Document generated on 04/23/2024 1:49 p.m.

Philosophiques

Relier la conscience phénoménale et éliminer le fossé explicatif

Maria Clara Dias

Volume 35, Number 2, Fall 2008

URI: https://id.erudit.org/iderudit/000440ar
DOI: https://doi.org/10.7202/000440ar

See table of contents

Publisher(s)

Société de philosophie du Québec

ISSN

0316-2923 (print)
1492-1391 (digital)

Explore this journal

Cite this article

Dias, M. C. (2008). Relier la conscience phénoménale et éliminer le fossé explicatif. *Philosophiques*, *35*(2), 513–524. https://doi.org/10.7202/000440ar

Article abstract

The main argument of this paper is that there is no gap between physicalist explanations and phenomenal consciousness, as a matter of principle. Based on an analysis of certain *sui generis* properties of phenomenal consciousness, I aim to demonstrate that the apparent explanatory gap between the mental and the physical can be eliminated. More specifically, I want to defend the claim that the so-called *sui generis* qualities of mental events, referred to as *qualia,* are physical properties. These properties would constitute, as Tye puts it, the nonconceptual representational content of our conscious experiences. Moreover, they are the input of physical processes whose output would be intentional states with propositional content.

Tous droits réservés © Société de philosophie du Québec, 2009

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

Relier la conscience phénoménale et éliminer le fossé explicatif *[1]

MARIA CLARA DIAS
Université fédérale de Rio de Janeiro
mcdias@ifcs.ufrj.br

RÉSUMÉ. — L'argument principal de cet article est qu'il n'y a pas d'incompatibilité entre les explications physicalistes et l'idée de conscience phénoménale. Ayant comme référence une analyse de certaines propriétés *sui generis* de la conscience phénoménale, nous avons l'intention de souligner que l'apparente incompatibilité entre le mental et le physique peut être éliminée. De façon plus précise, nous soutenons la thèse selon laquelle les propriétés *sui generis* des événements mentaux, connues comme *qualia*, sont des propriétés physiques. Ces propriétés constitueraient, comme l'affirme Tye, les contenus représentationnels nonconceptuels de nos expériences. En outre, ces propriétés sont des *inputs* des processus physiques dont les *outputs* seraient des états intentionnels avec des contenus propositionnels.

ABSTRACT. — The main argument of this paper is that there is no gap between physicalist explanations and phenomenal consciousness, as a matter of principle. Based on an analysis of certain *sui generis* properties of phenomenal consciousness, I aim to demonstrate that the apparent explanatory gap between the mental and the physical can be eliminated. More specifically, I want to defend the claim that the so-called *sui generis* qualities of mental events, referred to as *qualia,* are physical properties. These properties would constitute, as Tye puts it, the nonconceptual representational content of our conscious experiences. Moreover, they are the input of physical processes whose output would be intentional states with propositional content.

* J'aimerais exprimer mes remerciements à Carlos Eduardo Nogueira Loddo (UQTR) pour ses commentaires extrêmement utiles sur des versions antérieures de cet article. Ses observations stimulantes m'ont aidée à présenter mes thèses de façon plus claire, surtout en ce qui regarde les questions conceptuelles. Il m'a aussi aidée à formuler une interprétation précise (dans ce qu'il y a de pertinent pour la présente discussion) de la perspective proposée par Kaufmann avec qui il a travaillé étroitement, traduisant beaucoup de ses articles en portugais (incluant celui qui est commenté ici). Pour la révision anglaise du présent texte, je tiens à remercier Mário Nogueira et André Fuhrmann.

1. NdT. Le titre original anglais du présent article est « Tying Phenomenal Consciousness and Quining the Gap ». L'auteur s'en explique comme suit : Bien entendu, mon utilisation du néologisme « to quine » est tirée de D. Dennett, « Quining Qualia », in A. Marcel, and E. Bisiach, (eds.), 1988, *Consciousness in Contemporary Science*, Oxford, Oxford University Press. J'invente aussi un nouveau terme, « *to tye* », afin de tirer avantage de la coïncidence phonétique entre le nom de Michael Tye et le verbe anglais « *to tie* » (relier).

Dans son article portant sur le problème de la conscience et sur le fossé explicatif[2], J. Nicolas Kaufmann présente d'intéressantes contributions à la discussion des caractéristiques qualitatives des expériences subjectives (ou à la première personne) des phénomènes mentaux. Essentiellement, le problème du fossé explicatif émerge du fait que, d'un côté, les caractéristiques qualitatives semblent jouer un rôle causal dans les actions humaines alors que, de l'autre, aucune référence conceptuellement satisfaisante à ces éléments ne semble convenir à l'intérieur des structures explicatives. En effet, les demandes d'objectivité associées à ces structures paraissent ne pouvoir être rencontrées par les expériences qualitatives en raison de leur caractère subjectif. Par conséquent, il y a un fossé entre les qualités de l'expérience subjective et l'explication psychologique de l'action humaine acceptable dans un cadre physicaliste. Kaufmann admet la formulation de ce problème telle que proposée par Josef Levine[3].

S'il y avait un tel fossé, il poserait un sérieux problème au projet de naturalisation de l'explication psychologique. Le problème serait sérieux puisqu'il ne tiendrait pas simplement à une question de fait, relative à l'état présent de notre science (laquelle peut être amenée à se modifier dans le futur), mais à une *question de principe*. Le problème du fossé explicatif étant ainsi formulé comme un problème de principe, Kaufmann considère qu'il doit aussi être envisagé comme une question *conceptuelle*.

En raison de ce caractère conceptuel, Kaufmann commence son analyse en essayant de circonscrire précisément le *concept* de conscience. Il présente d'abord une série de distinctions et d'oppositions conceptuelles concernant la conscience et établit quel concept de conscience est approprié pour cette discussion. L'opposition de Ned Block entre *la conscience d'accès* et la *conscience phénoménale*[4] lui importe tout particulièrement. Selon cette opposition, la *conscience d'accès* est représentationnelle et définissable fonctionnellement. La *conscience phénoménale*, de son côté, ne peut être définie comme un concept général, bien qu'il soit possible d'indiquer quels objets peuvent être rangés sous cette notion. Kaufmann constate que la liste de ces objets varie d'un auteur à l'autre. Quoi qu'il en soit, c'est cette dernière notion de conscience qui est impliquée dans l'expérience subjective ou vécue à la première personne. Kaufmann localise donc ici le concept de conscience autour duquel le problème du fossé explicatif trouve sa formulation et en rapport auquel il doit être traité.

---

2. Kaufmann, J. N., 1999, « A Problemática da Consciência e a Lacuna Explicativa », *Princípios. Revista de Filosofia*, vol. 6, n. 7, 87-120.

3. Levine, J., 1983, « Materialism and Qualia : the Explanatory Gap », *Pacific Philosophical Quarterly*, 64, 354-361.

4. Block, N., 1995, « On a Confusion About a Function of Consciousness », *Behavioral and Brain Sciences*, 18, 227-247.

Par la suite, Kaufmann entreprend d'examiner deux perspectives opposées à propos du problème du fossé explicatif. D'abord, il étudie la position de Colin McGinn qui inclut des arguments favorables à l'idée qu'il y a un fossé[5]. Kaufmann s'arrête à deux des arguments de McGinn en employant les reconstructions de ces arguments offertes par Owen Flanagan[6]. Kaufmann retient les critiques de Flanagan et ajoute d'autres critiques provenant de la phénoménologie de Edmund Husserl.

Mais c'est l'examen par Kaufmann de la position de David Chalmers qui est le plus pertinent pour ce dont nous allons discuter dans ce texte. Chalmers favorise la « fermeture du fossé », ce qui revient à dire qu'il essaye de fournir une solution au problème. Comme le remarque Kaufmann, Chalmers présente dans différents textes des stratégies diverses pour parvenir à cette solution. Dans les premiers textes[7], Chalmers tente de compléter les explications physicalistes par une analyse conceptuelle typique de la phénoménologie. Dans un texte plus récent[8], afin de corréler la conscience à des processus neuronaux (ce qu'il cherche à établir), Chalmers propose ce qu'il nomme des « principes de rapprochement » (*bridging principles*). Des états fonctionnellement définissables sont reliés à des structures neuronales parce que les premiers peuvent être réalisés dans les secondes (cela constitue la prémisse empirique de l'argument de Chalmers). Nous pouvons avoir accès à la conscience grâce à des rapports verbaux ou par d'autres modes comportementaux d'expression (tel est le rôle dévolu au principe de rapprochement). Kaufmann analyse l'argument de Chalmers, soulève une série de critiques et, pour certaines d'entre elles, fait une fois de plus appel à la phénoménologie de Husserl.

Je n'ai pas pour but ici d'examiner le détail de l'argument de Chalmers ni de fournir un compte rendu complet de la critique qu'en fait Kaufmann. Je retiendrai toutefois une observation de Kaufmann. Il note que les principes de rapprochement de Chalmers ne peuvent réussir à relier la conscience à l'explication physicaliste que *si* la conscience est conceptualisée comme étant ce que Block nomme la conscience *d'accès* (le concept représentationnel et fonctionnellement définissable) ; ils en sont néanmoins *tout à fait* incapables si elle est conceptualisée en tant que conscience *phénoménale*.

---

5. McGinn, C., 1983, *The Problem of Consciousness*, Oxford, Blackwell.

6. Flanagan, O., 1992, *Consciousness Reconsidered*, Cambridge, Mass., MIT Press.

7. Chalmers, D. J., 1996, *The Conscious Mind : In Search of a Fundamental Theory*, New York, Oxford University Press ; Chalmers, D. J., « Facing up to the Problem of Consciousness » — « Response », in Shear, J. (ed.), 1997, *Explaining Consciousness — the Hard Problem*, Cambridge, Mass., MIT Press, 9-33.

8. Chalmers, D. J., « On the Search for the Neural Correlate of Consciousness », in Hameroff, S. R., Kaszniak, A. W. and Scott, A. (eds.), 1998, *Toward a Science of Consciousness II : The Second Tucson Discussions and Debates. Complex Adaptative Systems*, Cambridge, Mass., MIT Press.

Comme nous l'avons vu, c'est sur ce dernier concept que repose le problème du fossé explicatif selon l'approche de Kaufmann. Cependant, bien que cette observation soit critique à l'endroit de la tentative de Chalmers de fermer le fossé, et pour autant que le projet de le fermer soit considéré, l'attitude de Kaufmann est plus optimiste que son évaluation de la position de Chalmers ne le suggère. Selon lui, *si* Chalmers réussissait à montrer (contrairement à ce que Block suppose) qu'un concept de conscience fonctionnellement défini est systématiquement corrélé au concept de conscience phénoménale, *alors* son argument pourrait tenir le coup et le fossé serait fermé.

Dans la suite de ce texte, je vais examiner une perspective de rechange qui semble en mesure d'assurer cette fermeture. Je soutiendrai qu'il n'existe aucun fossé entre les explications physicalistes et la conscience phénoménale, aucun problème de principe. Cette perspective est celle qui a été suggérée par Michael Tye[9]. En un sens, ce que Tye propose correspond exactement à la suggestion faite par Kaufmann pour sauver l'argument de Chalmers. En effet, Tye met systématiquement en corrélation un concept de conscience phénoménale avec un concept de conscience fonctionnellement définissable. Plus radicalement, selon Tye, le concept de conscience phénoménale est fonctionnellement définissable.

En me basant sur une analyse de certaines propriétés *sui generis* de la conscience phénoménale, j'espère montrer que le fossé explicatif qui semble apparaître entre le mental et le physique peut être éliminé. Plus spécifiquement, je veux défendre l'affirmation selon laquelle les prétendues qualités *sui generis* des événements mentaux, celles auxquelles nous référons en parlant de *qualia*, sont des propriétés physiques. Ces propriétés constitueraient, comme Tye le propose, le contenu représentationnel non conceptuel de nos expériences conscientes. De plus, elles sont les inputs de processus physiques dont les outputs sont des états intentionnels dotés d'un contenu propositionnel.

En général, nous pouvons comprendre le physicalisme comme étant la perspective selon laquelle tout ce qui existe ou peut être connu appartient au monde physique, c'est-à-dire au domaine des objets et des lois d'une microphysique élémentaire, qui sera éventuellement développée dans l'avenir. Quelle que soit la définition du « monde physique » que nous pourrions adopter, il semblerait que le monde physique doit incorporer un ensemble de propriétés applicables à un groupe spécifique d'objets, lesquelles propriétés ne peuvent, du moins à première vue, en aucune façon être attribuées à des objets physiques de manière sensée. Par exemple, nous disons que « personne d'autre que moi ne peut sentir ma douleur » et semblablement que la douleur, la tristesse, la joie et autres états mentaux n'existent que lorsqu'ils sont vécus par un sujet ou pour autant qu'ils lui appartiennent. Mon expérience consciente, les sensations de mon corps, mes émotions, mes passions et mes humeurs semblent être liées à moi par une relation que je ne peux

9. Tye, M., 1995, *Ten Problems of Consciousness*, Cambridge, Mass., MIT Press.

entretenir avec mon automobile et mes vêtements, par exemple. Une autre personne pourrait posséder ma voiture et mes vêtements, et ces objets ne cesseraient pas d'exister s'ils n'étaient pas possédés par quelqu'un.

La « monopolisation » des phénomènes mentaux vécus par un sujet, par un soi, semble caractériser une sorte d'appartenance nécessairement intérieure qui est étrangère aux objets qui m'appartiennent seulement de façon contingente. Ne devons-nous pas alors distinguer les objets auxquels j'ai un accès privilégié, c'est-à-dire les objets qui peuvent seulement être « possédés » et « connus » dans la perspective de mon expérience et qui ne peuvent exister que s'ils sont en relation avec le sujet de l'expérience, des autres objets avec lesquels j'entretiens une relation simplement parce que je suis, de manière contingente, celle qui les possède ? Si les propriétés caractéristiques des objets du premier groupe — c'est-à-dire la propriété d'être privé, la propriété d'être nécessairement possédé et la propriété d'être dans la perspective de la première personne[10] — ne peuvent être attribuées aux objets physiques, ne devons-nous pas admettre l'existence d'entités non physiques ?

Nous nous heurterions à des difficultés si nous adoptions une telle supposition. Car elle semble violer un principe largement accepté, à savoir que le monde physique est causalement fermé. Ce principe dit que, pour n'importe quel événement physique, quelles qu'en soient les causes, elles sont d'autres événements physiques. Les entités non physiques n'ont donc aucun rôle à jouer dans les processus de réalisation des événements physiques. Bien entendu, les événements phénoménaux peuvent toujours être les effets d'événements physiques, mais jamais leurs causes (ce qui est la solution épiphénoméniste au problème des *qualia*). Mais sommes-nous vraiment disposés à admettre que les états phénoménaux sont toujours causalement non pertinents ?

Quoi qu'il en soit, nombreux sont ceux qui partagent l'intuition qu'il y a un gouffre infranchissable entre l'explication physicaliste de notre vie mentale et ce que nous croyons intérieurement qu'il se passe dans certains contextes. Est-il possible, en termes physicalistes, c'est-à-dire dans les termes d'une microphysique élémentaire, d'expliquer, voire seulement en principe, ce que signifie être le sujet de certaines expériences, ou d'exprimer d'une façon ou d'une autre le point de vue privilégié que nous présumons avoir relativement à nos sensations, émotions, passions et humeurs ?

Confrontés à ce problème, nous avons le choix entre trois solutions :

1. Accepter le caractère *sui generis* des objets phénoménaux et leur irréductibilité au domaine du monde physique ;
2. Montrer que les objets phénoménaux ne possèdent pas réellement ces propriétés *sui generis* ; ou montrer que ces objets n'existent pas et ainsi qu'il n'y a pas lieu d'appliquer ces propriétés ; ou

10. Sur ces propriétés, voir Tye, 1995.

3. Reconnaître l'existence des objets phénoménaux et de leurs propriétés caractéristiques, en refusant toutefois que celles-ci ne s'appliquent pas aux objets physiques.

La première solution est clairement dualiste. Elle rétablit un dualisme ontologique à l'intérieur duquel la relation entre le mental et le physique devient contingente, arbitraire ou même impossible. Au mieux, il serait possible de réintroduire une thèse épiphénoméniste selon laquelle le mental est causalement inerte. Cependant, nos croyances matérialistes du XXIe siècle rendent difficile de comprendre en quoi consiste penser à des objets qui ne produisent aucun effet sur le monde physique. Comment ces objets pourraient-ils même être perçus s'ils ne laissent aucune trace sur le monde physique ? Affirmer l'existence d'une chose semble impliquer qu'elle est en relation causale avec d'autres objets ou événements physiques.

La seconde solution est ou bien réductionniste ou bien éliminativiste. Respectivement, soit elle réduit les objets phénoménaux à des objets physiques, niant ainsi leurs caractéristiques particulières ; soit elle nie l'existence des objets phénoménaux et par conséquent la réalisation des propriétés *sui generis*. Cette solution laisse intact le physicalisme, mais pas nos pratiques courantes d'attribution d'un pouvoir causal aux états mentaux.

La troisième solution semble combiner le physicalisme avec une certaine forme de dualisme (ou de pluralisme) de propriétés, permettant ainsi une relative compréhension du caractère particulier de quelques-unes de nos expériences. La seconde et la troisième solution partagent l'essentiel de l'intuition physicaliste. La différence entre les deux est plus subtile qu'il ne paraît à première vue. Une certaine forme de réductionnisme ou d'éliminativisme semble inévitable même à la troisième solution. La différence entre elles se trouve dans le fait que la troisième solution est la seule qui se veut une tentative de conciliation du physicalisme avec certaines de nos intuitions les plus importantes concernant nos états subjectifs. Mais quels états subjectifs précisément ? Ou, autrement dit, lesquelles de leurs caractéristiques propres peuvent, de façon plausible, être combinées avec l'essentiel de nos croyances physicalistes ?

Nous devons ici clarifier notre position en rapport à l'*explanandum* du problème. Ce qui semble résister à une réduction physicaliste n'est pas les états mentaux en général mais certaines de leurs propriétés : caractère privé, appartenance nécessaire et subjectivité du point de vue. Ces propriétés expriment le caractère subjectif de notre expérience ; elles sont reliées à ce qui est vécu uniquement et exclusivement dans la perspective du sujet de l'expérience. Pour employer la formule de Thomas Nagel, ces propriétés constituent « l'effet que cela fait » [11] d'avoir notre expérience. Selon la conception de Nagel, ces propriétés sont présentes dans chaque événement intentionnel, c'est-à-dire dans

11. Nagel, T., 1979, « What Is It Like To Be a Bat ? », in *Mortal Questions,* Cambridge, Cambridge University Press.

chaque type de relation impliquant la conscience et les objets de conscience. Elles constituent le caractère irréductiblement subjectif des événements intentionnels.

Je vais ici m'en tenir à une position plus limitée, à une position qui exclut de ce domaine les états mentaux comme les croyances, les désirs et les intentions dont le contenu s'exprime par une clause-que. Afin de justifier une telle restriction, j'entends montrer qu'il est possible de fournir, au moins de prime abord, une description fonctionnelle de nos attitudes propositionnelles, et qu'une telle description semble n'omettre aucun élément pertinent de notre compréhension de ces états mentaux. Pour sa part, le caractère irréductiblement subjectif de l'expérience, ou le *quale*, sera considéré comme une qualité particulière de certains états mentaux. En raison de ces qualités particulières, la perspective du sujet de l'expérience semble à ceux qui défendent l'idée d'un fossé explicatif ajouter quelque chose d'irréductible à l'expérience elle-même. Je voudrais donc caractériser le problème du fossé explicatif comme portant spécifiquement sur les états mentaux dont la description apparaît insuffisante sans une référence directe à l'agent. Avant de présenter le problème lui-même, je donnerai une brève exposition de ce que pourrait être une explication fonctionnelle de nos attitudes propositionnelles, justifiant ainsi la restriction mentionnée ci-dessus.

La perspective fonctionnaliste identifie en général les états mentaux ou bien avec la fonction accomplie par des états cérébraux ou bien avec les états cérébraux accomplissant cette fonction. Ce qui fonde l'attribution d'un certain contenu *p* à un état de croyance, par exemple, est la relation que cet état entretient avec certaines stimulations sensorielles, avec d'autres états mentaux et avec le comportement. Les états mentaux sont donc à la fois des causes internes du comportement et des effets produits par le monde extérieur. Pour employer les mots de David Braddon-Mitchell et Frank Jackson :

> La théorie fonctionnaliste de l'esprit spécifie les états mentaux par trois types de clauses : les clauses relatives aux inputs, qui stipulent quel type de conditions donnent naissance à quel type d'états mentaux ; les clauses relatives aux outputs, qui stipulent quel type d'états mentaux donnent naissance à quel type de réponses comportementales ; et les clauses relatives aux interactions, qui stipulent comment les états mentaux interagissent typiquement les uns avec les autres.[12]

Afin d'éviter toute circularité, la caractérisation fonctionnaliste des états intentionnels doit se faire exclusivement en employant un vocabulaire physique non intentionnel. En particulier, les états intentionnels que nous avons l'intention de décrire en termes physiques ne doivent pas être subrepticement

12. Braddon-Mitchell, D., and Jackson, F., 1996, *The Philosophy of Mind and Cognition*, Oxford, Blackwell, p. 40.

réintroduits dans la caractérisation de l'output. À l'instar des inputs, les outputs doivent aussi être entièrement décrits en termes physiques. La caractérisation fonctionnelle des états mentaux spécifie ainsi un réseau de relations physiques qui implique des dispositions comportementales, des causes typiques, de même que d'autres états mentaux, pour autant que ces derniers sont aussi caractérisés en termes fonctionnels.

Les applications les plus évidentes de la théorie suivent le modèle d'une machine programmée pour exécuter différentes tâches spécifiques. Cela inclut des machines très simples comme les distributeurs automatiques qui donnent un cola après qu'on y a introduit des pièces de monnaie d'une certaine valeur, mais aussi des machines très complexes comme les machines connexionnistes de type PDP, ou les réseaux de neurones. Évidemment, la complexité des fonctions réalisées par les états mentaux exige des modèles flexibles et très complexes. Dans plusieurs cas, la complexité est telle que nous sommes encore incapables de spécifier les modèles en question. Cela, néanmoins, ne va pas à l'encontre de l'affirmation selon laquelle rien ne peut en principe nous empêcher de décrire les aspects pertinents des événements mentaux en termes physiques.

L'esprit humain décrit fonctionnellement opère ainsi comme une machine à calculer composée de plusieurs modules. Il y aurait un scanneur dans le premier module, celui qui est responsable de la réception des inputs. À partir de ce point, nous pouvons imaginer plusieurs modules. Parmi eux, il y aurait un module d'évaluation responsable de la sélection des données d'information produites par un générateur de solutions. Finalement, un module moteur fournirait une réponse comportementale spécifique en guise d'output. La particularité d'un tel assemblage de modules est son aptitude à l'apprentissage, c'est-à-dire sa capacité à altérer le résultat final en fonction des nouveaux inputs qui reflètent les effets produits par le comportement de l'agent dans les étapes précédentes. En d'autres mots, l'output produit favoriserait des réponses extérieures qui, à leur tour, seraient introduites en tant qu'inputs et évaluées par le programme afin qu'il puisse fournir un nouveau résultat. Par exemple, un programme capable de reconnaître les caractères de l'alphabet, comme celui qui est utilisé avec succès par les services postaux, transforme et améliore sa performance lorsque différents exemplaires du même caractère sont donnés et que le gestionnaire répond affirmativement ou négativement à leur identification. Ces programmes sont capables d'organiser l'information de telle sorte qu'un caractère précédemment mal interprété puisse plus tard être reconnu. En nous basant sur une analyse fonctionnelle de modèles au programme flexible, nous pouvons alors rendre compte de l'une des caractéristiques principales des êtres humains : leur capacité à apprendre. De la même façon, d'autres caractéristiques typiquement humaines, comme notre aptitude à délibérer et à infléchir le cours de nos actions — desquelles dépendent nos

attributions de responsabilité et de liberté à un agent — peuvent aussi être traitées avec succès.[13]

En supposant qu'une caractérisation fonctionnelle de nos attitudes propositionnelles est possible, et que rien de ce qui leur est essentiel ne soit perdu avec une explication fonctionnelle, le problème du fossé explicatif finit donc par se limiter à la question des *qualia*. Comment s'intègrent-ils à une description physique du monde ? À quoi réfèrent-ils exactement ?

Si nous ne sommes pas prêts à accepter une perspective réductionniste ou éliminativiste relativement aux *qualia*, nous devons être capables d'en fournir une description physique. Autrement dit, nous devons expliquer en termes physicalistes comment les propriétés liées à « l'effet que cela fait » peuvent jouer un rôle causal authentique dans nos dispositions comportementales. Afin de développer cette hypothèse, je fais appel à la théorie PANIC du caractère phénoménal élaborée par Tye, théorie selon laquelle « le caractère phénoménal est identique au Contenu Intentionnel non Conceptuel Abstrait Prêt pour l'usage »[14]. Bien que la théorie de Tye fasse la distinction entre l'aspect non conceptuel et l'aspect abstrait de l'expérience phénoménale, je vais exclusivement considérer le côté non conceptuel, celui qui est le plus directement relié à la présente discussion.

La théorie PANIC est une théorie qui traite les événements ou états phénoménaux comme des phénomènes intentionnels, bien que non conceptuels. Ces états peuvent s'unir à d'autres états intentionnels dotés d'un contenu propositionnel. Ils peuvent servir d'éléments de base aux processus cognitifs qui impliquent des états intentionnels pourvus d'un contenu propositionnel. Le point crucial ici est que les états intentionnels phénoménaux, par opposition aux attitudes propositionnelles, ne sont pas eux-mêmes conceptuels. Les sensations physiques, comme un mal de dents ou un orgasme, sont représentationnelles, de sorte que, dans des conditions normales, elles reflètent causalement les modifications de parties spécifiques du corps. Ces représentations (1) sont les outputs de récepteurs nerveux répondant à plusieurs stimuli dans notre corps et (2) s'intègrent aux inputs des processus cognitifs, qui sont pour la première fois révélés cognitivement au sujet de l'expérience. Tye affirme de plus que, par l'action des processus cognitifs opérant sur le plan des représentations non conceptuelles, le sujet devient conscient de ses états phénoménaux. Prendre conscience de ses propres expériences est un processus cognitif grâce auquel, pour la première fois, l'expérience phénoménale tombe sous des concepts. C'est seulement à ce moment que le sujet de l'expérience adopte une attitude épistémique (cognitive) par rapport à ses états phénoménaux. Sous ce niveau de conscience — traditionnellement nommé « introspection » — il y a la conscience phénoménale. Le contenu

13. Pour une défense du compatibilisme basé sur l'exposition du modèle de l'autocontrôle, voir Blackburn, S., 1999, *Think*, Oxford, Oxford University Press, p. 98 et suiv.

14. Tye, 1995, p. 137. (PANIC pour « Poised Abstract Nonconceptual Intentional Content » — NdT).

phénoménal serait donc mieux décrit comme étant une représentation non conceptuelle de changements physiques, prête à intégrer l'input des processus qui ont pour output des états intentionnels pourvus d'un contenu propositionnel (désirs et croyances). De la sorte, être le sujet d'une expérience phénoménale signifie posséder certains contenus intentionnels non conceptuels qui peuvent être les inputs de certains processus cognitifs, autrement dit des représentations non conceptuelles pouvant offrir différentes réponses.

L'explication fournie par la théorie PANIC permet de rendre le physicalisme compatible avec l'intuition selon laquelle l'expérience de voir du rouge et l'expérience de *savoir l'effet que cela fait de voir du rouge* sont des états différents irréductibles l'un à l'autre. Quelque chose d'assurément nouveau (une nouvelle expérience) arrive à quelqu'un qui voit du rouge pour la première fois en dépit du fait que cette même personne sache déjà tout ce qu'il est possible de savoir, en termes physicalistes, concernant *l'effet que cela fait de voir du rouge*. Ici, nous sommes confrontés à l'un des principaux arguments contre le physicalisme, l'argument de la connaissance. J'aimerais consacrer la suite de ce texte à montrer comment la description de la conscience phénoménale proposée par la théorie PANIC permet de rendre compte de la nouvelle expérience de Marie sans mettre en péril le physicalisme.

Marie est la fameuse scientifique imaginée par Jackson, celle qui sait tout sur le monde physique. Elle possède une connaissance complète des éléments et des lois de la physique du futur, son savoir comprend tout ce qu'il est possible de connaître dans le domaine de la physique. Après avoir présenté Marie, jetons un coup d'œil à son histoire. La brillante Marie a grandi et a tout appris dans une pièce noire et blanche. Tous ses contacts avec le monde extérieur se sont faits par le biais d'un moniteur noir et blanc. Pour les besoins de l'argument, nous pouvons même imaginer qu'un dispositif complexe la fait se voir elle-même ainsi que tout ce qui l'entoure dans sa chambre en noir et blanc. Un jour, Marie quitte sa chambre et voit pour la première fois une tomate rouge. Elle fait l'expérience de la couleur rouge pour la première fois. Marie connaissait non seulement tout des tomates, mais aussi tout de l'expérience visuelle de la couleur. Elle savait donc que les tomates mûres, les bouches d'incendie et les cabines téléphoniques anglaises sont rouges. Elle savait tout aussi des processus physiologiques qui sont à la base de l'expérience visuelle de la couleur. C'est alors que le physicalisme parvient à un point critique : supposons que nous partagions l'intuition que Marie a appris quelque chose de nouveau lorsqu'elle a vu les couleurs pour la première fois, ne devons-nous pas admettre que quelque chose d'inhérent à notre expérience ne peut pas être décrit en termes physicalistes ? L'argument de Marie semble indiquer exactement la distinction entre *savoir quelque chose* et *faire l'expérience de quelque chose*. Marie sait tout des couleurs, mais elle n'a jamais fait l'expérience de voir les couleurs avant de quitter sa chambre. De sorte que, bien qu'elle sache *ce qu'est voir des couleurs*, elle n'a jamais su *l'effet que cela*

fait d'en voir. Ce qui semble irréductible au physicalisme est donc ce caractère privilégié de la perspective de la personne qui est le sujet de l'expérience : le caractère subjectif de l'expérience.

Pour les besoins de l'argument, bien qu'ils puissent sembler très invraisemblables, j'accepterai les éléments de l'histoire. Il n'est pas dans mon intention de nier l'intuition de base selon laquelle Marie a fait l'expérience de quelque chose de nouveau lorsqu'elle a quitté la chambre. Ce que je veux nier en revanche est l'idée que ce quelque chose de nouveau ne puisse être décrit en termes physicalistes ou qu'il marque les frontières du physicalisme. Si Marie savait déjà tout des tomates rouges, évidemment, ce qui lui est arrivé au moment où elle les a vues pour la première fois ne peut avoir altéré son savoir. Nous pouvons cependant dire que quelque chose de nouveau a été introduit dans la formation de ce savoir. Selon la théorie de Tye, de nouvelles représentations non conceptuelles viennent s'intégrer à l'input des processus cognitifs de Marie. Bien que ces représentations ne changent pas le contenu cognitif de Marie, elles peuvent favoriser des changements dans son comportement ultérieur. Elles peuvent, par exemple, lui permettre d'acquérir la compétence de choisir plus rapidement les tomates qu'elle veut mettre dans une salade. Les éléments introduits sont des représentations d'éléments qui entrent dans la catégorie des faits que Marie connaissait déjà, étant donné qu'elle savait tout du monde physique. Mais l'introduction de ces éléments nouveaux permet une réorganisation des données disponibles qui détermine le comportement futur de Marie. Selon une telle interprétation, ce que Marie acquiert n'est pas un contenu cognitif nouveau, mais une nouvelle forme d'accès au contenu cognitif qu'elle possède déjà. Ainsi, il n'y aurait pas de contenu cognitif qui ne puisse être traité par une théorie physicaliste du monde, pas plus qu'il n'y aurait un accès non physique à ce contenu. L'hypothèse physicaliste resterait intacte.

Si la description fonctionnelle de nos états mentaux *en général* et la description que nous venons de donner du caractère prétendument subjectif de notre expérience *en particulier* sont toutes deux plausibles, serions-nous alors en position d'éliminer non seulement le fossé explicatif, mais aussi le vocabulaire mental et, avec lui, certaines pratiques explicatives largement acceptées et couronnées de succès ?

C'est là le point où il me semble raisonnable d'admettre que nous faisons face à un dilemme, et peut-être est-ce aussi celui qui sépare une simple perspective éliminativiste d'une explication fonctionnelle constructive des états mentaux. Nous ne voulons pas supprimer les caractéristiques *sui generis* de notre expérience, mais nous voulons les expliquer à partir d'un point de vue physique. Si nous obtenons du succès dans notre tentative d'offrir une explication plausible — laquelle pourrait faire ses preuves dans l'avenir — cela peut ne laisser aucune place à d'autres sortes d'explications. En effet, nos pratiques explicatives semblent entrer dans un processus graduel de substitution. Les événements et les états que nous sommes maintenant prêts à considérer

comme étant les causes réelles du comportement désorganisé d'un patient épileptique excluent les explications que nous donnions autrefois de ce phénomène. Nous plaçons l'épilepsie elle-même dans une histoire causale fonctionnellement conceptualisée et associée par des liens de réalisation à des processus neurologiques. La possession par le diable, à titre d'exemple, est aujourd'hui une hypothèse supprimée. Ce n'est pas une hypothèse acceptable dans la caractérisation des causes *véritables* du comportement épileptique.

Certains philosophes semblent voir des limites précises à cette substitution progressive des schèmes explicatifs.[15] Selon eux, nos pratiques explicatives largement acceptées et couronnées de succès ne peuvent être améliorées par aucun développement physicaliste. Jusqu'à maintenant, le présent article ne semble pas menacer ces pratiques. Mais nous ne pouvons pas bloquer ici l'émergence potentielle de ces substitutions. Pouvons-nous accepter comme une possibilité réelle que nous remplacions la façon bien enracinée que nous avons de décrire et d'expliquer grâce au vocabulaire mental les aspects de notre vie consciente par une théorie qui n'existe que dans l'imagination futuriste des philosophes d'orientation physicaliste ? Quelle sorte de schème explicatif fera place à notre conception quotidienne du monde des expériences ? Notre réponse est : une théorie de prime abord physicaliste, c'est-à-dire un pari pour la future compréhension fonctionnelle des états mentaux qui permettra de combler les fossés explicatifs actuels et qui pourra rendre notre description du fonctionnement humain compatible avec nos croyances matérialistes[16].

15. Je pense ici à des auteurs comme Lynne Rudder Baker qui, bien que d'un matérialisme assumé, a l'intention de sauver la crédibilité des pratiques explicatives courantes qui ont dans leur antécédent causal des faits ou des événements psychologiques, économiques et sociaux. Voir Rudder Baker, L., 1995, *Explaining Attitudes,* Cambridge Mass, Cambridge University Press.

16. Traduit de l'anglais par Pascal Lavallée.