous affecte réellement : malgré que nous affirmions réflexivement que l'illusion ne nous affecte pas, nous agissons intuitivement comme si c'était le cas. C'est d'ailleurs entièrement sur cela que repose le sentiment d'immersion dans un jeu vidéo.

Lorsque l'on est plongé dans un jeu dans lequel on contrôle un personnage en première ou troisième personne (je prends l'exemple du jeu vidéo car l'immersion y est bien plus forte que dans n'importe quel autre média), il n'est pas rare que nous adoptions des comportements à l'égard de celui-ci qui sont similaires à ceux que nous aurions pour notre propre corps ou pour celui d'un autre être vivant (le déplacement lui-même dans le monde virtuel s'effectue comme si le corps virtuel possédait une matérialité en raison de l'apparence de lois physiques générées par le programme, les dégâts ou échecs subis par le personnage sont susceptibles de faire sursauter ou réagir impulsivement le joueur, etc.). Il n'est pas nécessaire de postuler un état pathologique d'indiscernement du réel et du virtuel pour expliquer ce type de comportement, simplement tout se passe comme si le virtuel présenté à l'écran était incorporé d'une certaine façon à notre représentation inconsciente du réel, de façon à devenir comme une extension de celle-ci, sans que cela n'implique quoi que ce soit pour nos croyances réflexives.

À ma question de départ, « pourrait-on aussi facilement accepter l'illusion de l'utilisateur si l'on connaissait aussi peu le fonctionnement et la réalité physique de notre ordinateur que celle de notre cerveau ? », je répondrais que non. Ou plutôt nous pourrions aussi facilement l'accepter, mais cela ne garantirait pas que nous n'en soyons pas victime. Une personne qui aurait un niveau de connaissances et de compétences en informatique semblable à notre niveau de connaissances et de compétences en neurosciences serait sans doute très perplexe face à un écran ; dans le meilleur des cas, il serait amené à s'interroger sur toutes les fonctions mathématiques et les technologies qui rendent cette expérience possible, dans le pire il y verrait une démonstration de sorcellerie. Heureusement pour nous, nous connaissons généralement bien mieux l'informatique que le fonctionnement du cerveau, aussi sommes-nous plus à même de déjouer l'illusion induite par la première.

62 • Illusion riche ou légère ?

Comme mon exemple le soulignait, nous pouvons affirmer avoir accepté l'illusion de l'utilisateur, tout en y restant sujet d'une certaine façon. Pouvons-nous réellement dire nous être parfaitement « habitués » (Kammerer, 2022, p. 12) à l'idée que les entités sur l'écran ne sont que des représentations schématiques de ce qui se passe dans l'ordinateur, alors que nous interagissons avec la machine comme si ce n'était pas le cas ? Une fois prises en compte toutes les dimensions de la croyance, et en particulier sa dimension comportementale, nous voyons qu'il n'est pas si évident que l'illusion de l'utilisateur soit une illusion particulièrement légère. Plus légère que l'illusion de la conscience, peut-être, mais la différence qui peut exister entre les deux illusions serait alors plutôt une différence de degré qu'une différence de nature : il n'y a pas des illusions riches et des illusions légères, mais des illusions plus ou moins riches que d'autre. L'illusion de l'utilisateur n'est peut-être pas aussi légère que l'on pourrait le penser, et en tout cas affirmer qu'on ne se laisse pas prendre à celle-ci ne suffit pas à affirmer sa légèreté.

La force d'une illusion dépend de notre proximité avec le phénomène illusoire

J'en viens maintenant à ma seconde question : pourrait-on aussi facilement accepter l'illusion de l'utilisateur si l'on était aussi « proches » de l'interface de notre ordinateur que nous le sommes de notre conscience ? En effet, une autre différence importante qui sépare l'illusion de la conscience de l'illusion de l'utilisateur est que la première se produit *en* nous, alors que la seconde se produit *hors* de nous, nous n'avons donc pas la même proximité avec cette dernière.

En me questionnant sur la « proximité » d'une illusion, j'entends souligner qu'il peut exister entre deux illusions à la fois une différence de proximité géographique – une illusion peut se produire à distance de nous, à proximité de nous, ou même *en* nous, et sa prégnance peut varier selon cette distance géographique – et une différence dans la nature et le nombre de milieux « intermédiaires » séparant la réalité physique (source de l'illusion) de l'observateur (point final de l'illusion). Par exemple, dans le cas de l'illusion de l'utilisateur, l'écran joue un rôle d'intermédiaire important entre nous et le substrat physique de l'illusion (le matériel informatique), ou encore, lorsque l'on voit une paille brisée dans un verre d'eau, c'est la surface de l'eau qui joue le rôle d'écran déformant la réalité (dans ce cas, la partie de la paille immergée dans l'eau) aux yeux de l'observateur. J'envisage donc la possibilité que le niveau de proximité entre la réalité et l'observateur joue également un rôle important dans la force qu'aura une illusion sur cet observateur. La

plus grande proximité que nous pouvons envisager est (supposément) celle qui existe entre nous-même et notre conscience (nous avons un « accès direct » au produit de notre activité consciente), une moins grande proximité serait, par exemple, celle qui existe entre un observateur et un objet réfléchi par plusieurs miroirs déformants placés en série.

Nous ne sommes pas très efficaces pour remettre en cause ce qui se passe en nous, ou ce qui semble s'y passer. Nous avons d'entrée de jeu un recul sur l'illusion de l'utilisateur qui facilite le dépassement de l'illusion. Au contraire, prendre du recul est une étape supplémentaire obligatoire pour dépasser l'illusion de la conscience.

Mais enfin, voici une expérience de pensée qui pourrait aider à montrer pourquoi notre proximité avec une illusion est peut-être un critère plus pertinent pour expliquer sa ténacité que son origine cognitive ou sa richesse (telle que définie par Kammerer) : imaginons que soient commercialisées des puces électroniques destinées à afficher certaines informations en réalité augmentée dans le champ de vision de celui qui les reçoivent. De cette façon, il y aurait supposément un « effet que cela fait » de percevoir ces informations directement dans notre champ de vision, mais elles resteraient néanmoins l'équivalent pour nous d'éléments d'interface sur un écran d'ordinateur : il s'agit de symboles produits par une machine (mais rendus visibles sous forme d'images mentales produites par les aires visuelles du cerveau, plutôt que sous forme de pixels sur un écran).

Les illusionnistes pourraient affirmer qu'il s'agit là d'une illusion légère, puisque malgré que l'on ait transposé l'illusion de l'utilisateur dans notre crâne, nous sommes toujours au courant qu'elle trouve son origine dans un ordinateur, celui-ci étant simplement déplacé de l'extérieur à l'intérieur de notre corps. Ainsi, étant donné notre capacité à accepter les illusions légères, nous serions en mesure 1) de distinguer à volonté le caractère illusoire de l'effet de la puce (c'est-à-dire d'accepter le fait que l'effet de la puce soit un effet d'interface trouvant sa source dans le circuit imprimé de la puce, de reconnaître son caractère symbolique destiné à schématiser un état de fait bien différent en réalité) 2) de distinguer cette illusion de l'illusion de la conscience (puisque, par hypothèse, l'une étant dépassable et non l'autre, nous n'aurions pas le même accès aux deux illusions, et donc pas les mêmes informations à leur sujet).

Néanmoins, imaginons qu'on implante une telle puce à un enfant, qui grandisse dès lors avec, sans jamais qu'on l'informe de la présence de cette puce, ou que l'on y fasse référence. L'enfant développerait alors une plus grande proximité avec l'illusion : de son

64 • Illusion riche ou légère ?

point de vue, elle ferait intégralement partie de son expérience du monde. Pour les illusionnistes défendant une théorie de l'illusion riche, un problème se pose. En effet, il est probable que lorsqu'au bout d'une dizaine d'années (suffisamment pour que l'enfant soit en mesure de produire des croyances fermes à propos de son existence consciente et soit en mesure de comprendre les concepts de conscience, de puce électronique et de réalité augmentée – mais pas suffisamment pour qu'il ait eu le temps d'entamer des études de philosophie et de s'amuser à tenter de comparer ses *qualia* avec ceux de ses amis), l'enfant sera informé qu'une puce électronique modifiant son expérience du monde lui a été implantée, il aura bien du mal à croire à cette nouvelle information.

L'enfant aura beau être parfaitement au courant du fonctionnement d'un ordinateur (peut-être même est-il devenu un petit génie de l'informatique), cette idée le choquera sûrement. Le choc sera d'autant plus grand, je pense, que l'effet provoqué par la puce sera intégré à l'expérience visuelle de façon harmonieuse. Par exemple, la puce pourrait être chargée d'augmenter ou diminuer la saturation des couleurs dans le champ de vision en fonction de la température du corps ; un effet de ce type étant raisonnablement simple et suffisamment intuitif, on peut supposer qu'il s'intégrera harmonieusement à l'expérience humaine et sera suffisamment peu saillant pour que l'enfant n'ait pas remarqué sa présence en comparant ses expériences avec d'autres personnes. En principe il s'agirait toujours de la représentation symbolique visuellement accessible de données informatiques (ici issues d'un petit programme relié à une sonde-thermomètre), pourtant le fait que cette représentation soit à présent incluse directement dans le champ visuel, plutôt que par le biais d'un support (l'écran) semble produire (ou renforcer) l'intuition que ce symbole doit être différent de, ou quelque chose de plus que, la représentation d'une variable informatique.

Cet exemple est sans doute loin d'être exhaustif, mais je suis persuadé qu'il dit quelque chose de l'importance de la proximité (telle que définie plus haut) qui existe entre l'observateur et l'illusion pour expliquer la force de cette dernière. Plus nous sommes proches de l'illusion à laquelle nous faisons face, plus nous avons de mal à prendre le recul nécessaire à son dépassement. A fortiori, lorsque l'observateur et l'illusion appartiennent au même système, l'opération peut devenir quasiment impossible²².

²² Ici, je fais référence au mal que nous avons à « prendre du recul » sur certains processus cognitifs relativement transparents, la difficulté rencontrée (sans en avoir conscience) à surveiller suffisamment notre activité mentale pour en corriger systématiquement les erreurs.

Alors, « pourrait-on aussi facilement accepter l'illusion de l'utilisateur si l'on était aussi "proches" de l'interface de notre ordinateur que nous le sommes de notre conscience ? » Une fois de plus, je pense pouvoir répondre non, et une fois de plus, poser cette question m'a permis d'envisager un facteur explicatif de la différence apparente de richesse entre l'illusion de l'utilisateur et l'illusion de conscience qui ne recoure pas à l'hypothèse d'une machinerie introspective trompeuse. Par contre, on peut noter que cette notion de proximité de l'illusion ne contredit pas l'importance de la croyance, puisqu'on peut supposer qu'une illusion plus proche de nous est susceptible de produire ou faciliter l'acquisition de croyances plus tenaces à propos d'un état de fait, de la même façon que nous sommes plus influencés dans nos croyances par ce qui nous touche directement (le milieu familial, socio-économique, l'environnement culturel, etc.) que par ce qui nous est étranger.

3.3 Conclusion

Reprenons à présent l'argument de Kammerer et mes critiques à son sujet, afin d'en tirer les conclusions :

1. Si les théories de l'illusion légère sont vraies, alors « quand nous sommes informés du fait que nos représentations introspectives de nos états internes sont en fait schématiques et incomplètes, de sorte que nos états internes ont une nature intrinsèque qui ne nous est pas présentée à travers l'introspection (à savoir une nature physico-fonctionnelle), notre réticence à accepter cette nouvelle information ne devrait pas être substantiellement plus forte et plus résiliente que la réticence dont nous faisons l'expérience quand nous apprenons que n'importe laquelle de nos autres représentations incomplètes "standards" sont réellement incomplètes. » (Kammerer, 2022, p. 11)

Contre cette prémisse, j'ai affirmé qu'il est possible de trouver des exemples d'illusions légères (en l'occurrence, j'ai utilisé un exemple calqué sur celui que donne Kammerer quand il définit l'illusion légère) pour lesquels notre réticence à accepter la nature physico-fonctionnelle sous-jacente à une illusion est très importante, suffisamment *a priori* pour qu'une personne sujette à ce type d'illusion et adhérant aux croyances idoines rejette la révélation du fonctionnement de l'illusion pour conserver ses croyances. Il est donc possible que certaines illusions légères soient particulièrement tenaces, malgré que cette ténacité repose essentiellement, voire entièrement, sur la force de croyances. On peut donc dire que la réticence particulière à accepter une illusion n'est pas forcément un bon indicateur de la nature de cette illusion.

66 • Illusion riche ou légère ?

2. Or « nous sommes plus réticents à accepter la nature physico-fonctionnelle de nos états internes que nous le sommes à accepter que n'importe laquelle de nos autres représentations incomplètes "standards" soit incomplète. » (Kammerer, 2022, p. 12)

J'ai proposé ensuite que même s'il était vrai que l'illusion de la conscience est notablement plus prégnante que n'importe quelle autre illusion – ce qui n'est pas certain – il faudrait néanmoins, pour assurer la validité de l'argument de Kammerer, que la richesse de l'illusion de la conscience soit la cause nécessaire et suffisante de notre réticence à l'accepter, ce qui n'est pas certain non plus. J'ai proposé l'exemple de la proximité comme critère alternatif de la force d'une illusion, critère qui par ailleurs permet d'envisager des cas où l'on montrerait autant de résistance à une illusion légère (dans mon exemple, l'illusion de l'utilisateur) qu'à l'illusion de la conscience.

3. Donc les théories de l'illusion légère sont fausses.

En raison du problème de la première prémisse (dont seule la version faible est à peu près tenable, mais qui ne permet pas de conclure ce que Kammerer en conclut), et de celui de la seconde prémisse (Kammerer présuppose de la facilité du dépassement de l'illusion de l'utilisateur par rapport à l'illusion de la conscience), le passage à cette conclusion est mis en péril.

De plus, il y a un problème dans le passage des prémisses à la conclusion : on ne peut pas conclure du fait que l'illusion phénoménale est significativement plus dure à accepter que les autres, qu'elle soit *impossible* à accepter. Si l'on rajoute à cela le fait qu'il n'est pas certain que l'illusion phénoménale soit significativement plus difficile à accepter que les autres et que si elle l'est, il n'est pas certain que sa ténacité provienne de sa richesse (telle que Kammerer la décrit), on peut déceimment conclure que l'argument de Kammerer contre les théories de l'illusion légère ne tient pas la route.

Par ailleurs, je pense aussi avoir montré dans ce chapitre certaines limites des concepts de pénétrabilité et d'impénétrabilité cognitive dans le cadre qui nous occupe. Je pense effectivement que l'illusion de la conscience est plus difficile à accepter que d'autres illusions, telles que l'illusion de l'utilisateur, mais cette différence peut s'expliquer d'autres manières qui impliquent moins de présupposés que la théorie de Kammerer reposant sur le postulat d'un mécanisme trompeur cognitivement impénétrable (par exemple, par une différence de compétences au sujet du substrat de l'illusion, ou par une différence de proximité entre le sujet et le substrat de l'illusion). Je reviendrai à cette question au chapitre

6.3 afin d'exposer une autre conception que celle de Kammerer, qui me semble plus à même de rendre compte de la richesse apparente de l'illusion de la conscience.

Enfin, j'aimerais encore souligner, à la charge de cet argument, le caractère vague de certains termes utilisés par Kammerer, qui nuisent d'après moi à la qualité de ses observations à propos de la richesse des illusions. Kammerer parle de cas de représentations incomplètes « standards » – mot qu'il place lui-même entre guillemets – sans préciser ce qui définit ces cas et ce qui fait leur particularité par rapport à des cas non-standards (comme celui de la ou des représentation(s) qui donne(nt) lieu à l'illusion de la conscience). Il parle également de « s'habituer » à l'idée de l'illusion, ou encore « d'accepter » celle-ci, mais à quoi font référence ces termes ? Certes, ils sont facilement compréhensibles intuitivement, mais leur utilisation ici devrait renvoyer à une réalité bien particulière, à savoir à l'état (état neuronal, dispositionnel, intentionnel, comportemental, ou même phénoménal...) dans lequel serait une personne qui a été mise en contact avec une affirmation illusionniste et aurait entièrement digéré cette nouvelle information. Cet état mérite d'être décrit car il semble recouvrir une réalité plus large que celle à laquelle Kammerer fait référence.

Chapitre 4 Vers une théorie doxastique de l'illusion de la conscience

4.1 Biais cognitif ou illusion perceptuelle ?

En attaquant les présupposés sur la légèreté de l'illusion de l'utilisateur dans ma critique de la seconde prémisse de Kammerer, je souhaitais souligner la force que peuvent avoir les illusions dont on pourrait penser qu'elles sont si simples qu'elles ne pourraient avoir qu'un effet négligeable sur nous, et dont on ne verrait pas de raison de penser qu'elles sont câblées de façon innée dans notre cerveau. En sommes, des illusions qui ne retiennent pas longtemps notre attention pourraient avoir plus d'effet sur nous que nous le penserions.

L'illusion de l'utilisateur est un bon exemple puisque sous ses atours d'expérience de pensée très « en surface », elle est une illusion qui nous touche tous de façon non consciente et dont on ne réalise pas réflexivement à quel point elle peut nous affecter. En renforçant l'illusion de l'utilisateur (en mettant l'accent sur les formes différentes qu'elle peut prendre selon la personne touchée, ou en imaginant ce qu'elle serait pour quelqu'un qui la vivrait au sein même de son expérience consciente), j'ai tenté de réduire le gouffre qui semble à première vue exister entre elle et l'illusion de la conscience, en espérant que cela puisse nous aider à voir comment une illusion en apparence insurmontable peut en réalité l'être, à condition que l'on en acquière les moyens.

Accepter l'illusion de l'utilisateur n'a nécessité que deux choses dans mon exemple. Premièrement, pour accepter l'illusion du point de vue réflexif, il faut exister dans une culture qui, habituée aux machines, a intégré le fait qu'en dépit de leur enrobage (l'interface), elles restent un système fait de composants électroniques connectés de certaines façons. Deuxièmement, pour dépasser l'illusion du point de vue comportemental, il faut une connaissance un peu plus fine de la machine, de manière à être en mesure d'entrer en rapport avec elle selon une posture du design voire une posture physique, et non selon une posture intentionnelle, source de généralisation et d'illusions. Que nécessiterait le dépassement de l'illusion de la conscience ? S'agirait-il de moyens similaires à ceux évoqués ci-dessus ? Dans la mesure où je pense que la persistance de l'illusion de la conscience a été surestimée, je pense que oui.

70 • Vers une théorie doxastique de l'illusion de la conscience

Malheureusement, nous n'existons pas dans une culture habituée au discours physicaliste sur la conscience, en tout cas pas entièrement. Si l'idée du physicalisme concernant la conscience a gagné du chemin depuis quelques décennies (au moins depuis les années 1990, « décennie du cerveau »), nous subissons encore de toute part la pression d'une culture dualiste. Mais ce fait sera sans doute amené à évoluer à l'avenir. Pensons à cette fin du XIX^e siècle, marquée par un engouement pour l'occultisme et le développement d'une approche scientifique du surnaturel (Bergé, 1990, pp. 57, 68). À cette époque et dans certains milieux sociaux, les tables tournantes et autres démonstrations de médiumnité ont pu faire l'unanimité, convainquant même certains esprits scientifiques, en dépit d'indices et de certains cas notoires qui discréditaient régulièrement ce milieu. En lisant certaines histoires de ce temps²³, on pourrait être tenté d'accuser nos ancêtres d'un manque criant de jugement. Pourtant, c'est aussi cette période qui marqua les débuts de la psychologie en tant que science et, quelques décennies à peine plus tard, ces croyances qui semblaient si fortes, au point de tromper si radicalement les perceptions des hommes, vacillaient à mesure que les théories et techniques d'expérimentation scientifiques s'affinaient. Nous avons depuis tourné le dos à l'éther et à l'ectoplasme.

Malgré l'évolution des mentalités au sujet du physicalisme, le domaine plus spécifique des neurosciences semble encore largement ignoré du grand public. Je suis toujours étonné de voir des personnes affirmer avec conviction que « nous savons si peu au sujet du cerveau ! », avant de soutenir cette affirmation par un cortège de lieux communs depuis longtemps dépassés au sujet de la glande pinéale ou des expériences de mort imminente. À la décharge de ces derniers, il est vrai que, d'une part, les neurosciences sont encore jeunes (comparativement à, disons, la physique), et l'effort de vulgarisation qui devrait permettre de rendre leurs découvertes accessibles à tous est peut-être trop peu développé. D'autre part il s'agit d'une spécialisation assez poussée qui échappera donc toujours en partie à ceux qui n'ont pas déjà une certaine culture dans les domaines connexes.

Néanmoins, je pense que c'est bien au niveau de la croyance qu'il faut chercher l'origine de la force particulière de l'illusion de la conscience (comme d'autres illusions particulièrement tenaces). Ce qui m'amène enfin à considérer ce qu'affirme Kammerer au sujet des théories de l'illusions « ni-riche-ni-légère », c'est-à-dire des théories qui affirment que l'illusion de la conscience n'entre ni dans la catégorie des illusions riches, ni dans celle

²³ Voir, par exemple, le récit de l'expérience menée par William Crookes avec la médium Florence Cook rapporté par Christine Bergé (Bergé, 1990, pp. 81-84, 99-100).

des illusions légères (telles qu'elles sont définies par Kammerer). L'option envisagée par Kammerer et que j'aimerais explorer est celle d'une théorie qui se situerait plutôt du côté de l'illusion légère mais tenterait d'en pallier les limites en proposant des facteurs explicatifs de notre résistance intuitive au physicalisme, tels que « la préexistence de croyances anti-physicalistes tenaces, de nature philosophique, religieuse, etc. » (Kammerer, 2022, p. 15).

Ce type de position implique donc que nous soyons victimes à la fois d'une illusion « *introspective* » et d'une illusion « *doxastique* » (Kammerer, 2022, p. 15) dans nos jugements à propos de la conscience. La première serait légère, c'est-à-dire qu'elle relèverait d'une erreur produite par une inférence fallacieuse provoquée par une représentation schématique et incomplète. Quant à la seconde, selon Kammerer nous devons à nouveau nous poser la question de sa pénétrabilité cognitive (Kammerer, 2022, p. 14). Or, si cette illusion doxastique devait entrer dans la catégorie des illusions légères, nous nous trouverions à nouveau en difficulté pour expliquer la force particulière de l'illusion de la conscience ; et si l'illusion doxastique devait entrer dans la catégorie des illusions riches, alors la théorie qui en résulterait semblerait difficile à distinguer d'une théorie de l'illusion riche (Kammerer, 2022, pp. 14, 15). À ce stade, Kammerer renvoie la balle à d'autres pour ce qui est de développer une théorie crédible qui se déploie dans cet entre-deux. Remarquons déjà, néanmoins, qu'à défaut d'avoir tranché entre l'illusion riche et l'illusion légère, l'existence de cette possibilité déplace déjà le questionnement dans le domaine de la croyance, et non plus de l'introspection ou des représentations, ce qui me renforce dans l'idée qu'il s'agit-là d'une direction pertinente pour explorer plus avant la nature de l'illusion de la conscience.

Dans un autre article intitulé « *Does the explanatory gap rest on a fallacy ?* » (Kammerer, 2019d), Kammerer élabore une argumentation contre des théories qu'il relie au cas évoqué plus haut d'illusion introspective légère accompagnée d'une illusion doxastique riche (Kammerer, 2022, p. 14n13). Dans cet article à propos de la nature du gouffre explicatif entre le phénoménal et le physique, Kammerer affirme qu'il est faux d'interpréter celui-ci comme une erreur de raisonnement (*fallacy*) ou un biais cognitif (*cognitive illusion*), parce que cette stratégie ne serait pas « psychologiquement plausible » (Kammerer, 2019d, pp. 21-22).

Je ne souhaite pas élaborer au sujet des théories que Kammerer rassemble dans cet article sous le terme de « *fallacy accounts* » (c'est notamment la théorie de David Papineau

72 • Vers une théorie doxastique de l'illusion de la conscience

qui est visée) (Kammerer, 2019d, pp. 1-2), et qui envisagent le gouffre explicatif comme une certaine forme de raisonnement fallacieux. Je ne chercherai donc pas à défendre celles-ci, mais plutôt à rebondir sur la critique que Kammerer en fait pour, à nouveau, explorer plus avant l'influence de la croyance dans la formation de l'illusion de la conscience. Je pense, comme Kammerer, que l'illusionnisme de Dennett est un illusionnisme « léger » (Kammerer, 2022, p. 8), et que par conséquent il ne reconnaît pas l'existence d'un gouffre épistémique au sens où Kammerer le définit (c'est-à-dire avant toute chose une intuition persistante et résistante à la réflexion) (Kammerer, 2019a, pp. 55n3,57)²⁴. Je pense également que l'espace qui sépare actuellement notre compréhension du fonctionnement du cerveau et notre compréhension intuitive du vécu phénoménal sera amené à se réduire et même à disparaître à mesure que nos connaissances au sujet de la première augmenteront.

Néanmoins, certaines propriétés des biais cognitifs relevées par Kammerer sont de nature identique à des propriétés qui me semblent définitives de l'illusion de la conscience dans son ensemble²⁵ telle que l'envisage Dennett, par exemple le fait qu'ils proviennent d'un processus psychologique « automatique » et que par conséquent nous sommes prisonniers de notre tendance à reproduire ces erreurs, même si la réflexion attentive peut nous permettre de les surmonter (Kammerer, 2019d, pp. 11-12). De plus, je pense que la critique qu'en fait Kammerer répète certaines erreurs dénoncées par Dennett, il sera donc instructif de se pencher dessus.

Je souhaite néanmoins souligner que ce texte de Kammerer n'est pas censé être une critique d'arguments de Dennett, ni mon commentaire de celui-ci une défense de ce dernier contre les remarques de Kammerer. Donc, l'objectif ne sera pas d'argumenter contre Kammerer que le gouffre explicatif pourrait provenir d'une erreur de raisonnement plutôt que d'une propriété particulière de la machinerie introspective, mais simplement d'utiliser son texte pour appliquer à un cas concret la ligne de pensée dennettienne que je souhaite suivre dans ce travail, afin de montrer une fois de plus comment la croyance peut produire

²⁴ Chalmers catégorise quant à lui le travail de Dennett dans le « matérialisme de type A », c'est-à-dire un matérialisme qui ne reconnaît pas l'existence du gouffre explicatif (Kammerer, 2019a, p. 48).

²⁵ Je parle de « l'illusion de la conscience dans son ensemble » dans le sens où le problème de l'intuition de l'existence d'un gouffre explicatif entre le phénoménal et le physique est un aspect particulier du plus général problème difficile de la conscience. Aussi, il semble raisonnable de penser que l'intuition (trompeuse, par hypothèse, selon les illusionnistes) du gouffre explicatif respecte des mécanismes identiques à ceux d'autres intuitions (trompeuses, également) au sujet de la nature de la conscience (telles que l'intuition qu'il existe un aspect qualitatif de la conscience qui ne soit pas réductible à état physique du cerveau).

des intuitions fausses au sujet de la conscience qui ne soient ni insurmontables, ni facilement surmontables.

L'objectif de Kammerer dans cet article est d'argumenter en faveur d'une position selon laquelle notre introspection nous tromperait systématiquement en nous présentant des états phénoménaux représentés comme possédant des propriétés impossibles (Kammerer, 2019d, p. 8)²⁶, c'est-à-dire une position illusionniste selon laquelle l'intuition du gouffre explicatif ne proviendrait pas d'une erreur de raisonnement consistant à se tromper (de façon systématique et involontaire) sur le contenu de notre introspection, mais d'un raisonnement correct fondé sur une prémisse fautive fournie par la machinerie introspective, prémisse selon laquelle nous entrons dans des états phénoménaux possédant des propriétés impossibles. Kammerer illustre cette distinction par un cas de « *raisonnement valide*, où le raisonnement est fait sur base de fausses prémisses, et où les fausses prémisses sont obtenues par le biais d'un processus perceptuel illusoire et non fiable », afin de soulever les différences qui existent entre une illusion produite par ce type de raisonnement et une illusion produite par un raisonnement fallacieux (Kammerer, 2019d, p. 12). Sa description prend la forme d'une expérience de pensée :

« Supposons qu'une femme, Anne, est sur le point d'entrer dans une pièce. Des gens lui ont dit qu'il n'y a rien dans cette pièce si ce n'est une chaise blanche ; disons qu'elle les croit. Il se trouve qu'ils disent la vérité, et que la pièce est réellement vide, à l'exception d'une chaise blanche. Cependant, lorsqu'Anne entre dans la pièce, elle se trouve avoir une hallucination d'un chat noir assis sur la chaise blanche. Supposons également que, sur base de cette expérience visuelle non véridique, elle forme un jugement perceptuel du type : "il y a un chat noir dans cette pièce". » (Kammerer, 2019d, p. 12)

Kammerer conclut de cette situation qu'avec « un peu de raisonnement », Anne observera une contradiction entre sa précédente croyance (selon laquelle la pièce est vide à l'exception d'une chaise blanche), croyance qui se trouvait être vraie, et son expérience personnelle (expérience visuelle d'un chat noir dans la pièce), ce qui l'amènera possiblement à remplacer la première par une nouvelle croyance (selon laquelle il y a un chat noir dans la pièce qui devait être vide) qui, elle, est fautive (Kammerer, 2019d, pp. 12-13).

Anne a une expérience visuelle illusoire, mais c'est une illusion perceptuelle (« *perceptual-like* ») et non cognitive, « c'est son système perceptuel (visuel) qui

²⁶ Cette position correspond bien à ce que Kammerer appelle par ailleurs l'illusionnisme riche, que j'ai abordé plus haut.

fonctionne mal ici, et l'amène à former une croyance fausse » (Kammerer, 2019d, p. 13). Ce que Kammerer entend par là, c'est que nous serions face à une illusion qui ne se produit pas « au niveau du raisonnement (la manipulation de représentations conceptuelles) » comme les biais cognitifs, mais à un « niveau plus bas » (Kammerer, 2019d, p. 6). Contrairement aux biais cognitifs qui proviendraient d'une erreur d'inférence, les illusions perceptuelles n'impliqueraient pas d'erreur de manipulation de représentations conceptuelles et, si elles devaient reposer sur un mécanisme inférentiel, ce ne seraient que des inférences « subpersonnelles » dont les prémisses seraient « encodées dans un format non-conceptuel, et [auraient] un contenu non-conceptuel » (Kammerer, 2019d, p. 7).

4.2 L'erreur de Kammerer

Dans le cas des biais cognitifs, « l'élément illusoire survient au niveau cognitif », alors qu'il « survient à un niveau perceptuel » dans le cas de l'hallucination d'Anne décrite plus haut. Cette affirmation pose une question qui me semble importante : comment distinguer clairement le « niveau cognitif » du « niveau perceptuel » ? Si le premier se situe « au niveau du raisonnement », pourquoi nous est-il si difficile (voire impossible) de modifier les heuristiques responsables de ceux-ci ? Au contraire nous devrions, si c'était le cas, pouvoir apprendre simplement à raisonner correctement, plutôt que d'apprendre à reconnaître dans nos raisonnements intuitifs certaines erreurs systématiques. Or c'est plutôt de cette façon que nous apprenons à déjouer les biais cognitifs. Et peut-on être sûr que les illusions perceptuelles se produisent à un niveau « subpersonnel » et « non-conceptuel » ? Est-il vraiment juste d'appeler une perception « illusoire » tant que celle-ci n'a pas été affiliée à un concept qui, lui, amène à former un jugement faux à propos d'un état de fait réel ?

Par exemple, si j'ai l'impression d'entendre quelqu'un m'appeler (il me semble qu'il s'agit d'un type assez commun et généralement bénin d'hallucination auditive), peut-on dire que cette perception est illusoire avant que mon cerveau ait interprété ce son comme étant la voix de quelqu'un qui m'appelle ? Même s'il s'agissait d'une hallucination pure qui ne trouve sa source dans aucun *stimulus* extérieur (par exemple une voix que j'aurais cru reconnaître) mais ne soit qu'une production spontanée du cerveau, on ne pourrait affirmer son caractère illusoire qu'à partir du moment où nous serions conscients d'avoir une hallucination et donc seulement dès lors que cette production spontanée aura pris la forme d'une expérience consciente et donc d'un contenu chargé de concepts (dans

l'exemple précédent, je n'ai jamais conscience d'une sorte de sensation sonore pure – qui serait notre illusion « subpersonnelle » et « non-conceptuelle » – que j'interpréteraï ensuite comme étant un appel ; lorsque je deviens conscient du *stimulus*, j'en prends conscience *comme étant* un appel).

Pour Kammerer, l'information « obtenue par le moyen d'un appareil perceptuel dysfonctionnel » fait office de « prémisse fausse » à un raisonnement dont la conclusion logique s'en trouve elle aussi faussée. Or il semble que ce qui rend fausse cette prémisse dans le cas qui nous occupe, c'est précisément la conclusion qui lui est apposée. La perception d'Anne, qu'il s'agisse d'une perception tronquée par une illusion d'optique ou un élément extérieur (une drogue, ou un dispositif qui stimule son cortex visuel (Kammerer, 2019d, p. 12n12)), ou d'une véritable hallucination, ne peut pas être fausse *en soi*, elle n'est fausse que parce qu'Anne conclut de celle-ci une information fausse à propos du monde. Ce que Kammerer affirme être la prémisse fausse d'un raisonnement valide est déjà la conclusion fausse d'une inférence basée sur des données dont une partie au moins est déjà conceptuelle : on ne peut pas dire d'Anne qu'elle a vu un chat noir avant qu'elle ait elle-même apposé à son expérience visuelle le concept de chat noir.

J'aurai l'occasion de revenir dans la dernière partie de ce travail sur les questions psychologiques liées aux illusions et hallucinations²⁷. Pour l'instant, je souhaiterais me contenter d'appuyer le point de vue esquissé ci-dessus par un argument logique, puisque c'est avant tout de logique qu'il est question ici (on parle bien de conditions de vérité de prémisses et de validité de raisonnement) malgré le scrupule de Kammerer de rester proche du « psychologiquement plausible ».

Dans l'expérience de pensée, Anne est mise face à deux informations à propos de la pièce, qui donnent lieu à deux croyances à son sujet. Ce sont donc quatre représentations de Anne qui entrent en jeu :

(p) *L'information selon laquelle la pièce est vide*²⁸ ;

(p') *La croyance réflexive*²⁹ *selon laquelle la pièce est vide* ;

²⁷ La différence entre illusion et hallucination est un sujet qui sera également abordé en temps voulu.

²⁸ Pour alléger le texte, je ne ferai plus référence à une « pièce vide à l'exception d'une chaise blanche », mais simplement à « la pièce vide », étant donné que la présence de la chaise blanche n'a aucune importance dans le raisonnement.

²⁹ J'utilise ici le terme « croyance réflexive » pour souligner qu'il s'agit d'une croyance accessible à la conscience réflexive, contrairement à d'autres types imaginables de croyances telles que les « *alief* » ou autres représentations non-conscientes.

76 • Vers une théorie doxastique de l'illusion de la conscience

(q) Le jugement d'après lequel il y a un chat dans la pièce ;

(q') La croyance réflexive selon laquelle il y a un chat dans la pièce.

D'après Kammerer, dans cette histoire, Anne ne fait aucune erreur de raisonnement ; voyons donc quelles sont ses différentes inférences : Anne apprend p et en tire p' , ensuite Anne fait l'expérience de q , et elle observe que q entre en contradiction avec p' puisque q est contradictoire avec p^{30} (la pièce ne peut pas à la fois être vide et contenir un chat), donc si q est vrai p' doit être faux, or q est vrai (Anne semble avoir de bonnes raisons de le penser), donc Anne remplace p' par la nouvelle croyance q' tirée de q .

S'il serait correct d'affirmer la véracité de la contradiction entre p et q et entre p' et q' , il n'est pas juste d'affirmer la contradiction entre q et p' ($q \equiv \neg p'$), or c'est ce qu'Anne semble faire ici (si l'on en croit Kammerer). En effet, il n'est pas impossible qu'Anne voie un chat mais décide de croire néanmoins que la pièce était vide ($q \wedge p'$) : elle pourrait, par exemple, choisir de faire confiance aux personnes qui l'ont informée que la pièce était vide plutôt qu'à sa perception qu'elle sait trompeuse (parce qu'elle sait qu'elle souffre d'hallucinations récurrentes, parce qu'elle se souvient avoir pris une substance hallucinogène, ou plus simplement parce qu'elle sait qu'il peut arriver à n'importe qui d'avoir une hallucination dans de rares cas).

Pour que le raisonnement d'Anne puisse être valide, il faudrait qu'elle commence par produire une croyance q' sur base de son observation q , pour ensuite comparer cette croyance avec la croyance précédente p' , considérer le fait qu'elles sont en contradiction et dès lors faire un choix entre q' et p' au terme duquel, si elle opte pour la croyance q' , on pourra dire qu'elle a effectivement atteint une conclusion fautive (accepter une croyance fautive) sur base d'un raisonnement valide (consistant à tirer une croyance d'un état perceptuel, puis à comparer celle-ci à une précédente croyance).

Cette petite erreur est bénigne pour le raisonnement de Kammerer, puisqu'on pourrait aussi bien imaginer qu'Anne forme d'abord une croyance q' qu'elle compare ensuite avec sa précédente croyance p' pour en conclure à la vérité de la première étant donné la nouvelle information q , sans que cela change fondamentalement l'argument de Kammerer ou le déroulement de l'expérience de pensée. Néanmoins, elle permet

³⁰ « *A little bit of reasoning leads her to consider a contradiction between her previous belief about the room (which happened to be true) and the content of her new judgment that there is a black cat.* », littéralement « Un peu de raisonnement l'amène à considérer une contradiction entre sa précédente croyance à propos de la pièce (qui se trouvait être vraie) et le contenu de son nouveau jugement qu'il y a un chat noir. » (Kammerer, 2019b, p. 13)

d'entrevoir l'erreur (la « *fallacy* ») que commet Anne en réalité et que Kammerer n'a pas vue, qui se trouve aussi être l'erreur qui empêche aux réalistes de comprendre le point de vue illusionniste ; une erreur qui se situe dans le passage d'une expérience de quelque chose à l'affirmation de l'existence de ce quelque chose.

En effet, l'erreur de raisonnement décrite plus haut provient de ce que Anne a inféré $\neg p'$ sur la base de q , or si l'on devait donner sens à cette erreur en la dépliant, on obtiendrait un raisonnement de ce type :

$$q$$

$$\text{Or } q \rightarrow \neg p$$

$$\text{Et } \neg p \rightarrow \neg p'$$

$$\text{Donc } \neg p'$$

Ou, pour le formuler en français, Anne doit penser quelque chose comme : « J'ai vu un chat dans la pièce, or s'il y a un chat dans la pièce, alors la pièce n'est pas vide, et si la pièce n'est pas vide alors ma croyance selon laquelle la pièce est vide est fausse, donc ma croyance selon laquelle la pièce est vide est fausse ». Résumé comme cela, le raisonnement semble effectivement valide, et il semble bien que la fausseté de la prémisse (q) soit transmise à la conclusion ($\neg p'$).

D'après Kammerer, la fausseté de q implique la fausseté de la conclusion d'Anne selon laquelle p doit être faux. Mais quel sens y a-t-il à dire que q est faux ? Cela ne signifie pas qu'Anne n'a en fait vu aucun chat. D'une part cela reviendrait vraiment à « nier les données », d'autre part on voit bien par ailleurs qu'Anne est dans un état dispositionnel qui laisse à penser qu'elle a bien vu un chat : elle forme une croyance allant en ce sens, et rejette d'autres croyances incompatibles avec ce fait.

Je pense effectivement que le problème se situe dans la prémisse, mais pas dans le sens entendu par Kammerer. Si l'on s'interroge sur la condition de vérité de la prémisse q , on voit qu'elle comporte un présupposé important : lorsqu'Anne reçoit l'information qu'il y a un chat dans la pièce et produit un jugement à ce sujet, il semble qu'elle en infère immédiatement qu'il y a réellement un chat dans la pièce. Or rien n'est moins sûr !

En réalité, Kammerer n'a pas manqué cette importante nuance, puisqu'il affirme (en note de bas de page) que le raisonnement d'Anne comporte bien une étape supplémentaire qui consiste pour elle à « approuver » le produit de sa perception comme étant conforme à la réalité, étape qu'il ne considère pas comme une erreur de raisonnement :

78 • Vers une théorie doxastique de l'illusion de la conscience

« nous ne commettons pas d'erreur de raisonnement quand nous approuvons le contenu de notre perception (*du moment que nous n'avons pas de raisons de douter de nos perceptions*) [mes italiques] » (Kammerer, 2019d, p. 13n13)³¹.

Si l'on prend Kammerer au pied de la lettre, cette affirmation suggère que le raisonnement d'Anne deviendrait une erreur à condition que nous ayons une raison de douter de nos perceptions, comme si la connaissance du fait que nos perceptions *pourraient* être fausses était ce qui rendait le raisonnement faux.

Il est vrai que si nous sommes certains que nous ne pouvons pas faire confiance à nos perceptions, il serait fallacieux d'approuver leur contenu (c'est sans doute ce que Kammerer voulait dire), mais cela n'implique pas qu'il *ne serait pas fallacieux* d'approuver le contenu de nos perceptions lorsque nous n'avons pas de raison de douter de celles-ci, puisque le fait que nous n'ayons pas de raison de douter de nos perceptions n'implique pas qu'elles soient entièrement fiables (par exemple, même si Anne n'avait aucune raison de s'inquiéter d'un quelconque problème d'hallucinations, il se pourrait que quelqu'un lui ai administré à son insu une drogue hallucinogène). Je pense même qu'il est raisonnable de partir du principe inverse : par précaution, et étant donné ce que nous savons sur la perception (qu'il s'agit d'une reconstruction du réel produite par le cerveau, avec ce que cela implique de déformations et d'approximations malgré une haute-fidélité), il est raisonnable de ne pas lui donner une confiance totale en toute occasion. Par exemple, il est raisonnable de recevoir avec scepticisme une perception étrange lorsque l'on est dans un état de fatigue avancé. Il est tout aussi raisonnable de douter d'une perception dont le contenu échappe à ce qui est généralement admis comme crédible ou réaliste (par exemple, un fantôme), y compris si nous étions parfaitement alertes lorsque nous avons eu celle-ci.

Et c'est bien là ce que je souhaite défendre : l'erreur de raisonnement dont Anne est victime (une victime inconsciente, apparemment, mais il n'est pas nécessaire qu'elle le soit ; tout comme le fait d'être conscient ou non de l'existence d'un biais cognitif ne change rien à son existence ni à son caractère fallacieux) consiste en une seule inférence fallacieuse : l'inférence par laquelle le jugement visuel d'Anne est représenté comme portant sur quelque chose de réel. En effet, l'erreur que doivent souligner les illusionnistes ne tient pas à la fausseté des représentations (parler de vérité ou de fausseté dans leur cas n'a pas

³¹ Pour être exact, Kammerer dit « *arguably, we commit no mistake when we endorse the content of our perception* » (Kammerer, 2019b, p. 13n13), il ne considère donc pas cette affirmation comme une certitude, malgré la confiance qu'il place dedans...

vraiment de sens), mais plutôt à l'idée fautive selon laquelle si j'ai une représentation de quelque chose, ce quelque chose doit exister *quelque part*.

Pour rendre visible cette erreur, il est utile de distinguer deux sens différents des propositions q et p en fonction de leur contenu intentionnel : p et q peuvent renvoyer soit au contenu de notre expérience en tant que telle, soit au contenu de notre expérience en tant que celui-ci représente un état de fait réel. Il me semble important de distinguer ces deux sens, étant donné que l'un peut exister sans l'autre : toutes nos représentations ne portent pas la marque du réel. Nos capacités à imaginer, rêver, tisser des mondes fictionnels, sont autant de sources d'illusions bénignes qui ne portent pas directement sur un état de fait réel. Comme le soulignait Ryle dans *The Concept of Mind*, nous ne voulons pas dire la même chose quand nous affirmons voir quelque chose (percevoir quelque chose dans notre champ de vision) et quand nous affirmons « voir » cette même chose (le visualiser « dans notre tête », l'imaginer avec plus ou moins de détails). Et confondre ces deux sens amène à des confusions dans le langage et donc dans le raisonnement (Ryle, 1978, p. 234sv). Dans le cas qui nous occupe, q n'est faux que dès lors qu'Anne pense qu'il représente en quelque sorte un état de fait réel ; comme nous l'avons vu plus haut, q ne serait pas faux s'il ne portait que sur le contenu de l'expérience. Anne a raison de croire qu'il y a un chat dans la pièce telle qu'elle se l'est représentée, puisqu'il y en avait un (on pourrait dire : le contenu de l'expérience d'Anne contenait effectivement le concept de chat), mais elle a tort de croire que la pièce telle qu'elle se la représente possède une quelconque réalité (soit qu'elle identifie sa représentation à un objet réel existant dans le monde, soit qu'elle infère de sa représentation l'existence – dans le monde ou quelque part dans son Théâtre Cartésien – du contenu de celle-ci).

Sur base de son hallucination, c'est en réalité deux croyances que forment Anne, qui ne pense probablement pas à dissocier celles-ci : la croyance selon laquelle il y a un chat noir dans la pièce qu'on lui avait dite vide, et la croyance selon laquelle elle a vu un chat noir dans cette pièce. La première découle vraisemblablement de la seconde (pour affirmer l'existence de quelque chose, il faut d'abord se représenter cette chose), mais n'est pas identique à elle, et l'une peut parfaitement exister sans l'autre. Or seule la première croyance est problématique : comme nous l'avons vu, Anne a raison de croire qu'elle a vu un chat noir, mais pas que ce chat noir existe. Par conséquent, l'inférence qui nous intéresse ne donne une conclusion fautive que dans un des deux cas, alors que les deux découlent de la même prémisse (la représentation d'un chat noir dans la pièce), et celle-ci ne peut donc

80 • Vers une théorie doxastique de l'illusion de la conscience

pas être fausse. Si le problème venait de ce que notre introspection représente faussement (*misrepresents* (Kammerer, 2019d, p. 8)) nos états subjectifs de manière systématique, alors les conclusions que l'on tire à leur sujet devraient être constantes, or ce n'est pas le cas. Ceci suggère que le problème ne vient pas de la prémisse mais du passage (non-fondé) d'une affirmation à une affirmation plus forte (celle qu'il existe réellement un chat noir).

Revenons à présent au problème plus général de l'intuition du gouffre explicatif. D'après Kammerer, tout comme la croyance fausse d'Anne, l'intuition du gouffre explicatif trouverait sa source dans une illusion formée à un « niveau inférieur », « subpersonnel », indépendant des représentations conceptuelles et du raisonnement (Kammerer, 2019d, p. 6) et serait résistante à une réflexion ultérieure (Kammerer, 2019d, p. 14). S'agissant d'une « expérience visuelle de bleu », il rencontre l'intuition du gouffre explicatif lorsqu'il tente de rapprocher cette expérience d'une « activité électro-chimique ayant lieu dans [son] cortex visuel », intuition qui lui apparaît d'autant plus forte qu'il « médite sur sa subjectivité, son aspect qualitatif, le fait qu'elle est directement expérimentée, etc. » (Kammerer, 2019d, p. 15). Kammerer en conclut que « réfléchir avec attention » à une expérience produit l'intuition (voire en serait la cause unique) que cette expérience est irréductible à un processus physique.

Par conséquent, bien qu'une « réflexion supplémentaire » (Kammerer, 2019d, p. 15) puisse à nouveau me convaincre de la vérité du physicalisme, mon raisonnement lorsque je me concentre sur une expérience me mène irrésistiblement à conclure à l'existence du gouffre explicatif, prouvant ainsi que cette intuition n'est pas un raisonnement fallacieux (puisque si elle l'était, une réflexion attentive devrait permettre de déjouer l'erreur). Mais pourquoi ne pourrions-nous pas nous passer de la première « réflexion attentive » dont parle Kammerer pour adopter directement la « réflexion supplémentaire » qui nous amène à réhabiliter l'interprétation physicaliste ? Ou, à défaut, pourquoi ne pourrait-on pas considérer que la « réflexion supplémentaire » fait partie intégrante du processus qui consiste à confirmer (ou non) le fruit de notre perception ?³²

Outre le fait qu'il me semble arbitraire de différencier la « réflexion attentive » menant à l'erreur de la « réflexion supplémentaire » nous ramenant au physicalisme, je ne comprends pas pourquoi cette réflexion, qui serait supposée déjouer une erreur de

³² Comme le suggère Ryle : « La confirmation n'est pas un processus consistant à fonder une superstructure de devinettes sur une base de certitudes mais un processus qui consiste à confirmer ou à rendre une chose certaine. » « [Une certitude] est le fruit d'un travail, non l'objet d'une révélation. » (Ryle, 1978, p. 226)

raisonnement s'il y en avait une, ne permettrait pas de questionner la prémisse fautive du raisonnement, qui est à nouveau prise pour argent comptant dans cet argument. De plus, pour Kammerer, la « réflexion attentive » fait apparaître ou renforce l'intuition problématique au sujet de l'expérience (aspect qualitatif, etc.), ainsi que la perplexité face à la réalité du physicalisme (l'expérience n'est rien de plus que le résultat d'un processus cérébral), or c'est précisément ces intuitions (qui apparaissent donc bien au niveau du raisonnement) qui posent problème, et pas l'expérience elle-même. La « méditation » au sujet des propriétés phénoménales dont parle Kammerer (Kammerer, 2019d, p. 15), qui est censée supporter l'intuition du gouffre explicatif ne pourrait-elle pas également nous amener à douter de ces propriétés ?

Je pense que Kammerer décrit bien dans son texte le parcours de l'esprit perplexe face à sa propre expérience, parcours par lequel une question apparemment insoluble (comment cette expérience peut-elle se réduire à un processus chimique produit par un organe ?) amène l'introspection à buter contre des concepts hautement problématiques. Mais l'introspection amène à prendre le problème dans le mauvais sens : ce n'est pas l'expérience qui donnera réponse à notre question mais le processus qui la produit, car celle-ci est plus pauvre que celui-là (l'expérience est une représentation schématique d'un processus cérébral complexe).

Au contraire, se focaliser sur l'expérience nous amène à apposer à nos expériences des propriétés dont on ne peut rien dire qui ne soit pas vague (des propriétés telles que l'aspect qualitatif appellent une réflexion, qui ne permet pas de statuer clairement sur leur nature) ou inutile (la réflexion sur les propriétés de l'expérience ne semble pas apporter de réponse à la question de base, à savoir comment cette expérience peut être produite par le cerveau). En fait, tout ce que permet une telle réflexion, c'est de reproduire l'erreur décrite plus-haut : produire un concept de « bleu mental » qui soit un artefact de la confusion entre le réel et la représentation du réel.

En effet, puisqu'il est de la nature de cette représentation d'être une schématisation, une caricature, du processus physique, il est normal que l'un et l'autre ne soit pas dérivables *a priori*. Une « conversion » est nécessaire pour passer de l'un à l'autre, il faut pouvoir comprendre la relation qui existe entre les deux, ce qui demande un raisonnement supplémentaire, alimenté par des connaissances scientifiques. Sans cet effort particulier, c'est-à-dire si l'on ne base notre raisonnement que sur notre propre expérience, il est normal que nous ne trouvions aucune solution au problème.

82 • Vers une théorie doxastique de l'illusion de la conscience

Nous avons accès au contenu de nos expériences, mais nous n'avons pas accès à la machinerie interne qui nous fournit ces contenus. Il est impossible de se représenter précisément ce que serait « l'activité électrochimique dans le cortex visuel » nécessaire à la production d'une expérience particulière de bleu dans une situation donnée (à moins, peut-être, d'être un chercheur en neurosciences particulièrement informé et imaginatif – assurément, Mary doit en être capable). Il faut donc s'attendre à ce qu'il soit difficile de se représenter quelque chose comme identique à une telle activité, si le concept que nous avons de celle-ci est flou et, qui plus est, perçu intuitivement comme particulièrement complexe. Par conséquent, le fait qu'il ne semble pas y avoir de « réponse satisfaisante » (Kammerer, 2019d, p. 3) *a priori* à la question du rapport entre activité cérébrale et expérience, et donc le fait que leur identité soit contre-intuitive, ne devrait pas nous inquiéter.

Le gouffre explicatif représenterait un véritable problème si, étant capable de se représenter clairement quel devrait être le résultat d'un processus cérébral quelconque, et de définir clairement les limites de « l'effet que cela fait » de faire l'expérience qui en découle, la relation existante entre les deux nous laissait toujours perplexes. Or ni l'une ni l'autre de ces conditions ne semble remplie. Le gouffre ressemble donc plus à une « pompe à intuition », c'est-à-dire un argument en faveur d'une intuition naïve basé sur des concepts vagues ou mal définis.

Au contraire, une fois bien acceptés la différence entre réel et représentation du réel, et le fait que l'activité cérébrale est de l'ordre du premier tandis que l'expérience consciente est de l'ordre de la seconde, le problème se dissipera facilement. Personne n'aurait idée de postuler qu'il existe un gouffre entre l'interface colorée et interactive d'un programme informatique et le code qui est censé rendre celle-ci possible, parce qu'il est largement accepté et compris que l'ordinateur traite les commandes présentes dans le code et convertit toutes les opérations en informations visuelles rendues sur une matrice de pixels. De la même façon, il ne devrait pas sembler impossible que l'expérience soit le fruit d'un processus physique à condition de comprendre que les résultats de ce processus qui sont accessibles consciemment sont encodés dans un format prévu pour être utilisable par l'agent, qui n'a besoin que d'informations schématiques au sujet de lui-même et de son environnement. En réifiant l'expérience phénoménale, l'introspection nous pousse à chercher une identité ontologique entre notre expérience et ce qui se passe dans « la boîte noire », alors que l'identité qui existe en réalité entre les deux est fonctionnelle.

4.3 Conclusion

Il semblerait que la discussion au sujet de la nature du gouffre explicatif menace de nous précipiter vers les mêmes problèmes que rencontrent les discussions au sujet des *qualia*. Cela devrait, en soi, nous pousser à aborder cette thématique avec prudence, en prenant la peine de délimiter clairement les concepts en jeu et de les approcher avec une méthode neutre et débarrassée des présupposés de la méthode traditionnelle des phénoménologues. En d'autres termes, il est nécessaire de bannir les explorations introspectives *a priori*, qui embrouillent le débat en focalisant l'attention sur de faux problèmes provenant de concepts mal définis.

J'ai tenté de montrer comment un argument partant pourtant d'un postulat illusionniste pouvait retomber dans les travers du phénoménalisme et buter contre le même type de problème que rencontrent les qualophiles (la confusion entre réalité et représentation de la réalité). En ajoutant à cela les remarques que j'ai avancées auparavant au sujet de la position illusionniste « riche », j'espère avoir offert un argument satisfaisant pour montrer que considérer l'illusion de la conscience comme advenant au niveau de la croyance – c'est-à-dire adopter ce que j'ai appelé dans le titre de ce chapitre une « théorie doxastique de l'illusion de la conscience » – était une meilleure voie pour répondre au problème de l'illusion.

Je souhaite à présent conclure ce chapitre et cette deuxième partie en reprenant ce qui a été exposé jusqu'ici, ainsi qu'en introduisant la troisième et dernière partie de ce travail : dans la première partie de ce travail, j'ai passé en revue les différents concepts proposés par Daniel Dennett afin de montrer l'articulation de ceux-ci au sein d'un projet illusionniste (chap. 1 et 2). Ensuite, j'ai approfondi cet aperçu superficiel en abordant une question débattue au sein de l'illusionnisme au travers des travaux de Kammerer (chap. 3 et 4). J'ai conclu de cette discussion du problème de la « richesse » de l'illusion de la conscience que l'approche dennettienne soulignant l'importance de la croyance dans la prégnance de l'illusion me semblait la plus à même de répondre au problème de l'illusion.

Mes commentaires du travail de Kammerer étaient destinés à défendre l'hypothèse que l'illusion de la conscience est le résultat d'un processus systématique et probablement en partie câblé dans la machinerie cérébrale, certes, mais néanmoins altérable par une modification des croyances.

84 • Vers une théorie doxastique de l'illusion de la conscience

L'illusion de la conscience peut être surmontée de la même façon qu'un biais cognitif en acquérant des concepts qui permettent de déjouer de façon systématique l'inférence fallacieuse qui pousse le cerveau à réifier l'expérience consciente, à la place des concepts réalistes qui pousseraient au contraire à ratifier cette opération. Je pense que les intuitions amenées par ces concepts fallacieux (comme l'intuition du gouffre explicatif) perdront en force à mesure que les nouveaux concepts s'ancrent de façon définitive dans les schémas de pensée. Quels sont ces nouveaux concepts ? Les concepts de Dennett rappelés dans le premier chapitre en font partie.

Dans la troisième partie de ce travail, je souhaite montrer comment ces concepts peuvent être déployés pour rendre compte d'un phénomène particulier. Le moment est donc arrivé d'introduire le thème qui donne son titre à ce travail : l'expérience psychédélique. L'objectif du chapitre suivant consistera à définir celle-ci. Pour cela, je définirai d'abord, dans le chapitre 5.1, ce qu'est un état modifié de conscience (abrégé en « EMC » dans la suite de ce travail³³). Il y sera d'abord précisé en quoi un état modifié diffère d'un état habituel de conscience. Cette définition sera étayée par un aperçu d'un dispositif permettant de travailler en laboratoire sur les états modifiés de conscience, qui permettra aussi de dessiner un premier aperçu de leur phénoménologie. Le chapitre 5.2 abordera ensuite l'état particulier dont il sera question dans ce travail. L'expérience psychédélique y sera définie, puis contextualisée au travers d'un historique succinct des « recherches psychédéliques »³⁴, avant que sa phénoménologie particulière ne soit décrite.

³³ Traduction de l'acronyme répandu « ASC » (pour *Altered States of Consciousness*).

³⁴ Expression abrégée désignant les recherches au sujet des substances psychédéliques.

Partie III
HÉTÉROPHÉNOMÉNOLOGIE DE
L'EXPÉRIENCE PSYCHÉDÉLIQUE

Chapitre 5 Observer la conscience altérée

5.1 États modifiés de conscience

Définir ce qu'est un état modifié de conscience n'est pas évident, principalement parce que ce phénomène se définit par rapport à celui de la conscience, qui pose lui-même problème. Il est néanmoins nécessaire d'entamer cette dernière partie par une définition approximative de ce type de phénomène, dont fait partie l'expérience psychédélique. Celle-ci sera donnée en deux temps : premièrement une définition générique de la conscience et des états modifiés de celle-ci, et deuxièmement une exposition de facteurs constitutifs d'un état modifié de conscience mis en avant dans certains travaux académiques afin de repérer et quantifier les états modifiés de conscience dans le cadre d'études.

La conscience

Dans un article consacré à l'histoire de l'étude des EMC, Garcia-Romeu & Tart définissent la conscience de façon générique comme « l'appréhension (*awareness*) et l'expérience subjective de phénomènes internes et externes » incluant entre autres « les sensations internes, perceptions, pensées, émotions, et le sens du Soi (« *sense of self* »), ainsi que la perception de toutes classes d'objets externes, événements et autres *stimuli* » (Garcia-Romeu & Tart, 2013, p. 123).

La question de savoir comment le cerveau est capable de produire une telle expérience subjective est laissée en suspens par l'article, et est souvent considérée comme plus difficile que la description même des phénomènes et la recherche de leurs causes. Néanmoins, dans ce travail, j'ai choisi d'exposer la pensée de Daniel Dennett, qui propose pour sa part une réponse à cette question, que l'on peut ranger aux côtés des approches illusionnistes du problème de la conscience, et qui pourrait se résumer comme suit :

« Il n'existe pas de "flux de conscience" unique, définitif, parce qu'il n'y a pas de Quartier Général central, pas de Théâtre Cartésien "où tout se rejoint" pour le bénéfice d'un Signifieur Central. Au lieu d'un tel flux unique (aussi vaste soit-il), il existe de multiples canaux dans lesquels des circuits spécialisés essaient, dans des pandémoniums parallèles, d'accomplir leurs diverses tâches, en créant des versions multiples au fur et à mesure qu'ils opèrent. La plupart de ces versions fragmentaires de "récit" jouent un rôle éphémère dans la modulation de l'activité du moment, mais certaines se trouvent promues à un rôle fonctionnel nouveau, en

se suivant rapidement, à travers l'activité d'une machine virtuelle existant dans le cerveau. Le caractère sériel de cette machine (son caractère "à la von Neumann") n'est pas un trait d'organisation "câblée" ; c'est plutôt le résultat d'une succession de coalitions de ces divers spécialistes. » (Dennett, 1993, p. 318)

En d'autres termes, ce qui est habituellement perçu comme une conscience n'est en réalité que le produit de l'activité spontanée de l'ensemble des « circuits spécialisés » de notre cerveau (produits de l'évolution, mais dont l'organisation est en partie innée, et en partie produite par des « microhabitudes de pensée » permises par la culture humaine et le comportement auto-exploratoire des humains), dont les « Versions Multiples » (les produits de ces circuits, servant à moduler l'activité) peuvent produire en se succédant l'illusion d'un flux sériel (Dennett, 1993, p. 318). Ce flux est un modèle simplifié du fonctionnement interne du cerveau de l'agent, facilitant le contrôle du corps par celui-ci et produisant un Soi agissant comme un Centre de Gravité Narrative, un point (imaginaire) autour duquel gravite les Versions Multiples composant la machine virtuelle du cerveau.

Le Soi et son expérience du monde sont donc des illusions, des abstractions produites par le cerveau et permises par l'ajustement de celui-ci au cours de l'évolution. Par conséquent, pour les illusionnistes, le Problème Difficile de la conscience n'existe pas, ou plutôt il se fond dans les « problèmes faciles » et ne requiert pas d'explications particulières en termes de subjectivité, de propriétés intrinsèques, etc. Le mystère de la conscience est remis à la portée de la science, le physicalisme sauvé, et il ne reste aux philosophes qu'à s'en remettre aux chercheurs compétents pour déterminer les faits sur lesquels bâtir une théorie satisfaisante de la conscience. Accepter une définition minimale et peu ou pas controversée de la conscience est un premier pas dans ce sens, afin de ne pas surcharger celle-ci de présupposés problématiques et s'enliser dans des discussions pointilleuses, et ainsi rester proche des faits et d'une interprétation parcimonieuse de ceux-ci.

Les états modifiés de conscience

S'il existe des états modifiés de conscience, cela signifie qu'il existe plusieurs états, plusieurs formes, que peut prendre la conscience. Garcia-Romeu & Tart parlent d'un « spectre » d'états de conscience, chacun étant un état « discret » défini comme un « *pattern* ou une configuration unique et dynamique de structures psychologiques » (Garcia-Romeu & Tart, 2013, p. 123). Comme nous le verrons plus tard (chap. 6.3), ces structures sont identifiables à des réseaux neuronaux dont le fonctionnement peut varier, produisant une gamme d'états de conscience.

On parlera d'état *modifié* de conscience lorsque l'état usuel de la conscience identifié, par convention, à l'état d'éveil ordinaire – tel qu'il peut être défini statistiquement notamment par le recours à des techniques d'imagerie cérébrale mesurant l'activité cérébrale chez des sujets sains, en état de repos ou en activité – est affaibli ou modifié (par l'effet d'une substance, d'une anomalie cérébrale, d'autosuggestion, ou autre) (Garcia-Romeu & Tart, 2013, p. 123). Citant Adolph Dittrich, Studerus *et al.* définissent ainsi le phénomène comme une « déviation de l'expérience subjective ou du fonctionnement psychologique d'un individu normal de son état habituel de conscience éveillée » (Studerus, Gamma, & Vollenweider, 2010, p. 1), tout en soulignant sa différence avec les « maladies psychiatriques » – les EMC « durent peu de temps » et sont souvent (mais pas toujours) « auto-induites » – même si leur observation a souvent permis l'élaboration « d'hypothèses pour la recherche psychiatrique » (Studerus, Gamma, & Vollenweider, 2010, p. 1)³⁵.

De nombreux EMC ont été identifiés par l'homme, qui a semble-t-il toujours eu coutume d'explorer, et parfois même de rechercher, ceux-ci. En effet, leur influence sur la culture humaine s'étendrait autour du monde et à travers les siècles. On en retrouve les traces en Grèce Antique (dans le culte dionysiaque et les Mystères d'Eleusis), chez les peuples indigènes d'Amérique et de Sibérie (dont les pratiques chamaniques impliquent l'utilisation de drogues hallucinogènes), en Asie (où les pratiques Bouddhistes et Hindouistes de méditation visent explicitement ces états), ou encore dans le monde Occidental au Moyen-Âge (où les superstitions ont souvent identifié les EMC à une influence démoniaque ou divine), et même durant la préhistoire : l'archéologue sud-africain David Lewis-Williams suggère que certains dessins rupestres datant d'environ 15000 ans av. J.-C. présentant des motifs caractéristiques (zigzags, spirales...) auraient été influencés par des visions produites par des EMC (Garcia-Romeu & Tart, 2013, pp. 124-126).

Caractérisation des états modifiés de conscience

Une littérature abondante existe dans plusieurs domaines au sujet des EMC (psychologie, neurosciences, anthropologie, philosophie, etc.), explorant des états aussi variés que les expériences mystiques, l'épilepsie, les rêves, l'état méditatif... Chacun de ces états possède une phénoménologie propre, plus ou moins éloignée de celle de l'état de veille ordinaire. Rassembler cette pluralité d'expériences sous une seule dénomination est

³⁵ Il ne s'agit là que d'une définition parmi d'autres, qui pourra être sujette à débat. Elle sera néanmoins suffisante dans le cadre restreint de ce travail, dans lequel un seul type d'EMC sera abordé, auquel cette définition correspond bien.

90 • Observer la conscience altérée

problématique, tenter d'en extraire des constantes l'est tout autant. Néanmoins ce type d'effort est nécessaire pour donner une base empirique aux recherches menées sur ce terrain. Si les outils de mesures développés dans le cadre de recherches sur les EMC peuvent ainsi sembler réducteurs, ils n'en demeurent pas moins nécessaires.

Je souhaiterais passer en revue les différents facteurs de l'échelle *11 Dimensions of Altered States of Consciousness (11D-ASC)*, un outil de mesure utilisé pour repérer la survenue d'EMC et évaluer leur intensité. La pertinence et la portée de cet outil (par rapport aux autres outils psychométriques du même type, notamment) ne sera pas discutée, mais il semblerait qu'il soit en mesure de rendre compte adéquatement et avec une certaine précision des dimensions importantes d'EMC, y compris l'expérience psychédélique.

Comme son nom l'indique, l'échelle est composée de 11 facteurs³⁶ censés refléter les facettes possibles d'un EMC, quel que soit son type, bien qu'un nombre important de ces facteurs semblent refléter la réalité de certains types d'EMC plutôt que d'autres. En décrivant succinctement chacun de ces facteurs, grâce à des termes clairs, définis, ou en recourant à l'occasion à des témoignages éclairants à leur sujet, j'espère parvenir à donner une première idée au lecteur de la phénoménologie d'un EMC.

Sentiment d'unité (Experience of Unity)

Ce premier facteur est assez difficile à définir puisqu'il renvoie à une phénoménologie particulièrement contre-intuitive. Les *items* correspondants de l'*11D-ASC* décrivent un impression que les choses s'unifient en une seule réalité, une connexion particulière entre Soi et l'environnement, un sentiment d'unité temporelle ou d'éternité. Ce sentiment peut être vécu positivement ou non et ne représente pas en lui-même une expérience mystique (Letheby, 2021, pp. 48-49).

La définition courante de l'expérience mystique utilisée dans les *Mystical Experience Questionnaire (MEQ)* et le *Hood Mysticism Scale* (Letheby, 2021, p. 25) sépare l'unité en deux facteurs : unité interne et unité externe, lié à des *items* employant un vocabulaire proche du mysticisme. L'unité externe serait liée à un sentiment de connexion particulière avec l'environnement (objets ou personnes), ou une prise de conscience de la « présence vivante » des choses ou leur interdépendance et l'unité interne à un sentiment

³⁶ Ces 11 facteurs sont subdivisés en 42 *items*. Ceux-ci guideront mes descriptions et y seront soit cités explicitement, soit présents en filigrane. J'ai cependant choisi de ne pas tous les citer, par soucis d'économie de place et de moyens. La liste complète (en anglais) des 11 facteurs et leurs *items* tirée de (Studerus, Gamma, & Vollenweider, 2010, p. 9) est reprise en Annexe.

d'union ou de fusion avec un tout plus grand que Soi ou que la réalité, un sentiment d'éternité ou d'infini (Barrett & Griffiths, 2017, p. 17).

La fusion sujet-objet caractérisant le sentiment d'unité externe, ainsi qu'un rapport et une attention particuliers au vivant, sont présents dans ce témoignage rapporté par W. A. Richards :

« Il y avait de profonds sentiments d'unité, d'abord avec une tulipe rouge, ensuite avec une rose de la même couleur. En regardant la rose comme un objet, elle sembla s'animer devant mes yeux [...] jusqu'à ce que, soudainement, je connaisse la rose ; c'est-à-dire, je devins d'une façon Un avec la rose, n'existant plus comme un ego voyant passivement un objet dans son environnement. »³⁷

Tandis que le témoignage suivant, issu d'une expérience vécue par Michael Pollan durant une prise de 5-MeO-DMT (une substance psychédélique), décrit une unité interne : on y trouve une disparition des frontières du Soi, « éparpillé », et fusionnant avec la réalité tout entière, ce qui suggère également une unification temporelle :

« J'ai ressenti un élan d'énergie remplir ma tête... puis... "Je" n'étais plus, éparpillé en un nuage de confettis par une force explosive. Je ne pouvais plus [me] localiser dans ma tête, parce qu'elle avait explosé également, se dilatant pour devenir tout ce qui a jamais été. »³⁸

Ce témoignage dépeint aussi un effacement complet du « Je ». Le sentiment d'unité est relié à l'expérience d'une diminution ou disparition du sens de Soi, souvent appelée « dissolution de l'ego » (Letheby, 2021, p. 47).

Expérience spirituelle (Spiritual Experience)

Les expériences spirituelles (révélation, illumination ou grâce à caractère mystique) pourraient être considérées comme faisant partie des phénomènes paranormaux en raison de leur caractère extraordinaire et de leur proximité conceptuelle avec de nombreuses croyances métaphysiques non-naturalistes. Néanmoins, les études psychédéliques ont montré que l'administration de certaines substances permettait d'induire de manière fiable des expériences décrites par les sujets eux-mêmes, y compris les sujets naïfs à toute spiritualité, comme étant « spirituelles » (Letheby, 2021, p. 199). D'après Letheby, la régularité avec laquelle l'expérience psychédélique et la spiritualité sont mises en lien, y

³⁷ Richards, W.A., *Sacred knowledge : psychedelics and religious experiences*, 2015. Cité dans (Letheby, 2021, p. 48).

³⁸ Pollan, M., *How to change your mind : the new science of psychedelics*, 2018. Cité dans (Letheby, 2021, p. 120).

compris chez des sujets naïfs à toute spiritualité, montrent qu'il s'agit d'une expérience « paradigmatiquement spirituelle » (Letheby, 2021, p. 199).

C'est la raison pour laquelle les drogues hallucinogènes (psychédéliques en tête) ont une place de choix dans les recherches au sujet des expériences mystiques (voir, par exemple, les travaux du *Center for Psychedelic & Consciousness Research* de l'université Johns Hopkins, dont certains ont déjà été cités dans ce travail (Barrett & Griffiths, 2017)) : « Pour la première fois, les scientifiques sont en mesure d'induire des expériences paradigmatiquement spirituelles en laboratoire de façon sûre, fiable et répétée [...] chez des populations diverses avec différentes croyances préalables et des attitudes concernant la spiritualité », ce qui nous donne un accès privilégié à des « preuves de haute qualité concernant la nature, les causes et les conséquences » de ces expériences (Letheby, 2021, p. 199).

L'expérience spirituelle peut être dotée d'une signification religieuse. D'ailleurs, la tendance des sujets à qualifier leur expérience de « signifiante spirituellement » est corrélée à de hauts scores sur les échelles mesurant la survenue d'expériences mystiques (Letheby, 2021, p. 200). Cependant, ce type d'expériences n'implique pas forcément la présence préalable ou l'adoption par le sujet durant l'expérience de croyances métaphysiques non-naturalistes. Plutôt, la béatitude ainsi que d'autres aspects inhabituels de l'expérience sont mis en lien avec des connaissances préalables au sujet des états mystiques, ce qui permet de les définir de façon plus précise sans pour autant que le sujet identifie entièrement son expérience à la description qu'il en fait. Le témoignage montre clairement, selon moi, cette ambivalence entre la description donnée d'un phénomène et son sens réel pour le sujet :

« Submergé par l'amour et tout l'amour que j'ai pour ma famille et mes amis. J'ai senti que cela venait d'eux ; j'ai également senti que j'étais baigné dedans. Et si j'étais religieux, ça aurait définitivement été une expérience religieuse, j'aurais dit que j'étais baigné dans l'amour de Dieu. Et je ne pense pas que l'anglais ait un moyen de dire ça sans utiliser ce mot "Dieu", peut-être baigné dans un amour transcendant. Baigné dans l'amour universel. C'était un sentiment tellement fort. »³⁹

Le sentiment de révérence (« awe ») ou la connexion avec quelque chose de plus grand souvent décrits ne sont pas plus incompatibles avec une conception naturaliste du monde, d'après Letheby. Il existe plusieurs conceptions relevant d'une spiritualité naturaliste, qui partagent, selon Jerome Stone, trois dimensions (Letheby, 2021, p. 198) –

³⁹ Swift, T.C., et al., *Cancer at the dinner table : experiences of psilocybin-assisted psychotherapy for the treatment of cancer-related distress*, 2017. Cité dans (Letheby, 2021, p. 203)

un sentiment de connexion à quelque chose de plus grand, l'aspiration à de plus grandes choses et à réaliser ses idéaux, et enfin une réflexion sur les Grandes Questions (Letheby, 2021, pp. 200-201) – que l'on retrouve également dans de nombreux témoignages. Ces trois piliers consistent d'après Jones en une façon « d'agrandir nos Soi » (« *enlarging our selves* ») ou de « percer les murs étroits de l'ego » (« *breaking through the narrow walls of ego* ») (Letheby, 2021, p. 198), ce qui est consistant avec l'hypothèse de Letheby d'après laquelle le mode de fonctionnement principal de l'expérience spirituelle tien à une modification de la représentation de Soi.

État de béatitude (Blissful State)

Ce troisième facteur est relié à l'expérience d'un « plaisir sans limites », d'une profonde paix intérieure ou encore d'un « amour embrassant toutes choses » (« *all-embracing love* »). Cette extase est souvent décrite par son contenu sémantique à l'aide de termes chargés de sens parfois spirituels (« paix », « amour », « béatitude », « extase », entre autres (Studerus, Gamma, & Vollenweider, 2010, p. 9) (Barrett & Griffiths, 2017, p. 5), mais il est intéressant de noter qu'elle correspond également à un état cognitif plus simple délimité par le facteur « humeur accrue » dans l'échelle *EWL-60-S*⁴⁰ (Studerus, Gamma, & Vollenweider, 2010, p. 13). De la même façon que la géométrie simple peut se transformer en constructions complexes voire en véritables hallucinations, l'extase vécue comme hautement spirituelle serait en réalité du même ordre que le plaisir et l'euphorie bien terrestre recherchés par l'usage « récréatif » de drogues, et pourrait plus facilement que les précédents facteurs être identifiée à des processus neuronaux élémentaires (Barrett & Griffiths, 2017, p. 22). Ainsi, bien que la « joie et l'enchantement » accompagne souvent les expériences mystiques ou spirituelles (Letheby, 2021, p. 25), elle peut aussi apparaître seule et se suffire à elle-même.

Prise de conscience (Insightfulness)

L'*insightfulness* est la capacité à induire un *insight*. Ce terme est difficile à traduire en français, notre langue ne possédant pas d'équivalent clair. Le *Cambridge Dictionary* définit le terme comme « (la capacité à avoir) une compréhension claire, profonde, et

⁴⁰ L'*EWL-60-S* (pour « *Eigenschaftswörterliste* ») est une échelle mesurant l'état mental du sujet. Elle se compose de 60 adjectifs, groupés en 15 sous-échelles incluant : efficacité, concentration, inactivation, fatigue, hébétéude, extraversion, introversion, confiance en Soi, humeur accrue, excitation émotionnelle, sensibilité, colère, anxiété, dépression et rêverie. Il a été montré que cette échelle est apte à mesurer les changements d'état mental à court terme induits par des substances psychoactives (Studerus, Gamma, & Vollenweider, 2010, p. 5).

94 • Observer la conscience altérée

parfois soudaine, d'un problème ou d'une situation complexe »⁴¹. Une traduction possible serait « perspicacité », ce qui correspondrait à la capacité à comprendre intuitivement une situation ou un problème complexe, mais dans le cadre qui nous occupe, ce terme serait, il me semble, réducteur, étant donné son sens dans le langage courant, qui le rapproche souvent de qualificatifs liés à l'intelligence ou la finesse d'esprit.

Une personne perspicace est, dans le langage courant, une personne maligne, qui analyse avec facilité une situation, et dont les conclusions sembleront objectivement pertinentes. Or ce sens du terme peut faire oublier la notion d'intuition, qui a son importance dans l'*insight* permis par un EMC. En effet, si l'*insightfulness* permet une appréhension facilitée, éclairée, d'éléments auparavant problématiques, celle-ci ne se produit pas toujours par un raisonnement et dépasse parfois elle-même la compréhension. La perspicacité ressentie peut alors prendre la forme d'un état de clairvoyance, et l'*insight* celle d'une révélation. J'ai choisi de titrer cette section « prise de conscience » – puisqu'il fallait bien lui trouver un titre francophone – afin de ne pas occulter le caractère direct et subjectif de l'*insight*, mais dans la suite de ce travail, je conserverai le terme anglophone pour en conserver tout le sens. Notons également que ce facteur revêt une importance certaine pour ce travail, puisqu'il est au premier plan de l'hypothèse de Letheby, que j'aborderai plus en détails dans le chapitre 6.

L'*insightfulness* se caractérise donc par une compréhension ou une perspective nouvelle d'éléments de la vie d'une personne, dont l'acquisition, parfois soudaine, semble éclairante pour celle-ci. Un sujet pourra alors avoir l'impression de découvrir « comme des aspects de [sa] vie affectent [son] bien-être », ou de « voir [ses] problèmes avec plus de clarté », ainsi que d'acquérir une « compréhension plus profonde » de ses croyances, valeurs, ou de sa vie en général⁴². Un témoignage de sujet d'expérience sur les bénéfices thérapeutiques de la psilocybine repris par Letheby illustre bien ce facteur :

« Je pensais aux relations que j'ai eu avec d'autres personnes et je pensais que je pouvais les voir clairement pour la première fois. J'ai eu une nouvelle perspective ("fresh insight") sur les choses. C'était presque comme si soudainement,

⁴¹ *Insight*. Dans : *Cambridge Dictionary*. En ligne : <https://dictionary.cambridge.org/dictionary/english/insight>.

⁴² Les expressions reprises entre guillemets proviennent d'un questionnaire développé par Davis et. Al, 2020, repris par Letheby et mis en relation avec le facteur *insightfulness* de l'11D-ASC. Letheby note que bien que le recoupement entre les items de ce questionnaire et l'11D-ASC n'est pas clairement établi, certains éléments permettent d'affirmer leur proximité (Letheby, 2021, pp. 84-85).

les écailles m'étaient tombées des yeux, je pouvais voir les choses comme elles sont réellement. »⁴³

Nous retrouvons bien dans ce témoignage tous les éléments d'une expérience d'*insightfulness* : compréhension nouvelle, claire, soudaine, et, dans ce cas-ci, vécue avec une forme de révélation, comme pourrait le suggérer l'expression « les écailles me sont tombées des yeux » (« *the scales dropped from my eyes* »), tirée de la Bible (Actes 9,18) où elle décrit le recouvrement miraculeux de la vue de Saint Paul.

Désincarnation ou décorporation (Disembodiment)

Des sensations tactiles ou proprioceptives inhabituelles peuvent également apparaître durant un EMC. Parmi celles-ci, on retrouve la « disparition des frontières du corps » et le sentiment de quitter son corps ou que celui-ci a disparu (décorporation ou désincarnation).

La sensation la moins problématique parmi celles que regroupent ce facteur est sans doute la sensation de « flotter ». Une sensation d'apesanteur peut ou non s'accompagner d'une altération de la conscience de son propre corps : par exemple, dans les rêves, les sensations corporelles peuvent être conservées et accompagner une sensation de changement de gravité d'une impression de vitesse ou de mouvement horizontal ou vertical (on pourra alors avoir l'impression de voler, de tomber ou au contraire d'être propulsé dans les airs).

Contrôle et cognition affaiblis (Impaired Control and Cognition)

Dans la mesure où les EMC sont des états altérés de fonctionnement du cerveau, il n'est pas étonnant que ceux-ci altèrent également le contrôle précis que permet celui-ci en temps normal. Un EMC peut ainsi s'accompagner d'une difficulté à prendre des décisions ou encore à réaliser des actions ou mener à son terme une pensée, d'une sensation de perte d'agentivité, se traduisant par une impression de perte de volonté, ou de n'être qu'une marionnette (ce qui évoque une dissociation entre la volonté d'un agent et les mouvements qu'il produit), d'une incapacité à juger de l'importance des choses ou à entrer en contact avec les autres.

Anxiété (Anxiety)

Un EMC peut s'accompagner d'un sentiment d'anxiété, dont la source n'est pas toujours identifiée qui s'accompagne de symptômes tels qu'agitation, stimulation,

⁴³ Watts, R., et al., *Patients' accounts of increased « connectedness » and « acceptance » after psilocybin for treatment-resistant depression*, 2017. Cité dans (Letheby, 2021, p. 85).

difficultés de concentration, irritabilité ou tension musculaire (DSM-V, 2013, p. 818). Dans le cadre des recherches psychédéliques, cet état est souvent rencontré par les sujets recevant de fortes doses durant une session ; heureusement, le support offert par les encadrants de la session permet de faire face à ce type de situation, et ainsi permettre au sujet de rester en contrôle et d'éviter la panique (Letheby, 2021, p. 15).

Imagerie complexe (Complex Imagery)

La littérature au sujet des hallucinations distingue généralement plusieurs types d'anomalies perceptuelles qui sont regroupées ici sous le terme « imagerie ». Cette généralisation n'est sans doute pas dommageable aux facteurs retenus pour l'ASC, mais il me semble néanmoins intéressant d'étoffer quelques peu ceux-ci afin de rendre compte plus clairement de la phénoménologie complexe des hallucinations.

En effet les *items* retenus par Studerus *et al.* formant les facteurs relatifs aux perceptions visuelles anormales renvoient surtout à un type de visions particulier, apparaissant « dans le noir complet ou alors que [le sujet] avait les yeux fermés » (Studerus, Gamma, & Vollenweider, 2010, p. 9), laissant de côté une gamme de déformations du champ visuel souvent concomitantes d'EMC et qui ont également une place importante dans la phénoménologie de celles-ci.

Plusieurs phénomènes se rapprochent en principe de l'hallucination : la pseudo-hallucination, l'illusion, les rêves et l'imagerie mentale⁴⁴. Tous ont en commun d'être des perceptions ne représentant pas adéquatement un état de fait réel, et d'être reliés à la phénoménologie des EMC. Il est donc intéressant de les aborder dans le cadre d'une phénoménologie des EMC.

Le glossaire psychiatrique de l'*American Psychiatric Association* définit l'hallucination comme « une perception sensorielle ayant lieu en l'absence de tout *stimulus* externe » (DSM-V, 2013, p. 822), une définition insuffisante, qui ne permet par exemple pas de distinguer l'hallucination du rêve. Une définition émise par Mullen citée par Chiu (Chiu, 1989, p. 293) permet de mieux cerner ce phénomène : d'après Mullen, une hallucination est une « fausse perception involontaire ayant lieu en simultanéité avec les perceptions réelles (ce qui exclut les rêves), et ayant les qualités des vraies perceptions,

⁴⁴ Les deux premiers phénomènes, pseudo-hallucinations et illusions, seront abordés ci-après. Les rêves et l'imagerie mentale seront laissés de côté dans le cadre de ce travail. Notons simplement que le rêve se distingue de l'hallucination en ce qu'il se produit alors que le sujet est inconscient, tandis que l'imagerie mentale s'en distingue par le fait qu'elle soit sous le contrôle volontaire du sujet et ne s'impose donc pas à lui.

c'est-à-dire la vivacité, la substantialité, et la localisation dans l'espace objectif externe » (Chiu, 1989, p. 293).

Cette définition isole l'hallucination des autres phénomènes suscités en soulignant leur proximité avec la perception réelle. En effet, les hallucinations « parviennent à imiter la perception à tous égards » (Sacks, 2014, p. 10), ce qui les rend habituellement difficiles à discerner de la réalité et explique qu'elles puissent tromper les personnes qui en sont victimes.

Plusieurs pathologies s'accompagnent également d'épisodes hallucinatoires. Par exemple, le syndrome de Charles Bonnet, qui affecte les personnes « devenues aveugles ou dont la vue s'est détériorée » en l'absence de toute autre anomalie neurologique (Sacks, 2014, pp. 18-19), se caractérise par l'apparition d'hallucinations complexes dans le champ visuel. Une patiente suivie par Oliver Sacks lui raconta voir des « gens vêtus à l'orientale [...] montant et descendant des escaliers » entre autres éléments formant des scènes qu'elle pouvait décrire avec une grande précision (Sacks, 2014, p. 18). Cependant, malgré le réalisme des hallucinations, voire leur indiscernabilité du réel, leur présence ne trompe pas toujours celui qui les voit. Dans la plupart des cas, l'incongruité des visions suffit à faire prendre conscience de leur caractère hallucinatoire (Sacks, 2014, p. 26).

Les pseudo-hallucinations sont un second type d'expériences hallucinatoires proches des hallucinations par leur caractère involontaire et incontrôlable, mais aussi par leur complexité. Néanmoins, les pseudo-hallucinations ne sont pas projetées « dans l'espace extérieur » (Sacks, 2014, p. 10) mais restent confinées à l'espace « intérieur subjectif ». Elles ne possèdent en outre pas la « vivacité » et la « substantialité » des perceptions normales qui caractérisent également les hallucinations (Chiu, 1989, p. 293). Elles ne convainquent pas par leur réalisme et ne sont par conséquent pas confondues avec le réel.

Le troisième type d'expérience hallucinatoire que je souhaiterais détailler ici est l'illusion. Celle-ci se distingue de l'hallucination par le fait qu'elle est une perception distordue d'un *stimulus* externe réel (Chiu, 1989, p. 292). Les illusions font sans doute bien plus partie de nos expériences habituelles que les hallucinations et pseudo-hallucinations : certaines font partie de nos quotidiens, d'autres sont des dispositifs célèbres. Néanmoins, les illusions peuvent également faire partie d'un EMC, et y coexister avec d'autres expériences anormales. Dans un témoignage déjà cité, un sujet décrit une rose qui « sembla

s'animer » devant lui et dont les pétales « semblaient respirer »⁴⁵. Cette description est très certainement à prendre au premier degré, l'une des illusions communément vécues par les personnes sous l'influence de psychédéliques étant l'impression que les objets ou les plantes respirent, car ceux-ci sont perçus comme se contractant et se dilatant en rythme, rappelant les mouvements respiratoires d'organismes vivants. Ici, l'illusion coexiste avec le sentiment d'unité éprouvé ensuite, voire l'alimente, puisque la proximité émotionnelle avec la plante pourrait être renforcée par l'impression que celle-ci respire de la même façon qu'un animal.

Les différents types d'expériences hallucinatoires abordés ci-dessus peuvent se recouper, et distinguer de quel type relève quel phénomène particulier peut devenir difficile dès lors que les situations se complexifient. Sacks remarque ainsi que « chez quantité de [ses] patients, de véritables hallucinations coexistent avec des illusions et des perceptions erronées de la plus haute complexité, la ligne de démarcation entre les premières et les dernières pouvant être difficile à tracer » (Sacks, 2014, p. 11). Aussi, je ne me prononcerai pas sur la nature des expériences hallucinatoires relatées par les différents témoignages repris dans ce travail ; savoir si elles relèvent de l'hallucination pure ou de la pseudo-hallucination, ou même de l'imagerie onirique, ne m'intéresse pas.

Selon moi, une taxonomie des hallucinations permet surtout de prendre conscience de la grande variété d'expériences hallucinatoires, et notamment des différents degrés de complexité qui peuvent les caractériser. Le continuum qui sépare l'expérience habituelle, ponctuée d'illusions et d'hallucinations bénignes, de la plus grave psychose hallucinatoire permet en outre de se rendre compte de la contingence de l'expérience consciente, qui apparaît face aux EMC comme une façon (certes mieux adaptée à la majorité des activités humaines) parmi d'autres de percevoir le monde.

Imagerie élémentaire (Elemental Imagery)

L'imagerie élémentaire qui peut accompagner un EMC est proche en principe de l'imagerie complexe et s'en distingue, comme son qualificatif le laisse entendre, par sa moindre complexité. Ici, il n'est pas question d'hallucination car les images dont il est question ne sont clairement pas assez développées pour remplir les critères de réalisme, de « vivacité » ou « substantialité » des hallucinations.

⁴⁵ Richards, W.A., *Sacred knowledge : psychedelics and religious experiences*, 2015. Cité dans (Letheby, 2021, p. 48).

Cela n'implique pas pour autant que ce type d'imagerie soit simpliste. Bien que les *items* de l'échelle utilisent des termes qui puissent le laisser suggérer (ils font référence à des « motifs », « couleurs », ou « flashes de lumière »), je pense que ce facteur vise un type d'expérience hallucinatoire élémentaire dans le sens où n'y sont pas reproduits des objets complexes et réalistes (visages, personnes, lieux, *etc.*). Néanmoins, ces visions peuvent être très détaillées dans leur construction et très riches dans leur sens.

Les motifs géométriques sont des visions régulièrement rapportées lors d'EMC. Ces motifs peuvent prendre la forme de constantes récurrentes décrites et classées en quatre familles par Klüver en 1928 : les tunnels et entonnoirs, les spirales, les tressages (formant des alvéoles ou des triangles) et les toiles d'araignées (Bressloff, 2002). Mais souvent, une géométrie plus complexe est décrite, comme dans le témoignage suivant :

« Je découvris bientôt dans mon champ visuel l'émergence d'un réseau multi-dimensionnel de motifs géométriques complexes, semblables à des néons, qui attirait mon attention encore plus profondément. Je pouvais voir cela avec les yeux ouverts, mais découvrit qu'il était encore plus vivace et nettement présent quand je fermais les yeux. Je reconnus la vie dans ces motifs ondulants et commençai à me sentir comme si je pouvais entrer dans l'énergie qui s'écoulait en eux. Bientôt, je me sentis immergé dans une imagerie incroyablement détaillée qui pourrait être décrite comme de l'architecture islamique et des écritures arabes, à propos desquelles je ne connaissais rien. »⁴⁶

Une continuité pourrait exister entre les visions les plus simples et les expériences les plus complexes. Les migraines s'accompagnent souvent d'épisodes hallucinatoires appelés « auras ». Durant ceux-ci, une vision courante consiste en une forme en zigzag rappelant les fortifications d'un château fort (Sacks, 2014, pp. 149-151), mais cet effet peut s'accompagner d'autres expériences hallucinatoires, visuelles ou non. Sacks, décrivant ses propres auras migraineuses, évoque apercevoir à l'intérieur du scotome (l'aire du champ visuel rendue aveugle lors de l'aura) des « treillis, des damiers, des toiles d'araignées et des nids d'abeille », c'est-à-dire des formes du type des constantes de Klüver citées plus-haut, qui se déplaçaient, se déformaient ou se combinaient pour former des motifs plus complexes tels que « des décors de tapis turcs, des mosaïques complexes ou de minuscules pommes de pin et oursins en trois dimensions » (Sacks, 2014, p. 151). D'autres patients affirment que leurs auras s'accompagnent de véritables hallucinations visuelles (Sacks, 2014, p. 155).

⁴⁶ Richards, W.A., *Sacred knowledge : psychedelics and religious experiences*, 2015. Cité dans (Letheby, 2021, p. 41).

100 • Observer la conscience altérée

Enfin, notons que les visions peuvent prendre une place très importante dans le champ visuel ou, s'ajoutant à d'autres effets, prendre la forme d'une expérience totale et non plus seulement visuelle. Dans le témoignage repris plus haut, le sujet voit d'abord les motifs avant d'être « immergé » dedans d'une façon qui rappelle les descriptions liées au facteur « sentiment d'unité ».

Synesthésie audio-visuelle (Audio-Visual Synesthesiae)

La synesthésie est un phénomène perceptuel dans lequel la stimulation d'une modalité sensorielle mène à une expérience involontaire dans une seconde modalité. La synesthésie dont il est question ici est un phénomène temporaire, qui diffère donc de la disposition congénitale du même nom. Les phénomènes de synesthésie temporaire peuvent apparaître durant un EMC, lorsque des perceptions provenant de plusieurs modalités sensorielles s'influencent mutuellement. Les *items* de ce facteur renvoient en effet à une « influence » mutuelle entre les sons et l'image (« Les sons semblaient influencer ce que j'ai vu »), ce qui peut se traduire simplement par une influence de la musique sur l'imagerie, donnant l'impression que les visuels bougent en rythme avec celle-ci.

Changement de sens des perceptions (Changed Meaning of Percepts)

Le dernier facteur de l'échelle *IID-ASC* est d'une grande importance puisqu'il a trait au sens qui se dégage de l'expérience vécue durant un EMC. De nombreux témoignages – incluant certains des témoignages déjà cités ici – indiquent des perceptions saturées de sens, auxquelles se mêlent souvent des réflexions philosophiques ou spirituelles ou des souvenirs dont certains peuvent être représentés avec un très grand niveau de détail. À cela s'ajoute une implication émotionnelle inhabituelle qui donne à ces réflexions une dimension personnelle particulièrement profonde, à mettre en lien avec l'expérience d'*insight* évoquée plus haut.

L'abondance de sens des choses perçues durant l'expérience fait partie des mécanismes psychologiques, tels que l'*insight*, qui pourraient jouer un rôle dans l'effet thérapeutique des psychédéliques (Letheby, 2021, p. 125). Ce mécanisme, lié au fonctionnement du « réseau de la saillance » (« *saliency network* », ou *SN*), jouerait également un rôle dans la représentation de Soi (Letheby, 2021, p. 151), ce qui explique que les expériences de sens amplifié soient régulièrement reliées au Soi et à des considérations biographiques.

5.2 Expérience psychédélique

Le concept d'état modifié de conscience ayant été défini, nous pouvons à présent en présenter le type particulier qui fera l'objet de la fin de ce travail : l'expérience psychédélique. Avant d'aborder la place que celle-ci a tenu dans l'histoire des sciences et de décrire sa phénoménologie, une définition s'impose. Le terme « expérience psychédélique » désigne l'état modifié de conscience caractéristique qui se manifeste chez un sujet lorsque celui-ci est sous l'effet d'une substance dite « psychédélique ». Le terme « psychédélique » (signifiant « qui rend visible l'esprit », formé des racines grecques *ψυχή*, l'esprit, et *δηλείν*, rendre visible) a été proposé en 1957 par le psychiatre Humphrey Osmond pour désigner un ensemble de drogues ayant la particularité de causer de « dramatiques changements dans la perception, les émotions et la cognition » (Sessa 2012, cité par Letheby 8). Le terme est parfois utilisé pour désigner toute drogue hallucinogène susceptible de provoquer un EMC, telles que la sauge divinatoire ou la kétamine, mais les psychédéliques « classiques » possèdent, en plus d'une phénoménologie particulière, résumée dans ce travail par l'expression « expérience psychédélique », un mode d'action identique : les substances telles que le diéthyllysergamide (LSD), la psilocybine (présente dans les champignons hallucinogènes), la mescaline (présente dans plusieurs espèces de cactus) ou encore la diméthyltryptamine (DMT, présente dans certaines espèces de plantes tropicales sud-américaines), agissent toutes en imitant l'action de la sérotonine au niveau des récepteurs du sous-type 5-HT_{2A} (récepteurs à la sérotonine-2a), (Letheby, 2021, pp. 8-9) ce qui produit une large gamme d'effets constants entre ces différentes substances malgré quelques variations (liées à la puissance des composés, à leur mode d'administration ou à d'éventuelles modes d'action secondaires).

Aperçu historique des recherches psychédéliques

Dans les années 1950 et 1960, les psychédéliques ont été étudiés de plusieurs façons. Certains chercheurs y virent des substances « psychotomimétiques » (qui imitent la psychose) ou « psychotogènes » (qui causent la psychose), ce qui définit leurs deux premières applications majeures : « l'étude des mécanismes d'action des psychédéliques, dans l'espoir de trouver des bases biochimiques aux maladies mentales, et l'ingestion contrôlée de psychédéliques par les psychiatres, afin d'augmenter leur empathie pour les patients psychotiques » (Letheby, 2021, p. 9). Mais aux symptômes psychotiques s'ajoutaient une composante mystique ou spirituelle, qui piqua la curiosité de certains

102 • Observer la conscience altérée

intellectuels ayant un intérêt pour ces thématiques ; deux d'entre eux, Aldous Huxley et Alan Watts en vinrent ainsi à poser les bases de la « spiritualité psychopharmacologique de la contreculture des années 1960 » (Letheby, 2021, p. 10).

Ainsi, la conception des psychédéliques comme facilitateurs d'expériences spirituelles émergea et [entra dans] les milieux académiques. En 1962, Walter Pahnke réalisa une expérience, connue aujourd'hui sous le nom de « Good Friday Experiment », consistant à administrer à des étudiants en théologie une dose de psilocybine durant une messe du Vendredi Saint. Cette expérience, bien que défectueuse dans sa méthodologie, révéla le potentiel des psychédéliques dans l'étude des phénomènes religieux (Letheby, 2021, p. 10).

Enfin, c'est également à cette époque que le potentiel thérapeutique des psychédéliques fut remarqué. Conformément au paradigme psychanalytique dominant à l'époque, ce bénéfice fut interprété comme provenant d'une expression de « matériau inconscient : souvenirs, désirs, fantasmes refoulés ». Les thérapies pratiquées en Europe s'agrémentèrent donc de doses faibles de psychédéliques afin de « permettre l'accès à l'inconscient et d'améliorer le lien patient-thérapeute » (Letheby, 2021, pp. 10-11).

Aux États-Unis cependant, un autre type de thérapie était pratiqué, consistant à administrer, en lieu et place de doses faibles mais régulières, une seule dose forte de psychédéliques, dans le but de provoquer une expérience transformative intense. Comme l'expérience déjà citée de Pahnke, la plupart de ces travaux étaient problématiques, mais leurs résultats sont aujourd'hui corroborés par les recherches récentes (Letheby, 2021, p. 11).

Après une traversée du désert de près de 30 ans, les « recherches psychédéliques » ont repris dans les années 1990 dans un contexte scientifique ayant bénéficié d'avancées sur les questions méthodologiques et théoriques et d'un développement de technologies. Cette « renaissance psychédélique » a permis à ce champ de recherche de s'accélérer jusqu'à nos jours (Letheby, 2021, p. 12).

Les recherches actuelles montrent plusieurs choses à propos des psychédéliques : premièrement leur relative innocuité – administrées dans les bonnes conditions, à des sujets ne présentant pas de risques psychiatriques et suivis par une équipe de professionnels, ces substances sont très sûres autant sur le plan physiologique que psychologique (Letheby, 2021, pp. 13-16). Deuxièmement, de nombreux travaux actuels corroborent certaines

hypothèses émises dès les années 1960 à propos de l'efficacité des psychédéliques dans le traitement des addictions et de la dépression. Concernant les premières, deux études pilotes menées par Bogenschutz *et al.* en 2015 et Johnson *et al.* en 2014 (citées dans (Letheby, 2021, p. 17)) conclurent qu'administrer une à deux doses de psychédéliques, accompagnées d'un suivi adéquat, suffisait à réduire ou supprimer les addictions d'une majorité de sujets, avec des effets portant sur plusieurs mois. Peut-être plus étonnant encore, ceux qui répondaient le mieux au traitement montraient également de plus hauts scores d'expériences mystiques, suggérant qu'au-delà de l'action pharmacologique de la drogue, l'expérience elle-même jouerait un rôle dans les résultats du traitement (Letheby, 2021, pp. 17-18).

D'autres études-tests centrées cette fois sur l'utilisation de psychédéliques pour traiter la dépression montrent des résultats similaires : une dose unique de psychédéliques serait capable de réduire les symptômes dépressifs (y compris chez des patients atteints de dépression majeure ou résistante aux traitements) jusqu'à plusieurs mois après son administration. Ces résultats ont poussé la *U.S. Food and Drug Administration (FDA)* (l'agence chargée d'autoriser la commercialisation de médicaments aux États-Unis) à qualifier le traitement de la dépression à l'aide de psilocybine de « *Breakthrough Therapy* », autorisant ainsi l'utilisation de cette substance dans des études plus larges (Letheby, 2021, pp. 18-19).

Enfin, les psychédéliques se montrent efficaces pour réduire les symptômes de détresse psychologique – comprenant anxiété, dépression et détresse existentielle – chez les patients atteints de maladies en phase terminale. À nouveau, il semblerait que l'efficacité de ce traitement soit elle aussi corrélée à de hauts scores d'expérience mystique. Deux études particulièrement rigoureuses, réalisées en double-aveugle sur de larges groupes, menées l'une à la New York University, l'autre à la John Hopkins University, ont répliqué ces résultats, ajoutant à la confiance placée dans les traitements psychédéliques (Letheby, 2021, pp. 19-20).

D'autres études encore suggèrent que les psychédéliques pourraient bénéficier également à des personnes en bonne santé. Ces bénéfices pourraient être liés à une augmentation de l'Ouverture à l'Expérience, l'un des cinq traits du modèle « *Big Five* » de la personnalité, et s'étendaient également jusqu'à plusieurs mois après l'expérience (Letheby, 2021, pp. 21-23).

L'objectif de ce bref aperçu historique était double : je souhaitais d'une part montrer la légitimité acquise de l'expérience psychédélique en tant qu'objet d'étude et de recherche depuis la « renaissance psychédélique », d'autre part je souhaitais contextualiser quelque peu le lien déjà évoqué dans le chapitre précédent entre expérience psychédélique et altération du Soi, en montrant le pouvoir transformateur de ces substances, révélé par les recherches actuelles. Avant d'explorer ces thématiques, il nous reste à décrire la phénoménologie de cette expérience.

Phénoménologie de l'expérience psychédélique

Au début de ce chapitre, j'ai affirmé que l'action des psychédéliques sur le cerveau était responsable d'une large gamme d'effets, sans citer ni décrire aucun de ces effets. Une définition de l'expérience psychédélique ne serait cependant pas complète sans un aperçu de ceux-ci. L'objectif de cette section sera donc de donner une idée de la phénoménologie de l'expérience psychédélique.

Un effort particulier de classification des effets phénoménaux provoqués par la prise de psychédéliques guidera ma description : le *Subjective Experience Index* (Index des Expériences Subjectives, abrégé *SEI* par la suite) mis en ligne par l'encyclopédie communautaire en ligne *Psychonaut Wiki*⁴⁷, dédiée à la documentation au sujet des substances psychotropes et de la réduction des risques liés à l'usage de celles-ci. Cet index est entretenu par Graham Reed sur la base de littérature scientifique (lorsque celle-ci existe) ainsi que de comptes rendus d'expériences de contributeurs du site ou d'utilisateurs anonymes. Ces comptes rendus forment un corpus documentaire important bien que problématique : la grande variété de profils et d'intention, de leurs auteurs implique un possible haut niveau de déformation des faits. Pour remédier à ceci, en plus du recoupement, je conserverai le scrupule méthodologique de l'hétérophénoménologie consistant à « mettre entre crochets » (Dennett, 2012, pp. 71-72) le contenu des comptes rendus, c'est-à-dire l'accepter comme reflétant ce dont un sujet croit avoir fait l'expérience, sans préjuger de la réalité sous-jacente à ces affirmations.

J'ajouterai que le recoupement des différentes entrées du *SEI* avec les travaux académiques que j'ai pu lire, notamment l'ouvrage *Philosophy of Psychedelics* de Letheby,

⁴⁷ Le contenu de *Psychonaut Wiki* repris dans ce travail est édité sous licence *Creative Commons* (<https://creativecommons.org/licenses/by-sa/4.0/>), autorisant la copie, l'adaptation et le partage sous cette même licence de celui-ci. Le *SEI* est un contenu produit par Graham Reed, édité par l'équipe de *Psychonaut Wiki*.

ainsi qu’avec les nombreux comptes rendus que j’ai pu lire par ailleurs (notamment sur la base de données en ligne de l’organisation à but non-lucratif *Erowid*)⁴⁸ tend à soutenir la crédibilité du *SEI*. En outre, son contenu, bien que non revu par les pairs, ne devrait poser aucun problème majeur au regard des connaissances scientifiques actuelles sur les psychédéliques.

Enfin, je souhaite souligner que l’objectif de cette section est de décrire plusieurs éléments qui caractérisent l’état général dont un sujet pourrait faire l’expérience lors d’une expérience psychédélique. Il ne s’agit donc pas ici de décrire un épisode particulièrement intense de cette expérience, comme le font de nombreux extraits repris par Letheby dans son ouvrage. Rappelons en effet à ce sujet que l’action des psychédéliques s’étend parfois sur plusieurs heures (quelques heures pour la psilocybine, jusqu’à une dizaine d’heures pour le LSD), au cours desquelles le sujet pourra faire l’expérience d’une large gamme d’effets variant en intensité au cours du temps (Letheby, 2021, p. 44). En proposant une vision plus globale de cette expérience, j’espère souligner l’interdépendance de ses différents effets et donner une idée de comment ceux-ci coexistent pour créer un vécu complexe. Je ne souhaite pas en revanche dresser une liste exhaustive de l’ensemble des effets des psychédéliques – ce qui serait non seulement long et fastidieux, mais également trompeur, puisqu’un sujet ne rencontrera pas l’entièreté de ces effets au cours d’une seule expérience – mais d’en sélectionner un échantillon représentatif qui puisse esquisser de façon crédible une expérience-type.

Une expérience psychédélique peut être grossièrement divisée en trois phases : montée, pic d’intensité, descente. Celles-ci sont encadrées par une phase de « pré-montée » correspondant au temps nécessaire à la substance pour commencer à faire effet, et l’apparition d’effets secondaires à la fin de l’expérience, pouvant s’étendre durant plusieurs heures voire plusieurs jours. Durant la montée, les effets augmentent graduellement en intensité, jusqu’à atteindre un pic d’intensité, on entre alors dans la deuxième phase, surnommée « pic » ou encore « plateau », de durée variable et débouchant à son tour sur la

⁴⁸ Le contenu des comptes rendus que j’ai pu lire transparaîtra à l’arrière-plan du contenu de cette section, mais ceux-ci ne seront pas cités individuellement pour deux raisons. Premièrement parce que mon objectif est de produire une généralisation de la phénoménologie de l’expérience psychédélique, qui puisse refléter un grand nombre d’expériences particulières. Deuxièmement, pour des questions relatives au RGPD, puisque tout droit sur les comptes rendus publiés sur *Erowid* ainsi que sur *Psychonaut Wiki* sont détenus par leurs auteurs respectifs et il m’était impossible de demander une permission à chacun d’eux pour reprendre des extraits de leurs contributions. La citation générale d’*Erowid* et *Psychonaut Wiki* devrait suffire à soutenir mon propos, puisqu’on y retrouvera de nombreux comptes rendus soutenant mes affirmations, ainsi que le *SEI* auquel celles-ci renvoient toutes de façon plus ou moins directe.

descente durant laquelle les effets déclinèrent jusqu'à devenir presque imperceptibles – la majorité des effets s'étant effacés, certains effets résiduels faibles peuvent être ressentis durant plusieurs heures voire plusieurs jours après l'expérience. J'ai organisé la description qui suit en trois parties, titrées d'après les trois phases susmentionnées, et lui ai donné la forme d'un récit qui puisse rappeler les comptes rendus d'expériences psychédéliques que j'ai pu lire, bien qu'il soit écrit à la troisième personne. La personnification du sujet d'expérience, que j'ai choisi de nommer Sasha⁴⁹, par contre, est à vocation purement cosmétique – il m'a simplement semblé que cela rendrait le texte plus agréable.

Montée

Lorsque les premiers effets apparaissent, ceux-ci peuvent coexister dans certains cas avec des effets physiologiques désagréables. Sasha est ainsi victime de nausées et lourdeurs d'estomac. Celles-ci s'effacent généralement avec l'arrivée du pic mais peuvent créer un sentiment d'inconfort, surtout si elles s'accompagnent d'anxiété. De l'anxiété pourra être ressentie à n'importe quel moment de l'expérience, apparaissant spontanément ou provoquée par un élément extérieur. Cet effet est couramment rapporté durant la montée, où il peut s'accompagner d'un sentiment de catastrophe imminente. Si l'expérience se déroule dans un contexte favorable (par exemple en milieu contrôlé durant une thérapie), l'anxiété s'efface généralement, elle aussi, rapidement. Sasha ayant été correctement préparé à son expérience, il reste confiant malgré son état maladif.

La montée s'accompagne d'une altération plus ou moins importante de l'énergie physique : le LSD aura tendance à induire une stimulation (encourageant l'activité physique et pouvant favoriser la concentration à faibles doses, parfois accompagnée de bruxisme⁵⁰), mais la psilocybine que Sasha a ingérée entraîne en revanche une sédation (caractérisée par un manque global d'énergie et de motivation pour la moindre chose, et complémentée par des bâillements excessifs).

En parallèle de ces sensations, Sasha commence à ressentir les premiers effets caractéristiques des psychédéliques, les plus évidents étant sans doute les altérations de la perception visuelle, bien que celles-ci restent assez faibles durant la montée. Les couleurs

⁴⁹ En référence au surnom d'Alexander Shulgin (1925 – 2014), biochimiste, psychopharmacologue et auteur américain dont la contribution à l'étude des substances psychoactives est inestimable.

⁵⁰ Serrage ou grincement de dents, plus caractéristique des drogues stimulantes telles que les amphétamines et la MDMA. Outre la sensation physique l'accompagnant (sensation de pression, pouvant être gênante ou douleurs musculaires), cet effet modifie également l'expression faciale du sujet, qui peut alors apparaître crispé ou grimaçant.

lui apparaissent alors plus vives et plus contrastées, et un phénomène de diffraction (possiblement lié à la dilatation des pupilles) colore son champ de vision de traces rouges, vertes et jaunes lorsqu'il se tourne vers une source de lumière. Ces effets appliquent au champ de vision un filtre particulier, qui rappelle à Sasha les filtres intensifiant couleurs et contrastes que l'on peut appliquer aux photos grâce aux applications de *smartphones*. Ce filtre teintera sa vision durant l'entièreté de l'expérience. Des distorsions faibles arrivent également : les éléments se déplaçant dans le champ de vision de Sasha laissent une « trace » derrière eux, semblable à la trace laissée par les objets mouvants sur une photographie prise avec un temps de pose long, et les textures des objets lui semblent s'animer, comme si elles coulaient ou fondaient dans un mouvement infini.

Les autres sens subissent également des modifications : Sasha commence à percevoir des sons de façon inhabituelle et à ressentir plus intensément les sensations tactiles. L'hypersensibilité tactile qu'il ressent participe à un état général souvent désigné par le terme « *body high* » et s'accompagnant de sensations corporelles particulières : sensation de chaleur englobante émanant de l'intérieur du corps, frémissement ou chatouillement agréable englobant ou se déplaçant dans le corps, légères contractions musculaires localisées ou générales. Cet état donne à Sasha l'impression de ressentir son corps dans son entièreté et avec une grande intensité.

Plateau

Lorsque le pic d'intensité des effets est atteint, les effets décrits plus haut peuvent prendre une ampleur différente. Les sensations corporelles gagnant en intensité peuvent déboucher sur une puissante euphorie ou une sensation de confort maximal, mais elles peuvent également donner lieu à de véritables hallucinations telles que l'impression de fusionner avec une partie de son environnement (pour sa part Sasha sent son corps fondre et se mélanger au canapé dans lequel il se trouve) voire avec la réalité tout entière, quand l'altération des frontières corporelles s'accompagne d'un sentiment d'unité.

À ce stade de l'expérience, le pic s'est bien installé et Sasha se dit que l'expérience doit toucher à sa fin, mais en réalité il lui reste encore énormément de sensations à découvrir. C'est l'occasion pour lui de prendre conscience de la distorsion temporelle massive occasionnée par les psychédéliques : il pensait avoir fermé les yeux et être resté avachi dans son canapé pendant une éternité, mais en réalité l'expérience n'a commencé qu'il y a environ une heure et demi, il n'est resté immobile que quelques minutes.

Les effets négatifs de la montée s'étant estompés, l'envie vient à Sasha de partir à la découverte de son environnement. Comme il est accompagné d'amis, il profite alors des effets sociaux des psychédéliques. En lançant la discussion avec eux, il se rend compte que les mots lui viennent avec une aisance rare. Les traits d'esprits se multiplient alors et les convives se répondent du tac au tac avec une grande fluidité, provoquant régulièrement d'intenses crises de rire incontrôlables. La discussion s'éparpille, les idées fusent, mais finalement Sasha sent la tendance s'inverser. La conversation devient soudainement difficile à suivre : penser est devenu difficile, mener un raisonnement à son terme demande un effort considérable et il perd régulièrement le fil à cause de courtes absences. En réalité, au cours de l'expérience, il alternera régulièrement entre des épisodes d'amélioration et de diminution des capacités cognitives, ayant tantôt l'impression de penser à la vitesse de la lumière, d'acquérir une meilleure compréhension des choses ou des idées originales, et tantôt une grande difficulté à réaliser les opérations les plus simples, voire même l'impression d'avoir l'esprit parfaitement vide.

Sasha prend conscience à ce moment d'à quel point ses perceptions sensorielles sont altérées. Le filtre visuel s'est intensifié : à présent les arêtes des objets ressortent avec une incroyable netteté, de même que les textures des objets, les couleurs sont intenses et chaque micromouvement de l'environnement est retranscrit avec une grande fluidité. Le contraste entre sa vision en temps normal et son expérience actuelle lui rappelle celui qui existe entre l'image d'un écran à cristaux liquides un peu daté et un écran *LED* ultra-haute-définition. Cette vision « augmentée » s'accompagne par moments d'une grande attention à l'environnement. Il se sent soudain attiré par de nombreux détails de l'environnement : détails visuels (les rainures du bois d'un meuble, la texture particulière d'un mur blanc, etc.) mais aussi auditifs. Il entend distinctement le bruissement des feuilles d'une plante placée à l'autre bout de la pièce et les frottements infimes de vêtements contre le tissu d'un canapé. La musique lui apparaît toute différente : il lui semble entendre distinctement tous les instruments, comme s'il pouvait distinguer chaque nappe sonore composant le morceau, et il repère dans la mélodie des détails qui ne lui étaient jamais apparus. Par moment, certaines notes se répètent en boucle ou s'inversent sans que le morceau ne perde en cohérence. La musique semble également impacter d'autres éléments de sa perception : les sons extérieurs, et les couleurs de l'environnement vibrent en rythme sur la musique.

En se concentrant sur la texture d'un tapis, les mailles du textile lui apparaissent couvertes d'un motif symétrique. Alors qu'il observe avec attention celui-ci, il semble

encore gagner en complexité et se met à se déformer légèrement, se dilatant et se contractant en rythme, comme si le tapis respirait. Les éléments environnants dans son champ de vision périphérique semblent alors s'effacer. Bientôt il ne reste que le motif, englobant son champ de vision, dont les couleurs se mettent à changer au rythme de la musique, passant par toutes les nuances de l'arc-en-ciel, et dont les arrêtes se déplacent de manière fluide pour former des figures tantôt géométriques, tantôt plus figuratives comme des visages ou des animaux. La réalité semble alors disparaître : les frontières de son corps s'effacent et avec elles la distance qui le séparait du monde extérieur.

Sasha ne sait plus si les sons qu'il entend proviennent de son esprit ou de son environnement, et il ne sait pas s'il a les yeux ouverts ou fermés. C'est comme s'il flottait dans un bassin d'eau parfaitement calme dont la température imiterait parfaitement celle de son corps, de sorte qu'en se laissant flotter dedans il ne sentirait plus ni chaleur, ni froid, ni pression, ni gravité... Une grande euphorie l'envahit alors, un sentiment de paix et de bien être absolu qui lui semble dépasser tout ce qu'il a jamais connu. Une première idée lui vient à l'esprit, comme un rappel à la réalité : où se trouve-t-il actuellement ? Est-il seulement toujours là, ou son corps a-t-il réellement disparu ? Et où sont les autres ? Se sont-ils inquiétés de le voir disparaître ?

Il se souvient alors : il a pris de la psilocybine, et ce sont là des effets absolument normaux de cette substance. Rassuré, il se met à penser. À ce stade, Sasha est en train de vivre une expérience intense, aussi bien émotionnellement qu'intellectuellement, une « *peak experience* », comme on appelle parfois cela (Letheby, 2021, p. 11). Le contenu de celle-ci sera tout à fait particulier à Sasha, et dépendra aussi bien de ses tracas du moment que de ses croyances métaphysiques ou encore de ses peurs les plus profondément enfouies. Plusieurs descriptions d'expériences de ce type ont déjà été cités au chapitre 5.1, je ferai donc l'impasse sur cette partie de l'histoire.

Revenu à ses esprits, Sasha sent le contrôle de son corps lui revenir et ouvre péniblement ses paupières à l'intérieur desquelles semblent toujours imprimés des motifs géométriques vivaces. Il se sent envahi d'un immense sentiment de compassion alors qu'il retrouve la présence de ses amis.

Descente

Après quelques épisodes hallucinatoires moins intenses, Sasha commence à sentir qu'il « revient sur terre ». Il remarque à nouveau les effets physiologiques qui avaient

débuté en montée mais auxquels il ne prêtait plus attention durant le pic : un manque général de tonus, des vagues de froid accompagnées de tremblements et un rythme cardiaque irrégulier. Il lui vient l'idée de manger et boire pour combattre cet état de fatigue.

Il se rend alors compte que ses sens de l'odorat et du goût ont également été altérés. Les mets salés semblent lui attaquer le palais comme une gorgée d'eau de mer, les mets sucrés sont particulièrement doux, certains arômes ressortent fortement alors que d'autres lui semblent inexistantes. Il goûte un jus de fruit qui lui semble être une explosion de fruits tropicaux, tandis qu'un gâteau lui semble avoir un goût de carton en dépit de son odeur de vanille très distincte.

Tandis qu'il se restaure, les effets se font progressivement de moins en moins intenses. Malgré que les couleurs s'estompent et que l'euphorie disparaît, il conserve une certaine tranquillité. Alors que la fatigue commence à s'installer (une vraie fatigue cette fois, non plus un effet sédatif de la psilocybine), il se sent apaisé et confiant, comme s'il avait trouvé une nouvelle énergie aujourd'hui, ou qu'il avait rajeuni. Cet état, surnommé « *afterglow* », pourra perdurer pendant plusieurs jours, voire plusieurs semaines après l'expérience. Après quoi, l'humeur de Sasha reviendra probablement à un niveau standard, à moins que son expérience introspective lui ait permis d'en tirer un bénéfice à plus long terme, comme ceux qui ont été observés chez certains patients de thérapies psychédéliques.

Voici qui conclut cette description, et ce chapitre définitoire de l'expérience psychédélique. Le sixième et dernier chapitre développera finalement les questions qui m'ont intéressé dans ce travail, ayant trait à la nature de l'illusion de la conscience et à ce que les recherches psychédéliques peuvent nous apprendre à son sujet. Après une introduction (chap. 6.1) présentant les questions qui seront abordées et les enjeux qu'elles représentent pour la philosophie, le chapitre 6.2 établira un lien entre la méthode hétérophénoménologique décrite par Daniel Dennett (exposée au chap. 1.1) et les travaux issus des recherches psychédéliques. Enfin, le chapitre 6.3 exposera certaines idées développées dans ces travaux (en particulier par Chris Letheby dans son ouvrage *Philosophy of Psychedelics* (2021)) et en tirera des conclusions favorables au modèle illusionniste de la conscience exposé dans la première partie de ce travail et précisé dans la seconde.

Chapitre 6 Repenser la conscience grâce aux psychédéliques

6.1 Enjeux et méthode des recherches psychédéliques

De la même façon qu'observer les relations entre le fonctionnement cérébral d'une personne et l'état anatomique de son cerveau permet d'en apprendre plus sur la structure de celui-ci⁵¹, l'observation « d'états non-ordinaires » de conscience nous donne des informations « à propos de la structure et de la fonction de l'esprit ordinaire. »⁵² Étudier les EMC, c'est toujours étudier la conscience.

Pour ne citer qu'un exemple, l'observation de cas d'illusion de Capgras (dans laquelle ceux qui en souffrent affirment qu'un proche qu'ils reconnaissent pourtant a été remplacé par un sosie parfait) nous a permis de comprendre que le mécanisme cérébral qui permet la reconnaissance d'une personne n'est pas uniquement perceptif ou mémoriel, mais également affectif. Ceci souligne l'interdépendance de nombreuses fonctions dans le cerveau, œuvrant de concert pour produire l'expérience.

Dans le cas d'un EMC, le nombre et la complexité des perceptions anormales peut rendre la tâche de remettre dans l'ordre causes neurologiques, fonctions cognitives mobilisées et attributs phénoménaux difficile, c'est pourquoi le recours à des outils tels que l'échelle *IID-ASC* est important : premièrement, la caractérisation qu'elle offre permet de décomposer un phénomène complexe en unités plus simples qui pourront être analysées séparément des autres ou mises en relation avec différents EMC (par exemple, on peut mettre en lien le facteur « *disembodiment* » de l'*IID-ASC* avec d'autres travaux sur les expériences de sortie hors-du-corps (« *out-of-body experiences* »).

Deuxièmement, le travail de comparaison entre les différents dispositifs d'évaluation ou entre les différents *items* et facteurs d'un même dispositif permet de définir l'influence mutuelle de ces derniers et leur prévalence au sein d'un type d'EMC. Le phénomène se trouve ainsi de mieux en mieux circonscrit à mesure que ses traits

⁵¹ C'est par exemple de cette façon qu'au XIX^{ème} siècle, Wernicke et Broca découvrirent les régions du cerveau en charge, respectivement, de la compréhension et de la production du langage (Nutt, 2021, p. 98).

⁵² Graham, G. & Stephens, G.L., éd., *Philosophical psychopathology*, 1994. Cité dans (Letheby, 2021, p. 11).

112 • Repenser la conscience grâce aux psychédéliques

définitoires sont extraits de la masse d'effets anecdotiques ou épiphénoménaux, permettant de perfectionner plus avant ces outils qui, en retour, nous offriront une compréhension encore plus fine du phénomène, et ainsi de suite. Ainsi, l'*IID-ASC* est une évolution du *5D-ASC* elle-même une évolution de l'*OAV*, qui est sensée mieux rendre compte de la complexité des EMC et s'applique en tout cas très bien à l'expérience psychédélique (Studerus, Gamma, & Vollenweider, 2010).

Entre autres phénomènes que l'*IID-ASC* participe à explorer, l'expérience spirituelle a une place importante. Les dimensions spirituelles de l'EMC révélées par de nombreux témoignages sont englobées par les facteurs de l'*IID-ASC* (il s'agit des trois premiers facteurs, ainsi que l'*insightfulness*). Ce dispositif permet par conséquent de transposer ces données relevées par une étude qualitative de l'EMC (basée sur les comptes rendus d'EMC) dans un format propice à une utilisation quantitative de ces données. En d'autres termes, cela permet d'étudier scientifiquement l'aspect spirituel des EMC. L'opportunité d'expliquer un phénomène qui semble résister aux explications naturalistes est déjà alléchante en soi, mais elle l'est d'autant plus que son aspect spirituel semble jouer un rôle important dans les bénéfices thérapeutiques de l'expérience psychédélique. Nous avons vu au chapitre 5.2 que l'expérience psychédélique pouvait apporter des bénéfices thérapeutiques, dont la force est corrélée avec la survenue d'une expérience à caractère spirituel durant la thérapie psychédélique (Barrett & Griffiths, 2017). Or si expérience psychédélique et expérience mystique sont liées, il est nécessaire que l'on puisse comprendre comment elles le sont. Il s'agit là d'un enjeu médical, et donc éthique, majeur pour les recherches psychédéliques.

Cette recherche est heureusement rendue possible par la facilité de reproduction de l'expérience psychédélique. En effet, le fait qu'il soit possible de provoquer cet EMC « sur commande » en administrant simplement une substance à un sujet volontaire, le rend plus facile à étudier que n'importe quel autre. Ce fait garantit l'accès à quantité de données de première importance, à commencer par celles que nous donnent les techniques d'imagerie cérébrale aujourd'hui utilisées dans le cadre des recherches psychédéliques. Or les effets observés de l'expérience psychédélique, en plus d'être particulièrement spectaculaires, semblent devoir amener une réflexion sur notre conception de la conscience habituelle. Les changements drastiques de perception de Soi qui accompagnent cette expérience, et les effets de ceux-ci à long terme sur la personnalité sont remarquables et suggèrent que le Soi est bien plus fluide que ce qu'il n'apparaît.

La phénoménologie de l'expérience psychédélique, par son caractère spirituel, voire mystique, représente également un objet d'étude de premier plan. D'autant plus que nos connaissances actuelles suggèrent qu'elle jouerait un rôle prépondérant dans le potentiel thérapeutique des psychédéliques. D'aucuns pourraient voir dans ce fait un argument en faveur de l'irréductibilité de la conscience phénoménale. En effet, le maëlstrom de sensations et d'émotions que représentent l'expérience psychédélique en font un état dont l'aspect « qualitatif » ressort vivement. Pour beaucoup, l'expérience est « ineffable » (Letheby, 2021, p. 39), sa beauté et la profondeur de son contenu ne pouvant être épuisées par le langage.

Les tentatives d'évocation de cette réalité qui dépasserait l'entendement font écho de rencontres avec le Divin et d'extases mystiques, donnant l'impression d'une irréconciliabilité entre la science et le monde au-delà des « portes de la perception ». La réticence de certains à approcher ce phénomène « à la troisième personne », et le débat existant aujourd'hui sur le rôle joué par l'aspect « qualitatif » de l'expérience dans ses bénéfices thérapeutiques, rappellent les discussions entourant le problème difficile de la conscience. Mais le lien entre les discussions du domaine des recherches psychédéliques et les problèmes de philosophie de l'esprit ne s'arrête pas à cette simple ressemblance.

Dans les deux prochains chapitres, je renforcerai ce lien et en tirerai des conclusions. Le chapitre 6.2 abordera la façon dont la méthode hétérophénoménologique est déployée dans le champ des recherches psychédéliques. Nous y verrons que ni les données qualitatives, ni les données quantitatives obtenues par les recherches psychédéliques, ne nous permettent de rendre compte parfaitement du phénomène. Néanmoins, elles nous permettent de formuler de nouvelles hypothèses sur le fonctionnement de la conscience. Ensuite, le chapitre 6.3 permettra d'explorer le lien établi par Letheby entre la caractérisation de l'expérience spirituelle fournie par l'*IID-ASC* et l'altération de la représentation de Soi vécue sous l'influence des psychédéliques, et les implications de celui-ci pour le modèle de la conscience de Dennett.

6.2 Renvoyer la balle aux sciences cognitives

Le chapitre 5.2, présentait un compte rendu censé refléter l'effet que cela fait de vivre une expérience psychédélique. J'ai écrit celui-ci avec un scrupule méthodologique particulier : éviter les descriptions romancées, donner suffisamment de détails ou, lorsque

114 • Repenser la conscience grâce aux psychédéliques

ceux-ci ne suffisent pas, utiliser des métaphores aussi évocatrices que possible, et éviter l'interprétation des faits en termes issus de la psychologie naïve.

Ce compte rendu avait un double objectif : d'une part il me paraissait important de décrire certains éléments de la phénoménologie de l'expérience psychédélique afin de la rendre compréhensible au lecteur. Dans le même temps, en donner une description rigoureuse aura, je l'espère, participé à la démystifier. En effet, il n'est pas rare de lire des témoignages qui exagèrent l'intensité de l'expérience, ou la portée de celle-ci. Leur extravagance ne permet pas, je pense, de se représenter la réalité, ni ne lui rend honneur.

D'autre part, l'exigence de détail qu'il poursuivait devait permettre de relever les effets « de fond » de l'expérience psychédélique ; ceux qui, sans être toujours définitoires de celle-ci, y participent de façon substantielle, mais demeurent souvent absents des témoignages. Ces effets sont pourtant tout à fait notables, participent certainement à la qualité de l'expérience et n'ont rien d'ineffable. Mais une telle description peut-elle réellement approcher d'une description de l'effet que cela fait de vivre une telle expérience ?

Dennett affirme que les *qualia* sont en réalité un complexe de dispositions réactives. D'après cette conception, une expérience telle que nous nous la représentons (« voir du rouge », « sentir le vent caresser notre visage », « entendre quelqu'un appeler notre prénom », *etc.*) n'est pas définie par une « qualité » particulière mais par l'ensemble des dispositions qui la caractérisent, ce qui inclut aussi bien les perceptions que les jugements qui en découlent, mais également une foule d'états conscients et non-conscients qui peuplent notre cerveau au moment où cette expérience est vécue. Nous ne voyons jamais simplement « du rouge », nous voyons du rouge qui nous rappelle tel souvenir et s'accompagne de tel autre percept (l'odeur d'une rose rouge, le goût d'une pomme rouge...), ou évoque une infinité de sentiments particuliers... mais aussi, nous voyons du rouge présentant un certain contraste, une certaine texture, tandis que nous sommes dans un certain état de veille ou de somnolence, de faim ou de satiété, *etc.*

Décrire une telle expérience par le simple terme de « rouge » est réducteur, mais pas dans le sens où cette description omettrait une qualité particulière, intrinsèque voire ineffable de celle-ci. La description de l'expérience consciente est réductrice parce que ce qui est décrit est la représentation simplifiée d'une réalité extraordinairement complexe. Lorsque nous concluons avoir une expérience de rouge, derrière ce jugement, pléthore

d'informations ont été « compressées » par le cerveau à cet instant t pour produire un état informé qui permette le contrôle conscient du comportement.

Dennett a fait sienne l'idée populaire dans les sciences cognitives que le cerveau capterait et utiliserait l'information grâce à une forme de codage prédictif hiérarchique bayésien, une méthode de calcul basée sur les attentes préalables (Dennett, 2017, p. 167). Ce système est organisé en une « cascade de couches dans laquelle les couches les plus hautes font des prédictions bayésiennes à propos de ce que la couche inférieure “verra” ensuite » puis, si cette prédiction est fautive celles-ci « génèrent des signaux d'erreur en réponse » menant à des révisions de la prédiction, et ainsi de suite jusqu'à ce que le système puisse identifier correctement l'objet des prédictions (Dennett, 2017, p. 168). Il s'agit donc d'un système apprenant par ajustement de ses erreurs de prédictions permis par le feedback de ses couches inférieures (Dennett, 2017, pp. 168-169). Le cerveau crée en continu des modèles d'anticipation qui ne sont révisés que par les signaux d'erreur, qui peuvent se faire très rare lorsque nous évoluons dans un environnement familier (Dennett, 2017, p. 169). Les prédictions des niveaux les plus hauts deviennent alors notre « réalité par défaut » (Dennett, 2017, p. 169).

Les informations qui comptent pour le contrôle de notre comportement sont cataloguées dans notre ontologie grâce aux prédictions bayésiennes génératrices de possibilités (« *affordances* ») (Dennett, 2017, p. 356) : face à un objet pour notre attention, notre cerveau produit des attentes au sujet des interactions que celui-ci serait susceptible d'entraîner, dont la projection sur cet objet lui confère des propriétés attendues. Lorsque ces prédictions sont validées, l'objet est ensuite interprété comme possédant lui-même ces propriétés (Dennett, 2017, p. 357).

Ces propriétés ne sont que le « complexe de dispositions » que notre cerveau s'attendait à adopter face à l'objet, et ne devraient pas être considérées comme intrinsèques à celui-ci (Dennett, 2017, p. 357). Elles ne sont par conséquent pas les causes de nos expériences, ce qui *provoquerait* une représentation, mais en sont l'objet intentionnel, décrit dans des termes utilisables par l'agent pour le contrôle du comportement. Nous interprétons nos « jugements, convictions, croyances, inclinations » au sujet de nos expériences « comme provenant d'une propriété subjective (un *quale*, dans le jargon de la philosophie) qui est la source de notre jugement », alors que c'est exactement l'inverse qui se produit : c'est « notre capacité à décrire » cette expérience, « notre jugement, notre tendance à faire les affirmations que l'on vient de faire, et nos réactions émotionnelles »

116 • Repenser la conscience grâce aux psychédéliques

(Dennett, 2017, pp. 358-359) face à celles-ci, qui nous convainquent qu'il y a une expérience subjective.

Cette capacité nous a permis de développer au cours de l'évolution une image manifeste particulière du monde qui ne correspond pas à la réalité scientifique qui lui est sous-jacente. Aussi est-il vain de chercher à expliquer notre monde intérieur « à la première personne » – il nous faut adopter une posture à la troisième personne, telle que celle de la science. Néanmoins la science, pour fonctionner, a besoin de données, et ces données sont justement fournies par les descriptions d'états vécus en première personne...

Nous ne pouvons donc pas nous passer des données, mais plutôt que de tenter d'en tirer une science à la première personne, nous pouvons les exploiter à la troisième personne en nous servant des comptes rendus que les sujets d'expériences peuvent nous donner. En faisant cela, nous nous situons déjà du côté de la méthode hétérophénoménologique qui, comme nous l'avons vu au chapitre 1.2, permet effectivement de rendre compte de l'intégralité des données disponibles au sujet de l'expérience vécue par un individu, y compris les données accessibles uniquement par celui-ci. L'étape suivante de cette méthode sera d'essayer de tirer un maximum d'infos de ces comptes rendus pour ne pas rester limité par la posture intentionnelle que ceux-ci nous poussent à adopter (consistant à interpréter leur contenu en termes d'états intentionnels).

Pour dessiner une hétérophénoménologie approfondie sur la base de comptes rendus d'expériences, on pourrait chercher à décomposer les apparentes propriétés intrinsèques de ces expériences, en tentant d'interpréter les dispositions qui se cachent derrière de façon plus restreinte. Il s'agira toujours d'identification de ces expériences par leur contenu (Dennett, 2017, p. 367), mais le contenu ainsi désigné aura déjà perdu un degré d'abstraction, et donc d'indétermination ou d'imprécision (Dennett, 2017, p. 367).

Ce qui nous amène, enfin, à ma description de l'expérience psychédélique. J'ai tenté, dans celle-ci, de mettre en avant certaines des dispositions qui forment donc la toile de fond de l'expérience psychédélique et coexistent avec ses effets les plus spectaculaires, décrits dans des termes intentionnels difficiles à interpréter de prime abord (tels que les effets

relatifs au « sentiment d'unité », ou à un « état de béatitude »)⁵³. Ces dispositions, sans nécessairement figurer explicitement dans la « version finale » (« *final draft* »)⁵⁴ de l'expérience, ont néanmoins contribué à sa production et peuvent donc en retour contribuer à « décompresser » ce récit, c'est-à-dire à le rendre plus riche, plus vivant, plus proche de la réalité. J'ai également épuré mon compte rendu d'autant de descriptions abstraites et d'interprétations basées sur des croyances et dispositions personnelles que possible (tout en conservant quelques métaphores qui, bien qu'abstraites par essence et issues de mon imagination, ont été réfléchies pour être aussi parlantes que possible), leur préférant des descriptions brutes bien plus informatives.

En faisant cela, je n'ai rien fait de plus qu'un travail de description, avec un scrupule d'exhaustivité afin que mon objet puisse être conçu plus facilement, ce qui me semble un effort pédagogique nécessaire lorsque l'on cherche à décrire une réalité éloignée de notre expérience habituelle du monde. Ma description ne saurait être exhaustive, puisque chaque expérience est une « collection unique de dispositions fonctionnelles » impossible à faire tenir dans un « court catalogue d'effets » (Dennett, 2012, p. 243). Néanmoins elle permet de rendre compte d'un phénomène de façon suffisante pour avancer à l'étape suivante consistant à l'expliquer et, surtout, elle nous éloigne des idées reçues sur l'ineffabilité de cette expérience ; s'il manque des informations cruciales à une description, elles pourront encore y être ajoutées plus tard, mais, s'il semble que cette absence rend la description inadéquate, il faudra alors montrer pourquoi c'est le cas, sans en inférer que le phénomène doit être ineffable.

⁵³ En utilisant l'expression « toile de fond de l'expérience », je ne souhaite pas faire référence à l'expression « arrière-plan de la conscience » parfois utilisée pour désigner la conscience phénoménale (Block, 1995). Plutôt, j'envisage cette expression dans un sens très concret : elle désigne toute information disponible à la conscience et qui participe à l'état informé de celle-ci à un instant *t*, sans faire l'objet d'une représentation consciente (ou seulement partiellement). Ou encore : la « toile de fond de l'expérience » est « l'ensemble des contenus discriminés de la conscience ne faisant pas, ou seulement partiellement, partie de la « version finale » d'une expérience ». Par exemple : la température corporelle est un paramètre physiologique dont les fluctuations peuvent donner lieu à des représentations conscientes (coup de chaud, frisson, *etc.*), pourtant nous faisons rarement l'expérience d'une température idéale. En l'absence d'alerte lancée par les mécanismes thermorégulateurs, aucune représentation consciente de « température idéale » ne nous parvient, par contre ce facteur peut participer en filigrane à une sensation globale de confort (aux côtés de nombreux autres paramètres physiologiques – pression artérielle, (dé)contraction musculaire, état de fatigue, *etc.* – ou psychologiques – état attentionnel normal, équilibre émotionnel, *etc.*) qui, elle, fera plus souvent l'objet d'une représentation consciente.

⁵⁴ Précisons que cette « version finale » n'en est en réalité pas une, mais n'est qu'une version parmi toutes celles que le cerveau produit qui est promue temporairement par la « machine virtuelle » du cerveau à un rôle fonctionnel plus important. Elle coexiste donc avec les autres versions « en cours d'édition », et son contenu pourra subsister plus ou moins longtemps en mémoire avant d'être éventuellement remplacée par celui d'une autre version (Dennett & Akins, *Multiple drafts model*, 2008), (Dennett, 1993, pp. 148-149, 318-319).

Aux côtés de mon compte rendu détaillé (mais fictif) du chapitre 5.2, le chapitre 5.1 contenait plusieurs exemples de témoignages de personnes ayant vécu un EMC. Ces comptes rendus, dont beaucoup décrivaient des expériences psychédéliques, forment un corpus de données qui peut ensuite être analysé pour en abstraire les définitions et hypothèses qui permettront d'étendre notre compréhension du phénomène. À présent que nous possédons un corpus de données qualitatives, donc, nous pouvons passer à l'étape suivante consistant à les interpréter. Les outils psychométriques utilisés en psychologie, dont j'ai présenté un exemple adapté au contexte, l'échelle *IID-ASC*, illustrent bien ce mouvement : fondés sur les comptes rendus, ils permettent d'en extraire, après interprétation, une définition, qui pourra ensuite être révisée en la comparant à d'autres données, ou donner naissance à de nouvelles hypothèses concernant ce phénomène.

Les comptes rendus (et non les expériences elles-mêmes, comme Dennett le souligne (Dennett, 2012, p. 80) sont les données d'une science qui souhaiterait expliquer l'expérience consciente, et les outils psychométriques permettent leur exploitation. L'interprétation qui est faite des données pour l'élaboration d'échelles et de questionnaires n'est pas en elle-même une tentative d'explication de celles-ci, mais une mise en forme de données brutes enregistrées destinée à fournir une description adéquate d'une expérience, selon une posture intentionnelle. En d'autres termes, les outils psychométriques participent à élaborer un catalogue de ce qui doit être expliqué (Dennett, 2012, p. 2008) par les chercheurs en sciences cognitives et en neurosciences. En cela, ils apparaissent eux aussi comme une mise en pratique de la méthode hétérophénoménologique préconisée par Dennett, ce qui soutient en retour son affirmation d'après laquelle cette méthode est déjà celle de la psychologie expérimentale (Dennett, 2012, pp. 67-68).

Notons que si l'on considère que l'interprétation des données selon la méthode hétérophénoménologique implique une dégradation de celles-ci en raison du passage du pont de vue à la première au point de vue à la troisième personne, le recours aux outils psychométriques permet peut-être de limiter cette dégradation : puisqu'ils reposent sur un questionnaire que les sujets sont invités à remplir, le travail d'interprétation est en réalité remis entre les mains du sujet de l'expérience lui-même. Bien entendu, le questionnaire est un cadre limitatif⁵⁵ (puisque'il contient un nombre fini et déterminé d'*items*), mais il permet aussi de faciliter aux sujets le travail de conceptualisation nécessaire à la production d'un

⁵⁵ Ce cadre se doit néanmoins d'être souple, sans quoi son intérêt scientifique serait nul. En effet, si sa révision est impossible dès lors qu'un défaut lui est découvert, alors il ne satisfait pas à l'exigence de réfutabilité.

compte rendu (certains auront assurément l'impression que le questionnaire les aide à « mettre des mots » sur un aspect de l'expérience qui leur échappait), ainsi que de limiter leur tendance à « exagérer les prodiges de leur propre expérience consciente » (Dennett, 2012, p. 76). Par conséquent, les outils psychométriques permettent de respecter un certain degré de précision (les données sont transposées avec une grande fidélité) et de transparence (son influence sur les données est limitée), nécessaire au respect des données.

Ce type de dispositif s'est en tout cas révélé porteur dans le cadre des recherches psychédéliques. Les outils psychométriques ont en effet permis de délimiter des aspects de l'expérience psychédélique, rendant possible leur interprétation en termes scientifiques. En quantifiant ce phénomène complexe, ils permettent d'en tirer des données statistiques permettant l'élaboration rigoureuse de modèles explicatifs.

Certains résultats permis par ces outils mis en avant par Letheby sont d'importance dans le cadre de ce travail : en premier lieu, Letheby cite une étude ouverte explorant l'action de la psilocybine contre la dépression résistante aux traitements réalisée par Carhart-Harris *et al.* en 2018 (Letheby, 2021, p. 82), dans laquelle ces derniers ont utilisé l'échelle *IID-ASC* pour quantifier certains aspects de l'expérience psychédélique. Ce travail a établi « une corrélation significative entre les scores du facteur *USB* [une combinaison des facteurs *Unity*, *Spiritual Experience*, et *Blissful State* de l'*IID-ASC*, ndla] et une réduction durable des symptômes dépressifs » (Letheby, 2021, p. 82), mais également une corrélation plus forte encore entre une réduction des symptômes dépressifs avec les scores du facteur *Insightfulness*, ce dernier n'étant pas lui-même fortement corrélé à la survenue d'expérience de type mystique (Letheby, 2021, p. 83). Ce que ces résultats suggèrent, c'est que des expériences d'*insightfulness* sont au cœur du changement provoqué par les psychédéliques, mais en quoi exactement consiste ce type d'expérience ?

Les *items* de l'*IID-ASC* nous en disent peu, mais Letheby remarque que les résultats d'une enquête menée par Davis *et al.* en 2020 peuvent nous permettre d'approfondir notre compréhension de ce facteur. Cette étude menée auprès de personnes ayant bénéficié d'*insight* psychologique par le biais d'une expérience psychédélique montre une corrélation particulièrement importante entre expérience *insightful* et bénéfices psychologiques durables, reproduisant ainsi les résultats de Carhart-Harris *et al.* évoqués ci-dessus (Letheby, 2021, pp. 84-85), tout en décrivant l'expérience *insightful* de façon plus détaillée

par le biais de leur propre questionnaire, le *Psychological Insight Questionnaire (PIQ)*⁵⁶. Or, selon Letheby, les *items* de celui-ci décrivent tous « une forme de changement du Soi narratif⁵⁷ » (Letheby, 2021, p. 84). Ces résultats sont soutenus par plusieurs autres travaux semblables (Letheby, 2021, pp. 85-86). Ceux-ci apportent d'après Letheby un soutien à son hypothèse d'après laquelle des « changements du sens de Soi, et en particulier du Soi Narratif, jouent un rôle causal central dans la thérapie psychédélique » (Letheby, 2021, p. 86). D'autres travaux encore montrent que les psychédéliques « peuvent durablement augmenter les capacités liées à la pleine conscience telles que la flexibilité psychologique » (Letheby, 2021, p. 92), celles-ci étant également reliées à des changements dans le sens du Soi (Letheby, 2021, p. 92). Corollairement, ces résultats et les précédents renforcent également l'hypothèse selon laquelle l'expérience psychédélique se caractérise par une altération du sens de Soi, en particulier du Soi Narratif.

Enfin, une dernière batterie de données permet de relier expérience psychédélique et changement de représentation de Soi : celle issues de l'étude des « corrélats neuronaux des bénéfices psychologiques durables des psychédéliques » (Letheby, 2021, p. 92). L'utilisation de techniques de neuroimagerie a en effet permis d'observer le fonctionnement du cerveau durant l'état psychédélique, offrant des résultats divers et pas toujours répliqués (Letheby, 2021, p. 92). Néanmoins, Letheby remarque que « toutes les études par neuroimagerie qui ont investigué les corrélats neuronaux d'*effets thérapeutiques et transformatifs durables* [des psychédéliques] ont découvert des changements d'un des deux ou de chacun des deux systèmes neuronaux importants » que sont le *Default Mode Network* (abrégé *DMN*, en français « réseau du mode par défaut ») et le *Saliency Network* (abrégé *SN*, en français « réseau de la saillance ») (Letheby, 2021, p. 93). Ces deux réseaux neuronaux feront l'objet d'une description (succincte mais néanmoins suffisante à mon propos) dans le chapitre suivant, pour le moment je me contenterai de sauter aux conclusions avancées par Letheby à propos de ces données : il semblerait que des altérations de l'activité et de la connectivité du *DMN* et, dans une moindre mesure, du *SN*, sont impliquées dans « (a) des expériences aiguës d'*insightfulness* et de dissolution de l'ego, et (b) des bénéfices psychologiques durables – incluant des capacités de pleine conscience

⁵⁶ Letheby remarque qu'il n'est pas certain que le *PIQ* et le facteur *Insightfulness* de l'*IID-ASC* mesurent exactement le même phénomène, néanmoins, étant donné les résultats similaires de ces travaux et d'autres travaux reliés, nous pouvons « raisonnablement supposer qu'ils mesurent des phénomènes similaires de façon générale » (Letheby, 2021, p. 85).

⁵⁷ Le concept de Soi Narratif sera développé dans le chapitre suivant.

améliorées, des changements de personnalité positifs, et des symptômes psychiatriques réduits – résultant de l’usage de psychédéliques » (Letheby, 2021, p. 98), où (a) et (b) sont reliés à des changements de représentation de Soi parce que le *DMN* et le *SN* sont des substrats neuronaux plausibles du Soi Narratif et du Soi Minimal⁵⁸ respectivement (Letheby, 2021, p. 102).

Nous sommes passés au paragraphe précédent d’un aperçu des données quantitatives apportées par la psychologie, à un aperçu des données quantitatives apportées par les neurosciences à la même question (celle de la corrélation entre l’expérience psychédélique, l’altération concomitante de représentation de Soi durant celle-ci, et les effets thérapeutiques de la première, médiée par la seconde). Dans l’ordre, nous avons donc jusqu’à présent envisagé l’importance des comptes rendus et la question de savoir comment ceux-ci peuvent rendre compte de la phénoménologie d’un individu en cataloguant l’ensemble des croyances de celui-ci au sujet celle-là, puis nous avons vu comment la science peut se saisir efficacement de ce catalogue de données afin de commencer à expliquer le phénomène, enfin nous avons vu que les facteurs révélés par l’interprétation de comptes rendus pouvaient ensuite être reliés à des corrélats neuronaux. Mais ce passage du « qualitatif » au « quantitatif » psychologique, puis neurologique, semble ne donner que peu de réponses à ce stade de la recherche. L’on pourrait même affirmer que jusqu’à présent, tous les efforts développés par la science ont laissé le « gouffre explicatif » intact. Nous sommes laissés avec deux corps de données reflétant deux approches difficilement conciliables d’un même phénomène (une approche plutôt phénoménologique, et un approche neuroscientifique). Comment réconcilier les deux ?

Letheby propose une solution intéressante : pour lui, « il est possible d’expliquer des phénomènes psychologiques en termes neuroscientifiques, par le biais d’une *théorie neurocognitive* qui décrit les fonctions computationnelles ou de traitement de l’information des structures neuronales » (Letheby, 2021, p. 102). En d’autres termes, le passage du phénoménal au neuronal est possible en passant par une explication qui décrive en termes de fonctions cognitives la façon dont les informations représentées par le cerveau donnent lieu à une expérience donnée. La théorie en question proposée par Letheby sera abordée dans le chapitre suivant, puisqu’elle me semble pouvoir faire l’objet d’une interprétation illusionniste. Ce qui m’intéresse dans le présent chapitre, c’est de questionner l’enjeu de

⁵⁸ Le concept de Soi Minimal sera lui aussi développé dans le chapitre suivant.

cette méthode. Que doit-elle nous apporter de plus, concrètement, que la collection de données brutes et leur interprétation dans les termes cités précédemment ?

D'après Letheby lui-même, l'approche qu'il propose « ne constitue pas une solution au problème difficile de la conscience », qui restera donc – toujours d'après lui – « irrésolu » (Letheby, 2021, p. 105). En effet, bien que Letheby se range du côté des physicalistes, il reconnaît l'existence du problème difficile et le fait que celui-ci soit irrésolu, et risque de ne pas l'être avant longtemps (Letheby, 2021, pp. 37-38). Quel but poursuit alors son approche de plus humble qu'une solution au problème difficile ? Une « sorte d'explication contrastive dirigée, non pas vers la conscience elle-même, mais vers pourquoi certains aspects spécifiques des contenus conscients sont comme ils sont, et pas autrement », affirme-t-il, citant Hohwy⁵⁹, avant d'exemplifier son propos : « pourquoi [...] une lésion à l'aire corticale V5 devrait-elle mener à l'akinétopsie, ou cécité au mouvement visuel ? »⁶⁰ Et « pourquoi une activité renforcée dans l'aire 17 de Brodmann devrait-elle être corrélée à une augmentation de la vivacité de l'imagerie mentale plutôt que, disons, des sentiments de colère ou des hallucinations olfactives ? »⁶¹. D'après Letheby, citant à nouveau Hohwy, « de telles corrélations sont mystérieuses à moins que nous postulions que les structures neuronales appropriées sont spécialisées dans le traitement de différents types d'informations visuelles »⁶². Une hypothèse de ce type, comme celle proposée par Letheby, permet d'expliquer « les liens entre ces changements neurophysiologiques et phénoménologiques » (Letheby, 2021, p. 105).

Mais, dès lors que ces changements phénoménologiques et les changements neurophysiologiques correspondants sont identifiés, et que les « liens » qui les unissent sont compris, que reste-t-il à expliquer qui puisse relever du « problème difficile » ? Lorsque le « gouffre » s'étendant entre les « descriptions biologiques et psychologiques du cerveau/esprit humain » (Letheby, 2021, p. 106) se trouve enjambé par les neurosciences cognitives, que manque-t-il encore pour comprendre le lien entre cerveau et esprit ? Le travail conjoint de l'hétérophénoménologie, des sciences cognitives et des neurosciences semble être apte à résoudre l'entière des problèmes faciles de la conscience, or ceci nous

⁵⁹ Hohwy, J., *The sense of self in the phenomenology of agency and perception*, 2007. Cité dans (Letheby, 2021, p. 105).

⁶⁰ Zeki, S., *Cerebral akinetopsia (visual motion blindness) a review*, 1991. Cité dans (Letheby, 2021, p. 105).

⁶¹ De Araujo, D.B., et al., *Seeing with the eyes shut: neural basis of enhanced imagery following ayahuasca ingestion*, 2012. Cité dans (Letheby, 2021, p. 105).

⁶² Hohwy, J., *Functional integration and the mind*, 2007. Cité dans (Letheby, 2021, p. 105).

mènerait de fait, d'après Dennett, à l'explication recherchée de la conscience, qui apparaît comme un « vaste arsenal de tours » de magie (Dennett, 2012, p. 125) que nous prenons à tort pour un mystère unique. Ainsi, « dès lors que l'on prend la mesure de toutes les opérations non mystérieuses au moyen desquelles le cerveau peut créer d'innocentes "illusions de l'utilisateur", nous pouvons commencer à imaginer comment le cerveau crée de la conscience » (Dennett, 2012, p. 126).

Avec l'expérience psychédélique, dont il est question dans ce travail, nous avons l'exemple d'un phénomène qui est désormais bien connu et commence à être compris, pour lequel nous disposons déjà d'hypothèses et d'un projet méthodologique (déjà en cours de réalisation) permettant de vérifier celles-ci. Or, l'on se rend compte que ce champ de recherche particulier nous permet d'en apprendre plus sur les mécanismes de création de l'une de ces illusions de l'utilisateur : comme nous l'avons vu, les psychédéliques (de même que d'autres EMC, tels que les états méditatifs de pleine conscience) ont la capacité d'altérer la représentation de Soi, nous permettant en retour d'observer comment celui-ci est produit et maintenu par le fonctionnement cérébral.

En résumé, malgré les efforts d'observation de la conscience sous psychédéliques, ni une étude qualitative (permise par les comptes rendus d'expériences), ni une étude quantitative (permise par l'exploitation de ces comptes rendus au moyen d'outils psychométriques et leur comparaison avec d'autres corps de données), ne permettent de rendre compte de toute la complexité de cette expérience. Néanmoins, combiner ces deux approches à une théorie neurocognitive de la conscience nous permet d'organiser les faits pour en tirer une explication satisfaisante du phénomène.

Cette démarche que décrit Letheby et qui, d'après lui, anime déjà le champ des recherches psychédéliques, est en fait celle que Dennett décrivait dans *La Conscience Expliquée*, puis dans ses ouvrages suivants : une méthode qui démarre de l'hétérophénoménologie et tend vers l'explication de l'ensemble des « problèmes faciles » de la conscience, laquelle nous offrirait *de facto* l'explication de la conscience elle-même. Ainsi, d'après Dennett, Letheby, bien qu'il ne le reconnaisse pas, s'attaquerait déjà au problème difficile de la conscience. L'hypothèse particulière que ce dernier avance dans son ouvrage possède d'ailleurs une implication importante allant dans le sens de l'entreprise des illusionnistes : celle d'après laquelle le Soi est une construction du cerveau, rendue possible par l'activité du *DMN* et du *SN*. Il est temps d'en apprendre plus sur cette hypothèse, les arguments qui la soutiennent, et les conclusions que nous pouvons en tirer,

mais ce travail n'est plus du ressort de l'hétérophénoménologie. Pour aller plus loin, il nous faut « renvoyer la balle » aux sciences cognitives.

6.3 Perspective illusionniste sur le travail de Letheby

D'après Letheby, le Soi tel qu'il est conçu naïvement, le Soi « cartésien », n'existe pas (Letheby, 2021, p. 158). Une autre façon de poser cette affirmation est d'affirmer que le Soi est une illusion : une représentation trompeuse masquant une réalité sous-jacente différente de celle-ci. Letheby soutient une position éliminativiste à propos du Soi, et en cela il se rapproche de Dennett, même si leurs conceptions respectives diffèrent sur au moins un point crucial que j'aborderai dans ce chapitre : pour Dennett, le Soi est un « centre de gravité narrative » (Dennett, 1993, p. 518), pour Letheby, il est un « centre de gravité représentationnelle » (Letheby, 2021, p. 133). Après avoir exposé la position de Letheby à propos de l'origine et de la nature du Soi (elle-même assez proche de ce que défend Dennett), ainsi que ce que les psychédéliques nous apprennent sur celui-ci, je mettrai en perspective ces deux positions afin d'en dégager les différences et similitudes et, éventuellement, les réconcilier.

Dans un second temps, je tenterai d'élargir la portée des affirmations de Letheby pour en dégager une hypothèse éliminativiste concernant la conscience phénoménale, ce que ce dernier s'est refusé de faire (Letheby, 2021, p. 105). Ceci transformera de fait l'hypothèse de Letheby, issue d'une théorie cognitive solide, en argument favorable à l'illusionnisme. Enfin dans un troisième temps, je soutiendrai, à l'aide des éléments avancés précédemment à propos de la nature du Soi, une position illusionniste allant dans le sens de ce qui a été affirmé dans les chapitres 3 et 4 de ce travail, à savoir que l'illusion de la conscience tient sa force des croyances dont elle est issue.

Narrative Self, Minimal Self et Predictive Self-Binding

Letheby et Gerrans (Letheby & Gerrans, 2017) ont élaboré une « explication neurocognitive de la représentation de Soi dans le cerveau prédictif, sa disruption par les psychédéliques et les effets thérapeutiques potentiels qui s'ensuivent » (Letheby, 2021, p. 122). Afin de pouvoir dégager les implications d'une telle hypothèse pour l'illusionnisme de Dennett, il nous faut d'abord en expliquer les termes.

Avant de s’y attacher, je souhaiterais seulement souligner que les conclusions que Letheby tire de son hypothèse ont trait aux questions posées par les effets thérapeutiques prometteurs des psychédéliques. Par conséquent, une grande partie des éléments apportés par Letheby pour soutenir cette hypothèse sont issus de recherches centrées sur ces effets. Dans le cadre de ce travail, cet aspect du travail de Letheby m’intéresse moins, mais le modèle du « *predictive self-binding* » n’en reste pas moins pertinent de façon plus générale en tant que modèle du fonctionnement de la conscience. Il me semble que les conclusions tirées à propos de la conscience par ces recherches sont plus fondamentales que les conclusions ayant trait aux mécanismes des effets thérapeutiques des psychédéliques (puisque un modèle expliquant les effets d’une substance en se reposant sur un modèle du fonctionnement de la conscience doit d’abord soutenir ce dernier), par conséquent, utiliser leurs résultats spécifiques dans ce cadre plus large ne me semble pas poser problème.

Le cerveau prédictif

L’explication de Letheby (et Gerrans) repose sur la théorie de la cognition dite du « traitement prédictif » (abrégé par la suite en *PP*, pour « *predictive processing* »). Letheby décrit ainsi brièvement cette conception :

« D’après le PP, le cerveau est une machine à inférence qui construit des modèles hiérarchiques du monde et les utilise pour prédire ses prochaines entrées (“inputs”). Ces modèles sont mis à jour selon les principes de détection et de correction d’erreur. Les niveaux inférieurs modélisent des propriétés perceptuelles plus concrètes sur des échelles spatiotemporelles plus réduites, tandis que les niveaux supérieurs modèlent des propriétés de plus en plus abstraites sur des échelles plus grandes. Chaque niveau tente de prédire l’activité du niveau inférieur, et le traitement cognitif est motivé par la nécessité de réduire les erreurs de prédiction ascendantes (“upward-flowing”). » (Letheby, 2021, pp. 110-11)

Cette conception s’inscrit donc dans la même lignée que celle défendue par Dennett, et évoquée dans le chapitre 6.2 : d’après Dennett, le cerveau capterait et utiliserait l’information grâce à une forme de codage prédictif hiérarchique bayésien, une méthode de calcul basée sur les attentes préalables et la production de signaux d’erreurs lorsque celles-ci ne sont pas respectées (Dennett, 2017, pp. 167-168). Lorsqu’un signal d’erreur est généré, celui-ci doit être éteint « soit en mettant à jour le modèle, soit en agissant dans le monde pour modifier la source de l’entrée discordante » (Letheby, 2021, p. 113). Le cerveau doit donc constamment optimiser ses prédictions pour réduire le risque de signaux d’erreur dans le futur (Letheby, 2021, pp. 113-114).

Cette activité permet au cerveau de développer un « modèle génératif », c’est-à-dire un modèle permettant de générer des effets en représentant les « causes cachées » sensées

126 • Repenser la conscience grâce aux psychédéliques

entraîner ceux-ci. Ainsi, le modèle prédictif, alimenté par les *inputs* sensoriels recrée les « effets observés » en attribuant ceux-ci à des causes (les « objets, propriétés et événements qui peuplent le monde extra-crânien ») représentées par le cerveau (Letheby, 2021, p. 114).

Le modèle génératif créé par le cerveau est hiérarchisé entre des niveaux inférieurs, chargés de modeler des « régularités plus concrètes sur des échelles spatiotemporelles plus réduites, telles que des propriétés perceptuelles et des objets », et des niveaux supérieurs, encodant des « régularités plus générales ou abstraites », allant d'informations sociales ou autobiographiques permettant d'interpréter des situations et événements, jusqu'à des propriétés fondamentales telles que l'espace, le temps et le Soi (Letheby, 2021, p. 114).

Or d'après ce modèle, la perception est un processus « *top-down* », ce qui signifie que les niveaux supérieurs sont à la base du traitement et produisent les prédictions initiales, après quoi « chaque niveau dans la hiérarchie génère des prédictions d'activité dans le niveau inférieur ». Lorsqu'un *input* a été correctement prédit, celui-ci n'apportant aucune information nouvelle au modèle, il est « annulé » et ne remontera donc pas dans la hiérarchie. De cette manière, seuls les événements imprévus font l'objet d'un traitement particulier. Les processus prédictifs étant particulièrement efficaces, l'essentiel de notre expérience consciente provient des prédictions, donc de constructions produites par le cerveau et non de perceptions venues de l'extérieur. Par conséquent, notre expérience consciente s'apparente à une « hallucination contrôlée », dans le sens où elle est une production spontanée du cerveau, encadrée par un ensemble de règles (Letheby, 2021, p. 115).

Néanmoins, nous ne reconnaissons pas la nature représentationnelle de l'expérience ainsi produite : nous ne faisons pas l'expérience de nos modèles « *en tant que* modèles », mais en tant que réalité. Pour cette raison, l'expérience consciente est dite « transparente », par opposition à une représentation « opaque », car nous voyons la réalité même « à travers » les processus qui la produisent et qui nous échappent (Letheby, 2021, p. 115). La plupart du temps, notre expérience nous apparaît comme étant la réalité même, ce qui est permis en partie par le fait qu'elle tente d'émuler le plus parfaitement possible le monde extérieur, mais en certaines occasions, la nature construite de l'expérience peut nous apparaître (Letheby, 2021, p. 116).

Par exemple, face à l'illusion perceptuelle du « masque creux », constater qu'un visage qui apparaissait convexe, comme le sont habituellement les visages, est en réalité

concave, nous permet de distinguer métacognitivement⁶³ le caractère construit de l'expérience. Lorsque nous percevons que le masque est concave, cette perception contredit les prédictions du cerveau, qui s'attend, lorsqu'un objet possédant les traits d'un visage humain est perçu, à ce que cet objet soit convexe. Seulement, la prédiction de la convexité du masque se voit assigner un tel niveau de confiance que l'impression qui en résulte persiste malgré des données perceptuelles la contredisant radicalement (Letheby, 2021, pp. 116-117).

Ce type de fait montre également le filtre important qu'apposent les modèles prédictifs sur le contenu de l'expérience consciente : si les données de la perception entrent en conflit avec une prédiction, celle-ci ne se verra pas toujours révisée, voire même le sera rarement dans le cas des prédictions auxquelles sont assignées de hautes probabilités. Dans ce sens, les modèles prédictifs contraignent la perception autant qu'ils la permettent. Ils permettent de produire une expérience cohérente du monde sur la base d'*inputs* sensoriels chaotiques, mais cela « requiert de rejeter et d'ignorer (ce qui est considéré comme) inessentiel » (Letheby, 2021, p. 118).

Soi narratif et Soi minimal / Réseaux DMN et SN

Les recherches en neurosciences cognitives suggèrent que le cerveau humain est organisé en « coalitions de régions » caractérisées par une haute « connectivité fonctionnelle mutuelle », définie comme « une corrélation temporaire entre leurs niveaux d'activité », indiquant une « interaction fonctionnelle significative » (Letheby, 2021, p. 93). En d'autres termes il s'agit de réseaux formés par plusieurs régions du cerveau travaillant de concert à la réalisation de fonctions spécifiques. Deux de ces réseaux nous intéresseront dans ce travail : le *DMN* (« réseau du mode par défaut ») et le *SN* (« réseau de la saillance »), déjà évoqués dans le chapitre 6.2, dont l'activité formerait le corrélât neuronal de deux aspects du Soi (Letheby, 2021, pp. 101-102) : le Soi narratif, englobant les aspects « diachroniques » de la conscience (informations autobiographiques, sentiment d'être une entité étendue dans le temps, liée par une histoire et une personnalité) et le Soi minimal, plus simple et représentant les aspects « synchroniques » de la conscience (sentiment d'être un sujet d'expérience perceptuelle, corporelle ou affective, à un moment donné) (Letheby, 2021, pp. 50-52). En d'autres termes, le Soi minimal peut être conçu comme le « sentiment

⁶³ Perceptuellement, l'illusion reste entière, il s'agit en cela d'une illusion riche au sens où Kammerer définit ce terme.

immédiat et préréflexif de subjectivité », et le Soi narratif comme une « *construction* de Soi plus riche » (Letheby, 2021, p. 52).

L'activité dans les nœuds du réseau du mode par défaut (*DMN*) est associée à des tâches contribuant au maintien et au développement d'un Soi narratif : « jugements conceptuels à propos du caractère ou de la personnalité », « ruminations ou vagabondage mental » souvent centrés autour de « thèmes autoréférencés », tâches liées à la « théorie de l'esprit » ou encore « simulation autobiographique d'événements passés et futurs » (Letheby, 2021, p. 94).

L'hypothèse selon laquelle le « fonctionnement cohérent du *DMN* serait un substrat neuronal du Soi narratif » est supportée par les recherches psychédéliques : les régions du *DMN* contiennent beaucoup de récepteurs à la sérotonine de type 2A, sur lesquels agissent les psychédéliques. Cette activité est corrélée à la survenue chez les sujets d'expériences de « dissolution de l'ego », d'altération de la conscience de Soi, de réduction du « voyage temporel mental » (« *mental time-travel* », simulation autobiographique d'événements passés et futurs) ou encore d'*insightfulness* (Letheby, 2021, pp. 94-95).

Le réseau de la saillance (*SN*) est impliqué dans la « détection d'événements saillants », c'est-à-dire d'événements pertinents pour la gestion du comportement d'un agent étant donné ses intérêts et les objectifs qu'il poursuit et la « coordination de réponses adaptatives » à ces événements (Letheby, 2021, p. 99). Le cortex insulaire, l'un des nœuds du *SN*, remplirait des fonctions telles que « l'intéroception » (la perception d'informations relatives à l'état interne physiologique du corps), « l'empathie et les émotions conscientes », et travaillerait ensuite avec d'autres structures pour répondre à ces *stimuli*, notamment en contrôlant le passage du *DMN* au *TPN* (« *task-positive network* », actif lors de tâches spécifiques, qui représente peu ou prou le réseau opposé au *DMN*) en « allouant des ressources de traitement en conformité avec la nature et la localisation du *stimulus* saillant » (Letheby, 2021, p. 99).

Étant donné l'implication du cortex insulaire antérieur (*AIC*) dans l'intéroception et « le sentiment d'habiter ou de posséder un corps physique particulier », il a été suggéré qu'il serait le corrélat neuronal du Soi minimal : l'un et l'autre seraient liés dans la mesure où, pour déterminer correctement la saillance des *stimuli*, le *SN* a besoin d'entretenir une représentation du Soi, ses objectifs et ses intérêts. Cette hypothèse est soutenue par plusieurs travaux ; par exemple, l'étude des « pathologies de l'individualité » telles que la

dépersonnalisation a montré qu'une « hypoactivité des nœuds du *SN*, en particulier de l'*AIC* » donnait lieu chez les patients à un sentiment d'être mort ou de ne plus exister (Letheby, 2021, p. 99). Des études menées sur les effets de la psilocybine vont dans le même sens : l'expérience subjective de dissolution de l'ego et autres altérations du Soi pourraient également provenir d'une désintégration du *SN* (Letheby, 2021, p. 100), ces changements étant également corrélés avec les effets thérapeutiques des psychédéliques (Letheby, 2021, p. 101).

Les faits évoqués ci-dessus suggèrent par conséquent que les psychédéliques pourraient altérer différents aspects du Soi (le Soi narratif et le Soi minimal), via leur action sur le *DMN* ou le *SN*, ou les deux à la fois (L101-102), provoquant par là des expériences similaires (des expériences dites de « dissolution de l'ego ») bien que différentes conceptuellement et subjectivement.

Liage et déliage

D'après Letheby et Gerrans (Letheby & Gerrans, 2017), les modèles de Soi prédictifs « exécutent une fonction de *liage* cognitif : ils intègrent les informations à travers plusieurs modalités et niveaux de traitement en une représentation d'un sujet d'expérience unifié » (Letheby, 2021, p. 125).

En effet, nous ne percevons pas des agrégats de propriétés chaotiques, mais des objets possédant un ensemble de propriétés. La question de savoir comment ces propriétés sont intégrées dans une telle représentation cohérente constitue le « problème du liage »⁶⁴. L'approche *top-down* du *PP* offre une solution à ce problème (Letheby, 2021, pp. 126-127).

D'après le *PP*, le processus perceptuel commence par l'émission de prédictions par un modèle génératif. Or les modèles construits par le cerveau ont pour fonction de « rendre compte des régularités [présentes dans le monde] » (Letheby, 2021, p. 126). Aussi, pour produire une perception unifiée et cohérente, ceux-ci agrègent certains ensembles de propriétés, pourtant « traitées dans des régions corticales distinctes », en objets dotés de ces propriétés, se maintenant durablement et interagissant avec les autres objets produits par le cerveau prédictif (Letheby, 2021, pp. 126-127). Dans ce sens, le cerveau prédictif apparaît comme foncièrement « substantialiste », puisqu'il postule l'existence d'entités

⁶⁴ Ou « problème de la liaison ». Je préfère pour ma part traduire le terme anglais « *binding* » dont est issue cette expression par « liage ». Ce terme dénote mieux l'entrelacement particulier des propriétés, contrairement au terme « liaison » qui pourrait suggérer une simple relation de conjonction.

130 • Repenser la conscience grâce aux psychédéliques

particulières possédant ces ensembles des propriétés unifiées en un tout cohérent (Letheby, 2021, p. 126).

Cette ontologie substantialiste apparaît comme une « stratégie efficace pour minimiser les erreurs de prédiction », ce qui est le but premier d'un système prédictif, dans la mesure où elle permet de « détecter des régularités dans un flux complexe et bruyant d'inputs sensoriels » (Letheby, 2021, p. 127). Les modèles sont donc comme des hypothèses émises à propos du contenu de la perception, dont celles qui sont confirmées (par l'absence de signaux d'erreur) sont renforcées alors que d'autres seront abandonnées. Dans ce sens, on peut dire que le liage est un procédé d'inférence abductif (Letheby, 2021, p. 128).

D'après Letheby et Gerrans (Letheby & Gerrans, 2017), la représentation de Soi fonctionne de la même façon que la représentation d'objets décrite ci-dessus : tout comme le cerveau modélise l'existence d'objets peuplant le monde extérieur, il construit également une entité dans le monde intérieur (Letheby, 2021, pp. 127-128). L'organisme lui-même étant une source importante d'*inputs*, il est également nécessaire de donner une cohérence à ces données. Et, tout comme c'est le cas pour les objets externes, détecter et « attribuer [les régularités observées] aux actions et fluctuations de fortune d'une entité unique sous-jacente » permet d'obtenir cette cohérence. L'hypothèse du « *predictive self-binding* » (en français « liage prédictif de Soi », abrégé par la suite en *PSB*) de Letheby et Gerrans est donc un cas particulier de « liage cognitif » du type permis par le *PP*, selon lequel la représentation de Soi est une manière d'intégrer des informations en « une représentation unifiée d'une entité unique et persistante » (Letheby, 2021, p. 128).

Cette fonction de liage est réalisée par des nœuds du *SN*, qui « intègrent des informations corporelles et affectives en un modèle d'un sujet d'expérience momentané (le Soi minimal) », et du *DMN* qui « situent ces informations dans un contexte autobiographique, créant l'histoire d'un protagoniste persistant (le Soi narratif) » (Letheby, 2021, p. 125). L'interaction de ces deux réseaux permet l'intégration « verticale » de ces modèles à une entité unique qui, en état de conscience habituelle, est à la base de l'impression « d'être un sujet cohérent et durable, distinct de ses états mentaux et physiques, mais étant dans une relation de propriété avec eux » ; l'impression d'être « celui qui pense les pensées, ressent les sentiments, possède et habite le corps », « celui à qui tout cela arrive, et pour qui les événements du monde comptent » (Letheby, 2021, p. 132).

Le processus de liage prédictif de Soi agirait donc, d'après Letheby et Gerrans, comme un « centre de gravité représentationnelle », agissant comme un « principe organisateur du traitement cognitif » en régulant « l'allocation d'attention, l'attribution de la saillance, et la construction de représentations mentales », et ce à « tous les niveaux » et non seulement au niveau de la réflexion explicite comme le suggérait Dennett (Letheby, 2021, p. 133). Plus précisément, toujours d'après Letheby et Gerrans, « les représentations perceptuelles sont organisées dans un espace égocentrique ; les représentations intéroceptives sont interprétées comme des signaux d'événement pertinents pour l'adaptation [...] et les narratifs [autobiographiques] sont structurés autour des chances et perspectives d'un protagoniste. À tous les niveaux, la saillance est attribuée, l'attention dirigée, et l'information intégrée en accordance avec la pertinence [hypothétique] des informations vis-à-vis des objectifs de l'organisme » (Letheby & Gerrans, 2017, p. 9).

La relation qu'entretiennent le *SN* et le *DMN* permet l'élaboration de modèles d'attention et saillance relativement à un Soi envisagé comme une entité particulière dotée d'objectifs et d'intérêts propres, modèles qui, en retour, confirment l'hypothèse du Soi particulier en contraignant l'expérience de celui-ci (Letheby, 2021, pp. 133-134). Cette « dynamique circulaire » (Letheby, 2021, p. 135) produit et renforce à la fois les motifs qui cadrent notre expérience du monde, donnant l'illusion d'un « monde riche de sens et de significations » (Letheby, 2021, p. 134) centré autour du Soi.

Centre de gravité narrative ou centre de gravité représentationnelle ?

Le modèle du *PSB* et les faits apportés par les recherches psychédéliques qui le soutiennent poussent Letheby à soutenir une position éliminativiste au sujet du Soi (Letheby, 2021, p. 156sv). D'après lui, un modèle construit par le cerveau permet d'intégrer les informations relatives à un agent (« représentations perceptuelles, corporelles et narratives » (Letheby & Gerrans, 2017, p. 7)) à une représentation d'un Soi envisagé comme une « substance cartésienne : une entité mentale simple, indivisible et persistante » qui subit ces différents états (Letheby, 2021, p. 158). Or, aucune entité réelle ne correspond à cette représentation. Par conséquent, le Soi tel qu'il nous est représenté n'existe pas (Letheby, 2021, p. 156).

Cet argument ressemble crucialement à la position de Dennett au sujet du Soi (Dennett, 1993, pp. 531-532). Néanmoins, pour Dennett la substance cartésienne représentée agit comme un « centre de gravité narrative » (CGN) (Dennett, 1993, p. 518).

132 • Repenser la conscience grâce aux psychédéliques

Pour Letheby, il s'agit plutôt d'un « centre de gravité représentationnelle » (CGR) (Letheby, 2021, p. 133), intégrant non seulement les « processus narratifs », mais également les « processus affectifs, corporels et spatiaux de plus bas niveaux » (Letheby & Gerrans, 2017, p. 9).

Je ne souhaite pas mettre en cause le *PSB* ni les conclusions que Letheby et Gerrans en tirent. Simplement, en précisant la conception dennettienne du Soi comme centre de gravité narrative, je pense qu'il est possible de réconcilier ces deux points de vue.

Pour commencer, rappelons que Dennett ne nie pas qu'une intégration des propriétés d'un agent à un modèle intégré de cet agent en tant qu'entité unique et cohérente possédant ces propriétés est également réalisée par le cerveau à un niveau pré-réflexif, ou pré-narratif. Il affirme en effet que le « moi biologique » est déjà une abstraction, dont les frontières imaginaires permettent à un agent de distinguer ce qui relève du « moi » et ce qui relève du « reste du monde » (Dennett, 1993, pp. 513-514). Or ce type de « moi » est partagé par l'ensemble des êtres vivants sur la planète, jusqu'à « la plus humble des amibes », puisqu'il est nécessaire pour qu'un corps puisse se maintenir que celui-ci soit capable de « se distinguer d'autrui afin de *se* protéger » (Dennett, 1993, p. 513). Il existe donc bien, d'après Dennett, une forme de Soi minimal non-narratif.

Le Soi narratif est pour sa part particulier au genre humain et, bien qu'étant également un « produit biologique » (Dennett, 1993, p. 516), il se différencie de la forme minimale de Soi partagée par le reste du vivant par le fait que son contenu soit « représenté » (Dennett, 1993, p. 518). Cependant, Dennett parle plus précisément d'une représentation « en langage et en geste » (Dennett, 1993, p. 518), or je pense que cette précision est importante. Dans un sens large du terme, les propriétés intégrées au « moi biologique » peuvent être considérées comme « représentées » par l'agent, puisque la coordination des différentes parties de celui-ci ne pourrait pas avoir lieu si le système dans son ensemble ne représentait pas d'une certaine manière comme telles les propriétés nécessaires au contrôle du comportement : si même une amibe est capable de réagir à une menace, par exemple un corps étranger avec lequel elle entrerait en contact, c'est que les *stimuli* perçus à sa frontière extérieure portent une information capable d'influencer le comportement du système tout entier. Ceci est cohérent avec le modèle de Letheby et Gerrans : pour qu'une propriété soit intégrée au modèle prédictif de Soi, il faut qu'elle soit représentée par celui-ci (comme lui appartenant). Néanmoins, une représentation de ce type

peut avoir lieu de manière inconsciente, et même de manière « aveugle et ignorante » (Dennett, 1993, p. 513).

En revanche, le CGN dont parle Dennett apparaît comme une construction largement consciente, motivée par l'environnement communicationnel des êtres humains (Dennett, 1993, p. 518) (ce qui n'empêche pas que sa construction et sa consolidation soient basées sur, et viennent renforcer, des modèles qui sont, eux, largement non-conscients, y compris, par exemple, le Soi minimal). Dennett suggère en effet que le CGN provient de pratiques de rumination ou de vagabondage mental consistant à (se) raconter des histoires (fictionnelles ou non), comme les enfants le font à voix haute, destinées à étendre notre connaissance de nous-même (Dennett, 1993, p. 531).

Il semblerait donc que le CGN dont parle Dennett s'apparente spécifiquement au Soi narratif dont parle Letheby. De plus, Dennett établit bien une distinction ainsi qu'une continuité entre ce CGN et une forme minimale de Soi, également représentée par l'agent et également définitoire de l'unification de celui-ci en une entité abstraite cohérente. Par conséquent, le Soi CGN apparaît plus comme une partie, ou un cas particulier du Soi CGR, qu'une autre conception du Soi dans sa globalité.

Ce que Dennett nie en revanche, c'est que le Soi minimal soit *conscience de Soi* minimale ; le fait de posséder une représentation de Soi de ce type n'implique pas que celle-ci soit une composante de la conscience au sens commun du terme. Le CGN souligne que ce qui distingue la conscience humaine du type de conscience que l'on pourrait attribuer aux animaux provient de notre tendance à la narration de Soi. Sans celle-ci, il n'y a pas de conscience. Par conséquent, il n'existe rien de tel qu'un « arrière-plan de la conscience ».

Le fait que les psychédéliques rendent visibles les propriétés du Soi minimal en rendant opaques les mécanismes par lesquels celui-ci est produit ne signifie pas que le Soi-minimal ferait partie de l'arrière-plan de la conscience dans le sens où il serait constituant de celle-ci sans être représenté réflexivement par le Soi narratif. Au contraire, les propriétés du Soi minimal restent inconscientes jusqu'à ce que les psychédéliques (ou toute autre cause susceptible de rendre opaque de telles propriétés) les mettent en relief, motivant ainsi la formation de jugements à leur sujet. Ainsi, la disruption de la perception habituelle de Soi au sens du Soi minimal permet la conscience du Soi minimal.

Par exemple, l'altération de la perception des frontières de son corps rend visible cette composante habituellement peu représentée dans notre expérience du monde,

134 • Repenser la conscience grâce aux psychédéliques

permettant la formulation de jugements du type « il m'a semblé que je n'avais plus de corps » (Studerus, Gamma, & Vollenweider, 2010, p. 9). Mais c'est seulement lorsque ces jugements (ou les croyances desquels ils découlent) sont intégrés par le Soi narratif qu'ils peuvent être dits conscients.

Dennett ne nierait donc pas l'existence du Soi minimal ni son caractère représentationnel, mais nierait qu'il fasse partie de la conscience au sens du Soi narratif qui forme le CGN. En d'autres termes, le Soi minimal n'est conscient qu'en tant qu'il est intégré au Soi narratif, ou encore en tant qu'il est dans une interaction continue avec celui-ci (Letheby suggère pour sa part que le « modèle de Soi et les motifs d'attention et de saillance » de la conscience sont dans une relation « bidirectionnelle » de « renforcement mutuel » (Letheby, 2021, p. 134)). Il n'y a donc pas d'opposition majeure entre les conceptions de Dennett et de Letheby, et la première reste également cohérente avec les affirmations portées par l'hypothèse du *PSB*. Enfin, autant la conception de Dennett que celle de Letheby supportent l'élimination du Soi dans la mesure où il s'agit d'un concept problématique renvoyant à une substance cartésienne (une « entité simple, indivisible et persistante », « distincte de l'ensemble de ses propriétés » (Letheby, 2021, p. 153)) en réalité inexistante.

Qu'en est-il de la conscience phénoménale ?

Dans *La Conscience Expliquée*, l'élimination du Soi est, dans l'argumentation de Dennett, étroitement liée l'élimination des *qualia*. L'idée étant que l'aspect qualitatif de la conscience est inutile s'il n'y a « personne à la maison » (Dennett, 1993, p. 45) pour en profiter, et inversement il n'est pas nécessaire de postuler un Soi dont le rôle serait de faire l'expérience des *qualia* si ceux-ci n'existent pas. L'approche réductionniste de Dennett, et en particulier son modèle des versions multiples distribue la tâche du Soi et les propriétés qualitatives de l'expérience à travers le cerveau, soulignant de cette façon que l'un et l'autre sont des abstractions construites par celui-ci dans son effort de simplification à des fins de contrôle de l'information qu'il traite. Ainsi, si Letheby se veut éliminativiste au sujet du Soi, quelle serait sa position au sujet de la conscience phénoménale ?

Bien qu'il positionne son travail dans la lignée du naturalisme philosophique et qu'il adopte une conception physicaliste de la conscience phénoménale (Letheby, 2021, pp. 34-35), Letheby n'est pas ouvertement éliminativiste. Néanmoins, si l'interprétation dennettienne de son approche que j'ai esquissé au chapitre 6.2, est correcte et compte tenu

de son éliminativisme concernant le Soi, le réductionnisme au « sens faible et innocent » (Letheby, 2021, p. 108) du terme qu'il reconnaît comme sien se rapprocherait en réalité d'un éliminativisme concernant la conscience phénoménale.

D'après le *PP*, notre expérience est en réalité formée d'ensembles de propriétés émises par les modèles prédictifs du cerveau et projetées dans le monde sous forme d'objets substantialisés (Letheby, 2021, p. 126). Les substances projetées n'existent par conséquent pas réellement, et peuvent être considérées comme des illusions au sens illusionniste du terme : il nous *semble* qu'elles existent, mais il ne s'agit que d'une apparence qui cache une réalité bien différente. Or Dennett ne dit rien de plus au sujet des *qualia* : ce sont des ensembles de propriétés réifiés par le cerveau, et démystifier le procédé qui rend cette réification possible soutient et valide donc le point de vue dennettien.

En fait Letheby ne tire de son modèle que des conclusions concernant le Soi (en l'occurrence, il conclut que le cerveau produit la représentation erronée d'une entité unifiée, cohérente et persistante, sujet conscient des expériences), mais d'après ce même modèle l'expérience elle-même est en fait un ensemble de propriétés disparates représentées de façon erronée. Par conséquent, en tirant toutes les conclusions du *PP*, et en envisageant les *qualia* selon une définition dennettienne de ceux-ci, le travail de Letheby constitue un argument en faveur de l'illusionnisme, d'après lequel l'illusion de la conscience serait définie comme la représentation erronée de la conscience phénoménale en tant qu'expérience unifiée, cohérente et persistante, vécue par un sujet également unifié, cohérent et persistant.

Argument en faveur de la légèreté de l'illusion

Dans la seconde partie de ce travail, j'ai défendu l'hypothèse que l'illusion de la conscience n'est pas le fait d'une représentation incorrecte donnée par un processus cognitivement impénétrable du cerveau, mais qu'elle provient d'une mauvaise interprétation de données non-problématiques selon un processus certes systématique et probablement en partie câblé dans la machinerie cérébrale, mais néanmoins altérable par une modification des croyances. J'aimerais, pour clore ce chapitre, confronter une telle hypothèse aux idées avancées par Letheby dans son travail sur l'altération du Soi par les psychédéliques.

Comme nous l'avons vu dans la première section de ce chapitre, Letheby, adoptant une définition de la « transparence » donnée par Thomas Metzinger d'après laquelle celle-

ci « désigne une propriété des représentations conscientes, à savoir qu'elles ne sont pas expérimentées comme des représentations » mais comme la réalité elle-même⁶⁵, reconnaît que la conscience est, en temps normal, transparente (Letheby, 2021, p. 115). Par conséquent, les mécanismes qui la produisent ne sont pas représentés consciemment, et restent donc hors de portée des sujets qui ne possèdent dès lors aucun moyen pour remettre en question leur expérience. J'aimerais montrer que penser les mécanismes responsables de l'illusion de la conscience en termes d'opacité ou de transparence est plus à même de nous permettre de comprendre celle-ci, et pointe en outre vers une conception de l'illusion plus proche de celle de Dennett.

Le modèle du fonctionnement de la conscience présenté par Letheby semble en effet adéquat pour décrire le fonctionnement de l'illusion de la conscience : il offre une explication en termes cognitifs de la façon dont le cerveau pourrait trompeusement représenter la réalité, sans que nous soyons en mesure de le savoir. De plus, Letheby applique ce modèle à l'illusion du Soi (la représentation fautive d'après laquelle il existe réellement une entité particulière, unifiée, cohérente et persistante, qui soit la bénéficiaire de l'ensemble de représentations conscientes au sein du cerveau), dont on a vu qu'elle était très reliée à l'illusion de la conscience elle-même, telle que Dennett la conçoit. Le concept de transparence phénoménale pourrait donc par conséquent offrir une autre façon d'expliquer notre réticence à accepter la nature réelle de la conscience que celle proposée par Kammerer en termes d'impénétrabilité cognitive.

En effet, d'après Kammerer, l'illusion de la conscience est impénétrable cognitivement, c'est-à-dire qu'elle ne peut être expliquée qu'en termes d'architecture fonctionnelle, sans référence à des entités sémantiques (telles que les croyances) (voir chap. 3.1). Selon lui, l'illusion provient d'une représentation incorrecte donnée de façon systématique et cognitivement impénétrable par un processus perceptuel jouant en quelques sortes le rôle d'un malin génie voué à nous tromper au sujet de la nature réelle de nos expériences. L'illusion serait en quelques sortes « câblée » dans le cerveau et il serait par conséquent impossible de s'en extraire, d'où la prégnance d'intuitions anti-physicalistes concernant la conscience.

⁶⁵ Metzinger, T., *How does the brain encode epistemic reliability? Perceptual presence, phenomenal transparency, and counterfactual richness*, 2014. Cité dans (Letheby, 2021, p. 115).

J'ai proposé au chapitre 3 que le critère d'impénétrabilité cognitive n'était pas à même de déjouer une argumentation en faveur d'une théorie doxastique de l'illusion de la conscience. Pour résumer, j'avais subverti un exemple d'illusion légère donné par Kammerer afin de montrer que ce type d'illusion pouvait être particulièrement difficile à surmonter, puis donné un exemple d'illusion légère que sa proximité avec le sujet, et non sa pénétrabilité cognitive, rendrait particulièrement difficile à surmonter, afin de montrer que l'argument donné par Kammerer contre les théories de l'illusion légère échouait. En montrant qu'envisager le problème de l'illusion en termes de pénétrabilité cognitive ne permettait pas de trancher en défaveur de celles-ci, cette argumentation permettait également de remettre en question ce critère et d'ouvrir la voie à une autre théorisation du problème de la nature de l'illusion de la conscience.

D'après Letheby, la nature représentationnelle de notre expérience nous échappe car nous ne faisons pas l'expérience de nos modèles « *en tant que* modèles », mais en tant que réalité. Pour cette raison, l'expérience consciente est dite « transparente », par opposition à une représentation « opaque », car nous voyons la réalité même « à travers » les processus qui la produisent et qui nous échappent (Letheby, 2021, p. 115). L'illusion de la conscience viendrait donc de ce que les mécanismes qui produisent l'expérience sont transparents. Néanmoins, ceux-ci peuvent devenir opaques dans certaines conditions. Dans ces conditions, la nature représentationnelle de la conscience devient donc une propriété traitable par le sujet au niveau sémantique : il peut se la représenter consciemment, former des jugements et des croyances à son sujet. En d'autres termes, il lui est possible de se représenter consciemment la vraie nature de la conscience, et donc de surmonter l'illusion. Notons déjà que le fait qu'il soit possible *dans certains cas* de surmonter l'illusion de cette façon, suffit à prouver que celle-ci n'est pas entièrement impénétrable cognitivement.

Le fonctionnement des psychédéliques sur le cerveau tel qu'il est conçu par Letheby montre comment des représentations devenues opaques permettent de surpasser une illusion au niveau sémantique par la recalibration des croyances de haut niveau (« *high-level priors* »)⁶⁶ qui forgent certains modèles prédictifs (Letheby, 2021, p. 125). Il affirme en effet que l'altération du soi à l'origine des effets thérapeutiques des psychédéliques est rendue possible par le fait que ces croyances deviennent opaques, permettant au sujet de se

⁶⁶ Ou encore « *hyperpriors* », il s'agit de nos « croyances les plus fondamentales et abstraites à propos du Soi et du monde », celles-ci sont implémentées aux niveaux les plus haut du modèle génératif du cerveau et donc « largement inconscientes » en temps normal, que Letheby associe, entre autres, à la notion de catégories kantienne ou encore à la « *Web of Belief* » de Quine (Letheby, 2021, p. 119).

les représenter consciemment et de les manipuler comme n'importe quelle croyance. Au niveau phénoménologique, cela se traduit par une impression pour les sujets de gagner un accès particulier à des informations au sujet d'eux-mêmes ou de leur relation à leur environnement qui leur étaient inaccessibles auparavant (qui correspond selon l'échelle *IID-ASC* à la survenue d'une expérience de type « *insightfulness* » ou une révélation transpersonnelle renvoyant aux facteurs « unité », « expérience spirituelle » et « état de béatitude »).

Toutefois, cette position n'implique pas qu'il soit possible de changer tous les modèles prédictifs du cerveau, ce qui reviendrait à soutenir que la conscience peut se recalibrer entièrement par elle-même ! En fait, seuls les mécanismes déjà opaques ou qui peuvent être rendus opaques par certains moyens sont susceptibles de modification. Ensuite, la modification dont il est question ici n'est pas nécessairement une modification des représentations produites par le modèle en question mais plutôt une modification du rapport à ces représentations.

En effet, les représentations conscientes ne sont pas données par une seule prédiction originelle, mais par un ensemble de prédictions données par les mécanismes parallèles du cerveau hiérarchisés en cascade. C'est l'ensemble du système qui donne l'*output* représenté par le CGN, par conséquent il est possible qu'une prédiction particulière problématique soit corrigée par le système bien avant d'être représentée consciemment. Cela est rendu possible par la modulation du poids statistique conféré aux prédictions, qui détermine leur importance dans le système. Dans un système de ce type, une trop grande confiance accordée aux *outputs* d'un modèle prédictif problématique risque d'influencer négativement le système. Par exemple, la confiance maximale accordée à une prédiction du type « les visages sont convexes » (donnée par un module spécialisé dans la reconnaissance des visages et de leurs propriétés) empêche le système de percevoir un contre-exemple manifeste dans l'illusion des masques creux (Letheby, 2021, p. 116).

Mais, à côté de la modulation du poids statistique accordé à un modèle particulier, on peut aussi imaginer que d'autres systèmes pallient à cette trop grande confiance en donnant des *outputs* contradictoires à haute certitude. Dès lors, s'entraîner à surmonter une illusion pourrait consister à modifier le modèle qui l'a produite ou à entraîner un autre modèle à corriger les résultats du modèle problématique, sans que cela ne change fondamentalement le résultat : dans les deux cas, la modification de croyances de haut-niveau permet d'entraîner le cerveau à repérer et surmonter de façon systématique un

certain type d'illusion. Dans le cas de l'illusion des masques creux, toujours, des données non-publiées⁶⁷ suggèrent que « les psychédéliques diminuent la susceptibilité à cette illusion chez des sujets saints », ce que Letheby interprète comme le résultat d'une diminution de leurs croyances préalables que « tous les visages humains sont convexes », permettant un traitement plus adéquat des données perceptuelles habituellement filtrées (« *filtered out* ») par des modèles auxquels une haute confiance est accordée (Letheby, 2021, p. 165). Si ces résultats venaient à être confirmés, et que l'interprétation que leur donne Letheby est correcte, il pourrait s'agir d'une confirmation de l'hypothèse proposée ici : une altération suffisante du Soi pourrait permettre de surmonter des illusions autrement très prégnantes en rendant opaque certaines croyances associées.

Dans le cas de l'illusion de la conscience, ce qui serait modifié par une altération des croyances de haut-niveau n'est pas le fait que le modèle prédictif lie les propriétés en objets ni sa tendance à substantialiser ceux-ci, mais notre propre capacité à nous représenter consciemment les produits de ces modèles comme des représentations et non comme des réalités. Cette capacité pourrait évoluer soit en réduisant la confiance apportée aux modèles responsables de l'illusion, soit en augmentant celle apportée aux modèles qui pourraient donner des réponses contradictoires. Par conséquent, même si certains modèles particuliers restent transparents, et donc cognitivement impénétrables⁶⁸, le degré de confiance accordé par d'autres modèles ou par le système dans son entièreté aux *outputs* de ceux-ci pourrait quant à lui varier, permettant malgré tout de surmonter l'illusion.

En d'autres termes, surmonter l'illusion de la conscience serait possible sans rien changer à la manière dont le monde nous apparaît, en parvenant à modifier nos croyances de haut-niveau concernant le statut ontologique de notre propre expérience. Or Dennett n'affirme rien de plus : il ne dit pas qu'on peut imaginer « soulever le Voile de Maya » et voir le monde tel qu'il est réellement (essentiellement, en dehors de tout point de vue, « en soi ») mais qu'il est possible de surmonter nos intuitions problématiques concernant la conscience grâce à un travail de révision de nos croyances au sujet de celle-ci.

⁶⁷ Citées dans Milliere, R., *et al.*, *Psychedelics, meditation, and self-consciousness*, 2018. Cité dans (Letheby, 2021, p. 165).

⁶⁸ Ce qui restera très probablement le cas. Comme Pylyshyn le note, « être cognitivement pénétrable n'empêche pas un processus d'avoir certaines composantes impénétrables qui fassent partie de l'architecture fonctionnelle [du cerveau] », c'est même forcément le cas puisque l'ensemble des processus sont réalisés par cette architecture cognitivement impénétrable (Pylyshyn, 1985, p. 134). Par conséquent, on tombera toujours sur un processus cognitivement impénétrable en remontant à la source d'un phénomène, mais cela ne signifie pas que ce processus soit impossible à interpréter au niveau sémantique.

Il nous reste maintenant à voir comment atteindre cette opacité ou en tout cas comment permettre la modification de notre rapport aux représentations présentées par les mécanismes cognitifs du cerveau. Bien entendu, les psychédéliques ne sont pas le seul moyen en vue de cette fin. Le passage de la transparence à l'opacité est, en fait, une prise de recul du cerveau par rapport à son propre fonctionnement (et, corrélativement, une prise de recul du sujet par rapport à son expérience habituelle)⁶⁹. Dans ce sens, toute pratique qui permet de distancier le sujet de ses croyances pourrait permettre un tel travail. Dans le cas de la conscience, cette possibilité est consistante avec une hypothèse concernant la méditation selon laquelle les bénéfices de celle-ci (semblables à ceux des psychédéliques, bien qu'ils demandent un travail différent pour être atteints) proviendraient du fait qu'elle permette aux sujets de prendre du recul sur leurs affects, adopter une attitude non-réactive vis-à-vis de ceux-ci rendue possible par une désidentification au contenu du Soi narratif (Letheby, 2021, pp. 88-90). La méditation apparaît donc également comme un moyen de rendre opaques certains processus transparents du cerveau prédictif.

Dans le chapitre 3.2, je proposais également qu'une amélioration de la compréhension des mécanismes sous-jacents à une illusion permettraient *de facto* de surmonter celle-ci : l'idée étant que comprendre un objet entraîne une « mise à distance » de celui-ci qui permet en principe une altération des croyances qui lui sont reliées allant dans le sens d'une plus grande complexité. Par exemple, comprendre le fonctionnement d'un ordinateur permet de résister à l'illusion de l'utilisateur puisque cela permet de se représenter les icônes affichées à l'écran non plus comme de véritables fichiers et dossiers, mais comme des *représentations* de fichiers et dossiers, renvoyant au fonctionnement de l'indexation des données stockées dans la mémoire de la machine, qui sont en réalité des contenus réalisés physiquement par des champs électro-magnétiques à certains endroits précis des plateaux du disque dur...

Si mon approche du problème de l'illusion dans les chapitres 3 et 4 est correcte, et si celle-ci reste cohérente avec les concepts de Dennett, les points évoqués ci-dessus suggèrent que les conclusions de Letheby se prêtent effectivement bien à une relecture

⁶⁹ Notons que cela rejoint ce que j'affirmais dans mon commentaire de Kammerer au sujet du critère de proximité. Un état de proximité maximale avec un phénomène équivaldrait à une transparence totale de ses mécanismes (dans ce cas, le phénomène est totalement confondu avec la réalité), au contraire un état de proximité minimal équivaldrait à une opacité totale des mécanismes (dans ce cas, la représentation serait non seulement correctement expérimentée comme telle, mais ses mécanismes seraient devenus parfaitement opaques et donc entièrement traitables au niveau sémantique – j'ignore si un tel état de proximité minimale existe.

dennettienne. Dans la théorie défendue par Dennett, tout ce qui est conscient est cru, et inversement, donc la question ne se pose pas de savoir si une propriété particulière de la conscience est pénétrable cognitivement ou non, puisque par définition ce qui est conscient est représenté sémantiquement, donc pénétrable cognitivement⁷⁰. En revanche, il est possible que certains processus produisant ces propriétés soient transparents, et que celles-ci soient donc représentées en temps normal de manière rigide et résistante à la réflexion, entraînant la formation de croyances tenaces à leur sujet.

Par conséquent, le modèle défendu par Dennett (le MVM, et sa conception du Soi comme CGN), envisagé comme un modèle de type *PP* tel que décrit par Letheby, permet de rendre compte de la prégnance des intuitions réalistes au sujet de la conscience, et par extension la réticence aux thèses illusionnistes, sans en appeler à l'action d'un mécanisme unique produisant de telles intuitions, mais en les expliquant par la présence de modèles prédictifs générant un centre de gravité narratif (une abstraction découlant d'une simplification de l'activité réelle du système) altérable au niveau sémantique par l'acquisition de croyances à leur sujet qui correspondent mieux au fonctionnement réel du cerveau. Un processus de ce type, et la possibilité qui en découle d'altérer le CGN, suggère que l'illusion de la conscience – définie comme la représentation erronée de la conscience comme étant une expérience unifiée, cohérente et persistante, vécue par un sujet également unifié, cohérent et persistant – serait une sorte d'illusion doxastique légère, définie comme un croyance découlant de représentations incorrectes données par des mécanismes cognitifs habituellement transparents mais pouvant néanmoins être rendus opaques sous certaines conditions.

⁷⁰ Cela est confirmé par l'affirmation de Dennett selon laquelle le MVM est un « opérationnalisme de la première personne », niant « par principe la possibilité d'avoir la conscience d'un stimulus en l'absence de la croyance du sujet en cette conscience » (Dennett, 1993, p. 170). Voir aussi la critique adressée à Levine par Dennett, évoquée au chapitre 1.2. Notons également que Letheby affirme également une telle position, affirmant pour sa part qu'une telle conception est « congénitale au *PP* » (Letheby, 2021, p. 145)

Conclusion

La portée de ce travail est certes limitée, mais j'espère que celui-ci aura au moins participé à convaincre de la pertinence d'une approche illusionniste de la conscience, et de l'intérêt des recherches psychédéliques pour alimenter cette dernière. J'aurais aimé aller beaucoup plus loin sur les sujets abordés, mais aussi aborder bien d'autres sujets. Outre une enquête sur les autres états modifiés de conscience, c'est surtout un questionnement éthique qui manque à ce travail. Une démarche éthique guide en effet le travail de Dennett (qui questionne l'impact de l'élimination de la conscience sur la notion de libre-arbitre (Dennett, 2004)), ainsi que celui de Letheby (dont l'ouvrage *Philosophy of Psychedelics* a traité avant tout à une question éthique soulevée par les recherches psychédéliques). Croiser ces questionnements aurait certainement enrichi ce travail d'une dimension supplémentaire, mais en aurait trop largement dépassé le cadre. Je laisserai donc pour l'heure ces pistes de côté et me contenterai de donner une conclusion aux contenus qui sont parvenus jusqu'à cette « version finale » de mon enquête.

Après avoir rappelé les concepts nécessaires à la compréhension de l'approche illusionniste de Dennett, et après avoir défendu et approfondi celle-ci en m'appuyant sur les travaux de Kammerer, la troisième et dernière partie de ce travail présentait sous forme d'analyse de cas une mise en application des concepts abordés dans les deux premières. Le chapitre 5 aura permis au lecteur, je l'espère, d'appréhender l'expérience psychédélique et de comprendre les enjeux d'une étude scientifique et philosophique de ce phénomène.

Après cette introduction nécessaire, le chapitre 6 constituait le cœur de cette analyse de cas, dans laquelle j'entendais tirer des conclusions d'une lecture croisée de Letheby et Dennett. J'y ai d'abord établi un lien entre la méthode hétérophénoménologique décrite par Dennett et les travaux issus des recherches psychédéliques. J'en ai conclu que les enjeux et la méthodologie des recherches psychédéliques tel qu'ils sont mis en perspective par Letheby vérifient l'affirmation de Dennett selon laquelle la méthode hétérophénoménologique est *de facto* d'application dans les domaines scientifiques.

Ensuite, j'ai esquissé la théorie neurocognitive proposée par Letheby et en ai tiré des conclusions favorables au modèle illusionniste de la conscience exposé dans la première partie de ce travail et précisé dans la seconde. En résumé : Letheby tire des données qu'il analyse dans son ouvrage que le Soi tel qu'il est conçu naïvement, le Soi

« cartésien », n'existe pas (Letheby, 2021, p. 158). Or, cela revient en réalité à affirmer que le Soi est une illusion (au sens illusionniste du terme) : une représentation trompeuse masquant une réalité sous-jacente très différente de cette représentation. Letheby soutient donc une position éliminativiste à propos du Soi qu'il est possible de rapprocher de celle de Dennett.

La théorie de Letheby m'a également permis de renforcer la position illusionniste élaborée dans la deuxième partie du travail, d'après laquelle l'illusion de la conscience tient sa force des croyances dont elle est issue, qui proviennent elles-mêmes des modèles prédictifs du cerveau, et peuvent être rendues opaques (et donc accessibles consciemment) dans certaines conditions. L'expérience psychédélique semble permettre d'obtenir ces conditions, puisque cet état modifié de conscience semble s'accompagner d'une altération des croyances de haut-niveau, notamment celles qui forment le Soi narratif.

Cet état permet par conséquent une recalibration de ces croyances qui serait à l'origine des bénéfices thérapeutiques exceptionnels des psychédéliques, mais pourrait également permettre l'acquisition de nouvelles croyances métaphysiques au sujet du monde ou de la conscience elle-même. La possibilité d'une altération des croyances de haut-niveau formant la structure du Soi, et les effets qui en découlent, sont suffisants pour affirmer le rôle que jouent celles-ci dans la formation de l'illusion de la conscience, mais cette position pourrait trouver une confirmation supplémentaire à condition que l'on puisse montrer que les psychédéliques peuvent faciliter explicitement l'acquisition d'intuitions illusionnistes. Or, cela pourrait bien être le cas, si l'on considère que certaines croyances reliées à l'usage des psychédéliques ainsi qu'aux pratiques méditatives sont proches de doctrines philosophiques bouddhistes (Letheby, 2021), qui ne sont elles-mêmes pas sans lien avec l'illusionnisme (notamment le concept d'*anatta*, ou « non-Soi »).

Enfin, ce travail poursuivait également un objectif transversal : soutenir qu'une étude scientifique de la conscience phénoménale (en tant qu'elle est une illusion), détachée des préjugés cartésiens, est porteuse, et que celle-ci pourrait bénéficier des résultats des recherches psychédéliques. Celles-ci permettent en effet d'affiner notre connaissance du fonctionnement de la conscience. Malheureusement, leur approche subit les mêmes préjugés de la part de ceux qui pensent l'expérience psychédélique irréductible, que l'approche éliminativiste de la conscience de la part des qualophiles. Or, rejeter l'étude de l'expérience psychédélique en raison de sa présumée « ineffabilité », tout autant que s'accrocher aux vieux mythes de la conscience, revient à décourager la science d'enquêter

sur un sujet parce que celui-ci *semble* échapper aux règles naturalistes. Pour éviter de répéter l'erreur des vitalistes, il est nécessaire de surmonter les intuitions dualistes et de mener l'entreprise scientifique à son terme.

En réalité, lorsque l'on s'y attaque, l'expérience psychédélique fait un objet de science plutôt docile et en tout cas très prometteur, si l'on en croit Letheby (Letheby, 2021, p. 33). Une fois surmontée l'intuition défaitiste d'après laquelle un phénomène aussi riche que la conscience ou l'expérience psychédélique ne peut se laisser approcher par la science, nous découvrons alors à quel point nous nous sommes trompés à leur sujet. Inviter le lecteur à surmonter cette intuition était un objectif que poursuivait Dennett dans *La conscience expliquée*, et c'est dans cette lignée que j'ai souhaité inscrire mon travail.

C'est d'ailleurs là ce qui donne son sens au titre de ce travail. Le terme « braver » réfère à l'expression anglaise « *facing up to* », reprise à Chalmers⁷¹ par Frankish dans son article programmatique de l'illusionnisme « *Illusionism as a theory of consciousness* » (Frankish, 2016)⁷², tandis que les « portes de la perception » sont une métaphore des limites de la perception humaine du monde, reprise par Aldous Huxley au poète William Blake dans son ouvrage programmatique de la culture psychédélique, *Les portes de la perception* (1954). Braver les portes de la perception, c'est donc se dresser contre les intuitions introspectionnistes refusant de céder du terrain à une approche « à la troisième personne » des phénomènes, en reconnaissant l'imperfection de la perception humaine et la contingence de notre expérience consciente – un attitude qui constitue, je pense, le cœur du projet illusionniste.

⁷¹ Dont l'article introduisant le concept du problème difficile de la conscience s'intitule « *Facing up to the hard problem of consciousness* » (Chalmers, 1995).

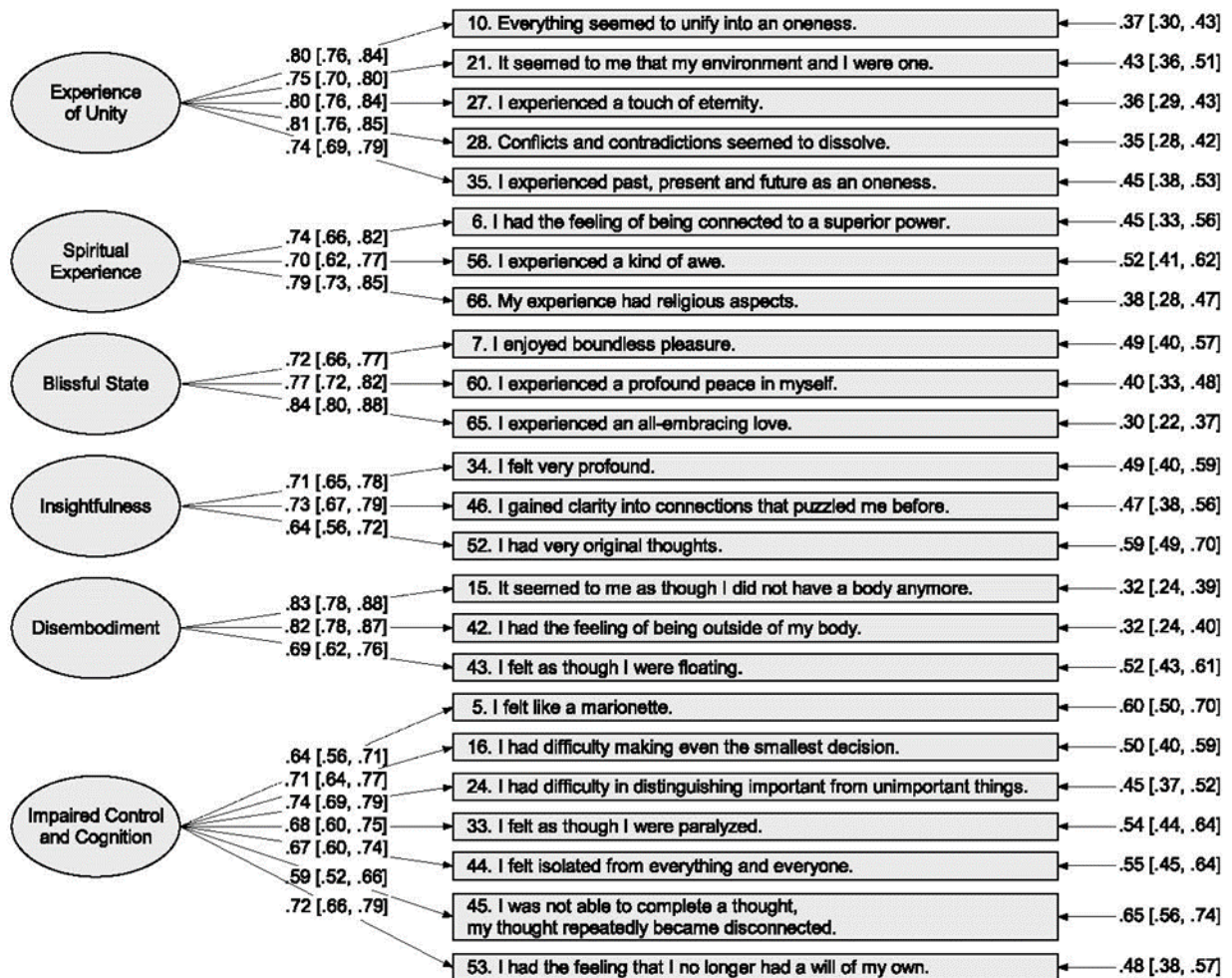
⁷² J'aimerais préciser ici qu'une traduction en français plus exacte de cette expression serait « faire face à ». J'ai choisi de dévier de celle-ci, au risque d'obscurcir la référence, pour des raisons esthétiques. Si le titre anglais « *Facing up to the doors of perception* » semble bien fonctionner, la traduction littérale « Faire face aux portes de la perception » me semblait par contre peu vendeuse.

Annexe

Structure factorielle du modèle *11D-ASC*

Tiré de : Studerus E, Gamma A, Vollenweider FX. *Psychometric Evaluation of the Altered States of Consciousness Rating Scale (OAV)*. PLoS ONE, vol. 5(8), Août 2010.

(Annexe 1 – Partie 1)



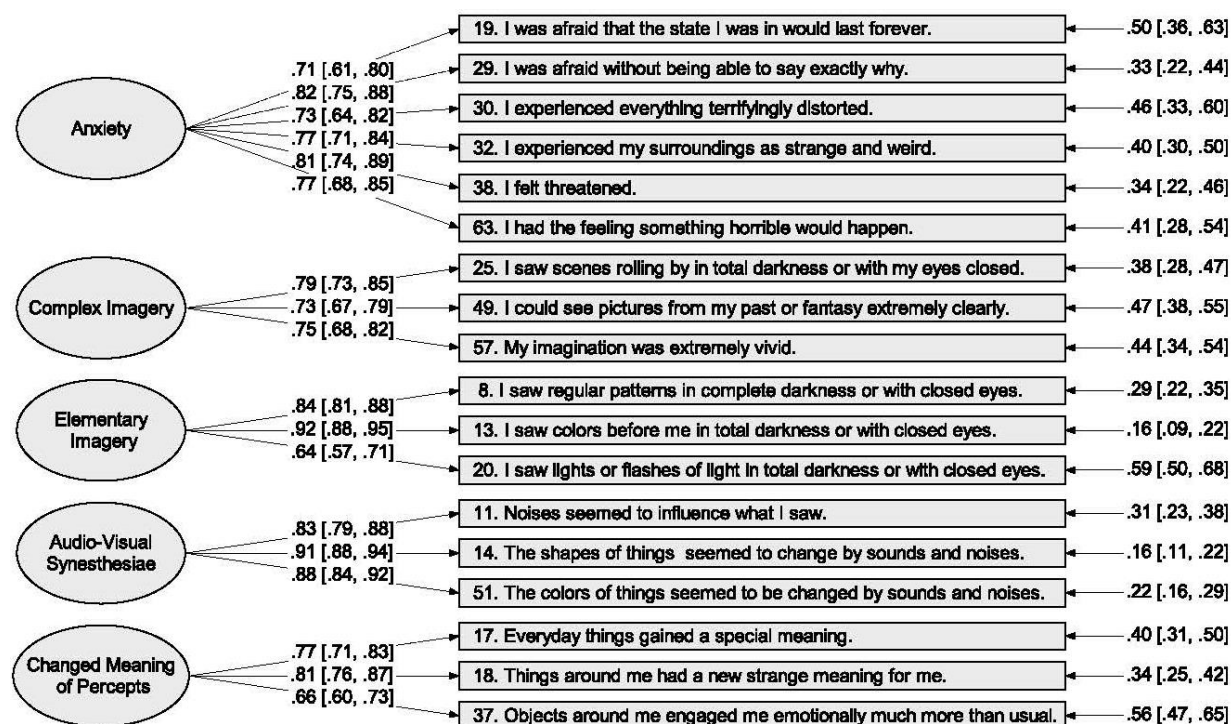
(Annexe 1 – Partie 2)

Figure 1. Final confirmatory factor analysis model with completely standardized loadings and error variances. Numbers in brackets are the 95% confidence intervals for the estimates. Covariances between factors were freely estimated and are shown in Table 3.
doi:10.1371/journal.pone.0012412.g001

Références

(Aydede & Murat, 2019)

AYDEDE, & MURAT. (2019). *Pain*. (E. N. Zalta, éd.) *The Stanford Encyclopedia of Philosophy*. Récupéré sur The Stanford Encyclopedia of Philosophy : <https://plato.stanford.edu/archives/spr2019/entries/pain/>.

(Barrett & Griffiths, 2017)

BARRETT, F. S., & GRIFFITHS, R. R. (2017). *Classic hallucinogens and mystical experiences : phenomenology and neural correlates*. (A. Halberstadt, F. Vollenweider, & D. Nichols, éd.) *Behavioral neurobiology of psychedelic drugs (Current topics in behavioral neurosciences, vol. 36)*. Springer.

(Bergé, 1990)

BERGÉ, C. (1990). *La voix des esprits : ethnologie du spiritisme*. Paris : Métailié.

(Block, 1995)

BLOCK, N. (1995). *On a confusion about a function of consciousness*. *Behavioral and brain sciences*, 18(2), pp. 227-287.

(Bressloff, 2002)

BRESSLOFF, P. C. (2002). *What geometric visual hallucinations tell us about the visual cortex*. *Neural Computation*, 14(3), pp. 473-491.

(Chalmers, 1995)

CHALMERS, D. (1995). *Facing up to the problem of consciousness*. (J. Goguen, éd.) *Journal of Consciousness Studies*, 2(3), pp. 200-219.

(Chalmers, 2003)

CHALMERS, D. (2003). *The character of consciousness*. Oxford : Oxford University Press.

(Chalmers, 2018)

CHALMERS, D. (2018). *The meta-problem of consciousness*. (V. G. Hardcastle, éd.) *Journal of Consciousness Studies*, 25(9-10), pp. 6-61.

(Chiu, 1989)

CHIU, L. P. (1989). *Review articles : diferencial diagnosis and management of hallucinations*. *Journal of the Hong Kong Medical Association*, 41(3), pp. 292-297.

(Crick & Koch, 1990)

CRICK, F., & KOCH, C. (1990). *Towards a neurobiological theory of consciousness*. *Seminars in the Neurosciences*, 2, pp. 263-275.

(Dennett, 1976)

DENNETT, D. C. (1976). *Are dreams experiences ?* *The Philosophical Review*, 85(2).

(Dennett, 1978)

DENNETT, D. C. (1978). *Why you can't make a computer that feels pain*. *Synthese*, 38(3), pp. 415-456.

150 • Références

(Dennett, 1982)

DENNETT, D. C. (1982). *Styles of mental representation. Proceedings of the Aristotelian Society*, 83, pp. 213-226.

(Dennett, 1990)

DENNETT, D. C. (1990). *La stratégie de l'interprète : le sens commun et l'univers quotidien*. (P. Engel, Trad.) Paris : Gallimard.

(Dennett, 1993)

DENNETT, D. C. (1993). *La conscience expliquée*. (P. Engel, Trad.) Paris : Odile Jacob.

(Dennett, 2003)

DENNETT, D. C. (2003). *Who's on first ? Heterophenomenology explained*. (V. G. Hardcastle, éd.) *Journal of Consciousness Studies*, 10(9-10), pp. 19-30.

(Dennett, 2004)

DENNETT, D. C. (2004). *Théorie évolutionniste de la liberté*. (C. Cler, Trad.) Paris : Odile Jacob.

(Dennett, 2012)

DENNETT, D. C. (2012). *De beaux rêves : obstacles philosophiques à une science de la conscience*. (C. Pichevin, Trad.) Paris : Gallimard.

(Dennett, 2016)

DENNETT, D. C. (2016). *Illusionism as the obvious default theory of consciousness*. (V. G. Hardcastle, éd.) *Journal of Consciousness Studies*, 23(11-12), pp. 65-72.

(Dennett, 2017)

DENNETT, D. C. (2017). *From bacteria to Bach and back*. Londres : Penguin Books.

(Dennett, 2019)

DENNETT, D. C. (2019). *Welcome to strong illusionism*. (V. G. Hardcastle, éd.) *Journal of Consciousness Studies*, 26(9-10), pp. 48-58.

(Dennett, 2020)

DENNETT, D. C. (2020). *A history of qualia*. (B. Ball, F. Nagle, I. Votsis, S. Stammers, & L. Bortolotti, éd.) *Topoi*, 39(1), pp. 5-12.

(Dennett & Akins, 2008)

DENNETT, D. C., & Akins, K. (2008). *Multiple drafts model*. 3(4). Récupéré sur http://www.scholarpedia.org/article/Multiple_drafts_model.

(DSM-V, 2013)

American Psychiatric Association, éd. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5^e édition). Arlington, Virginie : American Psychiatric Association.

(Frankish, 2012)

FRANKISH, K. (2012). *Quining diet qualia*. *Consciousness and Cognition*, 21(2), pp. 667-76.

(Frankish, 2016)

FRANKISH, K. (2016). *Illusionism as a theory of consciousness*. (V. G. Hardcastle, éd.) *Journal of Consciousness Studies*, 23(11-12), 11-39.

(Garcia-Romeu & Tart, 2013)

GARCIA-ROMEU, A. P., & TART, C. T. (2013). *Altered states of consciousness and transpersonal psychology*. (H. L. Friedman, & G. Hartelius, éd.) *The Wiley-Blackwell handbook of transpersonal psychology* (pp. 121-140). John Wiley & Sons, Ltd.

(Humphrey, 2011)

HUMPHREY, N. (2011). *Soul dust : the magic of consciousness*. Princeton : Princeton University Press.

(Johnson, Richards & Griffiths, 2008)

JOHNSON, M., RICHARDS, W., & GRIFFITHS, R. (2008). *Human hallucinogen research : guidelines for safety*. (B. A. Psychopharmacology, éd.) *Journal of Psychopharmacology*, 22(6), pp. 603-622.

(Kammerer, 2018a)

KAMMERER, F. (2018a). *Can you believe it ? Illusionism about consciousness and the illusion meta-problem*. *Philosophical Psychology*, 31(1), pp. 44-67.

(Kammerer, 2018b)

KAMMERER, F. (2018b). *Qualia (A)*. (M. Kristanek, éd.) Consulté en août 2022, sur L'encyclopédie philosophique : <https://encyclo-philos.fr/item/176>.

(Kammerer, 2019a)

KAMMERER, F. (2019a). *Conscience et matière : une solution matérialiste au problème de l'expérience consciente*. Paris : Éditions Matériologiques.

(Kammerer, 2019b)

KAMMERER, F. (2019b). *Does the explanatory gap rest on a fallacy ? Review of Philosophy and Psychology*, 10(4), pp. 649-667.

(Kammerer, 2021)

KAMMERER, F. (2021). *The illusion of conscious experience*. *Synthese*, 198(1), pp. 845-866.

(Kammerer, 2022)

KAMMERER, F. (2022). *How rich is the illusion of consciousness ? Erkenntnis*, 87(2), pp. 499-515.

(Leclercq, 2010)

LECLERCQ, B. (2010). *What is it like to be a bat ? Phénoménologie "à la troisième personne" de Wittgenstein à Dennett*. *Bulletin d'analyse phénoménologique*, 6(2), pp. 278-312.

(Letheby, 2021)

LETHEBY, C. (2021). *Philosophy of psychedelics*. Oxford : Oxford University Press.

(Letheby & Gerrans, 2017)

LETHEBY, C., & GERRANS, P. (2017). *Self unbound : ego dissolution in psychedelic*. *Neuroscience of consciousness*, 2017(1), pp. 1-11.

(Nutt, 2021)

NUTT, D. (2021). *Brain and mind made simple*. Gloucester : Waterside Press.

(Pylyshyn, 1985)

PYLYSHYN, Z. W. (1985). *Computation and cognition : towards a foundation for cognitive science* (2nde édition). Cambridge, Londres : Bradford Books.

(Ryle, 1978)

RYLE, G. (1978). *La notion d'esprit : pour une critique des concepts mentaux*. (S. Stern-Gillet, Trad.) Paris : Payot.

(Sacks, 2014)

SACKS, O. (2014). *L'odeur du si bémol : l'univers des hallucinations*. (C. Cler, Trad.) Paris : Éditions du Seuil.

(Stone, 2012)

STONE, J. A. (2012). *Spirituality for naturalists*. *Zygon : Journal of Science and Religion*, 47(3), pp. 481-500.

(Studerus, Gamma & Vollenweider, 2010)

STUDERUS, E., GAMMA, A., & VOLLENWEIDER, F. X. (2010). *Psychometric evaluation of the altered states of consciousness rating scale (OAV)*. (V. Bell, éd.) *PLoS ONE*, 5(8).

(Van Gulick, 2021)

VAN GULICK, R. (2021). *Consciousness*. (E. N. Zalta, éd.) Consulté en août 2022, sur The Stanford Encyclopedia of Philosophy : <https://plato.stanford.edu/entries/consciousness/>.