Diversité du représentationnalisme de la conscience

Paul Bernier

Volume 41, numéro 1, printemps 2014

URI : <https://id.erudit.org/iderudit/1025722ar>

DOI : <https://doi.org/10.7202/1025722ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Bernier, P. (2014). Diversité du représentationnalisme de la conscience. *Philosophiques*, 41(1), 37–56. <https://doi.org/10.7202/1025722ar>

Résumé de l'article

Cet article discute de diverses versions du représentationnalisme de la conscience. L'objectif principal est de défendre une interprétation de la théorie auto-représentationnelle de la conscience (TARC) selon laquelle le contenu d'un état mental conscient serait une proposition *de re* qui est constituée, en partie, par l'état mental conscient lui-même. Je souligne d'abord certains problèmes importants auxquels est confrontée une des théories de la conscience les plus influentes, soit la théorie représentationnelle de la conscience (TRC) et soutiens que la principale théorie rivale, soit la théorie de la conscience d'ordre supérieur (TCOS) doit lui être préférée. Je montre que les versions standards de la TCOS sont confrontées à un problème de régression à l'infini intolérable, et je propose la TARC comme une version non standard de la TCOS.

Diversité du représentationnalisme de la conscience

PAUL BERNIER

Université de Moncton

paul.bernier@umoncton.ca

RÉSUMÉ. — Cet article discute de diverses versions du représentationnalisme de la conscience. L'objectif principal est de défendre une interprétation de la théorie auto-représentationnelle de la conscience (TARC) selon laquelle le contenu d'un état mental conscient serait une proposition *de re* qui est constituée, en partie, par l'état mental conscient lui-même. Je souligne d'abord certains problèmes importants auxquels est confrontée une des théories de la conscience les plus influentes, soit la théorie représentationnelle de la conscience (TRC) et soutiens que la principale théorie rivale, soit la théorie de la conscience d'ordre supérieur (TCOS) doit lui être préférée. Je montre que les versions standards de la TCOS sont confrontées à un problème de régression à l'infini intolérable, et je propose la TARC comme une version non standard de la TCOS.

ABSTRACT. — This article discusses various versions of Consciousness Representationalism. Its main purpose is to defend an interpretation of the Self-Representational Theory of Consciousness (SRTC) according to which the content of a conscious state is a *de re* proposition which is constituted, in part, by the very conscious state itself. I first underscore some important problems for the Representational Theory of Consciousness (RTC), which is one of the most influential approach in the literature. I argue that the main rival theory, the Higher-Order Theory of Consciousness (HOTC) should be preferred. I argue that the standard versions of the HOTC are faced with an intolerable infinite regress and I propose the SRTC as a non-standard version of the HOTC.

MOTS CLÉS: Conscience phénoménale, qualia, représentationnalisme, théorie représentationnelle de la conscience, théorie de la conscience d'ordre supérieur, théorie auto-représentationnelle de la conscience

1. Introduction

Le problème de la conscience phénoménale concerne le caractère qualitatif de nos expériences conscientes ou ce que cela fait que d'avoir une expérience consciente particulière, pour reprendre l'expression consacrée de Thomas Nagel¹. Par exemple, lorsque vous faites l'expérience consciente de voir une tomate mûre sur le comptoir, il y a quelque chose que cela vous fait que de percevoir la couleur rouge de la tomate: votre expérience a un certain caractère qualitatif. On peut difficilement nier que le caractère qualitatif de votre expérience visuelle est quelque chose dont vous êtes conscient (*aware of*), au

1. Nagel, 1974.

sens où votre expérience est dirigée, du moins en partie, vers ce caractère qualitatif. C'est précisément cette intuition forte qui motive l'idée générale selon laquelle la conscience phénoménale serait essentiellement intentionnelle, c'est-à-dire qu'elle se réduirait à l'intentionnalité. Dans la littérature récente, les théories de la conscience phénoménale les plus influentes, qui s'inscrivent dans une telle perspective, sont des formes de représentationnalisme au sens où le caractère qualitatif d'une expérience consciente se réduirait à son contenu représentationnel. Ces diverses théories constituent donc ce qu'on peut appeler le représentationnalisme de la conscience phénoménale. Cette approche a pris deux formes principales soit : i) la théorie représentationnelle de la conscience (TRC), selon laquelle c'est l'objet du monde extérieur et ses propriétés, représentés dans l'expérience, qui permettent de rendre compte de la conscience phénoménale; et ii) la théorie de la conscience d'ordre supérieur (TCOS), selon laquelle c'est la représentation de l'état mental conscient lui-même qui serait cruciale pour une théorie de la conscience².

Mon principal objectif dans cet article est de défendre une version non standard de la TCOS, soit une version particulière de la théorie auto-représentationnelle de la conscience (ou TARC)³. La TARC diffère de façon importante des versions standards de la TCOS. En effet, selon ces dernières, pour qu'un état mental soit conscient, celui-ci doit être accompagné d'un état mental d'ordre supérieur numériquement distinct qui représente l'état mental conscient (d'ordre inférieur). Selon la TARC, pour qu'un état mental soit conscient, celui-ci n'a pas besoin d'être accompagné par un état mental distinct; il suffit que l'état mental se représente lui-même. Pour cette raison, les versions standards de la TCOS sont parfois appelées des approches à deux états, alors que la TARC est appelée une approche à un seul état.

Dans la section suivante, je rappelle certains problèmes notoires auxquels la TRC est confrontée, et dans la section 3 je souligne les avantages de la TCOS relativement à ces problèmes. Dans la section 4, je montre que les versions standards de la TCOS sont confrontées à un important problème de régression à l'infini. Dans la dernière section, après avoir noté que la TARC n'est manifestement pas confrontée à ce problème de régression à l'infini, je formule différentes interprétations de cette théorie et propose qu'une de ces interprétations doit être préférée, soit celle selon laquelle le contenu (auto) représentationnel est une proposition *de re*.

2. Pour la TRC, voir Dretske, 1995, 2003, Tye, 2000 et Hill, 2009. Pour les versions standards de la TCOS, voir Rosenthal, 1997, 2005, Lycan, 1996 et Carruthers, 2000. Notons que dans la littérature le terme « représentationnalisme » est normalement utilisé pour référer à la TRC. Cependant, la TCOS est également une forme de représentationnalisme comme je le précise plus loin.

3. Pour d'autres versions de la théorie auto-représentationnelle de la conscience (*Self-Representational Theory of Consciousness*), voir notamment Van Gullick, 2006 et Kriegel, 2009.

2. Certains problèmes notoires de la théorie représentationnelle de la conscience

Selon la TRC, une expérience consciente représente le monde comme étant d'une certaine manière, et ce que cela fait que d'avoir cette expérience est complètement épuisé par son contenu représentationnel. Considérons, par exemple, le cas où vous voyez une tomate mûre sur le comptoir. Il y a quelque chose que cela vous fait que d'avoir cette expérience, et ce que cela vous fait est différent de ce que cela vous ferait si vous regarderiez un ciel sans nuages. Comme vous êtes visuellement conscient de la tomate mûre sur le comptoir, et comme ce que cela vous fait que d'avoir cette expérience est quelque chose dont vous êtes conscient, il est plausible de croire que votre expérience consciente est un état représentationnel, soit un état qui représente le monde comme étant d'une certaine manière, et que le caractère qualitatif de votre expérience se réduit tout simplement au contenu représentationnel de l'expérience, à savoir que certaines propriétés sont instantciées par un certain objet particulier.

Comme un contenu représentationnel est quelque chose qui peut être vrai ou faux, la TRC a l'avantage de pouvoir rendre compte des cas d'illusion et d'hallucination. Lorsque Macbeth hallucine une dague en face de lui, il est dans un état mental conscient qui représente la présence d'une dague avec certaines propriétés, et cette représentation est tout simplement fausse. Selon la TRC, ce contenu représentationnel faux permet de rendre compte du caractère qualitatif de l'expérience de Macbeth. Cela dit, la TRC est confrontée à des problèmes notoires.

D'abord, il y a le problème des apparences visuelles⁴. Ce problème peut se présenter sous diverses formes; l'exemple suivant de Christopher Peacocke l'illustre bien :

Imaginez que vous vous trouvez dans une chambre et que vous regardez un coin de cette chambre formé par deux de ses murs. Ces murs sont couverts de papier peint d'une couleur uniforme (même teinte, même brillance, même saturation), mais un des murs est plus illuminé que l'autre. Dans ces circonstances, votre expérience peut représenter les deux murs comme étant de la même couleur : un des murs ne vous paraît pas être peint d'une couleur plus claire que l'autre. Cependant, c'est également un aspect de votre expérience visuelle elle-même que la région de votre champ visuel dans laquelle un des murs vous est présenté est plus claire que celle dans laquelle l'autre mur vous est présenté⁵.

Au sens proprement phénoménologique du mot « paraître », les murs vous paraissent comme étant de différentes couleurs, mais vous vous les représentez comme étant de la même couleur puisque vous ne croyez sûre-

4. Voir Bernier, 2013, pour une discussion détaillée de ce problème.

5. Peacocke, 1983, réédité dans Block *et al.*, 1997, p. 345.

ment pas qu'ils sont de différentes couleurs, sous peine de vous les représenter comme étant à la fois de la même couleur et de différentes couleurs. Selon la TRC, une expérience visuelle représente des propriétés objectives du monde, et rien d'autre. Par conséquent, la TRC semble incapable de rendre compte de la façon dont certaines choses peuvent nous apparaître, en un sens purement phénoménologique⁶.

Le deuxième problème important pour la TRC est une variante du problème précédent, qui s'appuie sur une expérimentation. Il s'agit du problème de la saillance phénoménale. Comme David Chalmers l'a souligné, lorsqu'un sujet regarde deux points rouges placés sur un fond noir, si le sujet déplace et focalise son attention vers un des deux points, il y a une différence significative dans le caractère qualitatif de ses expériences avant et après le déplacement de l'attention⁷. On ne peut pas rendre compte de cette différence par une différence de contenu représentationnel puisque la scène visuelle demeure la même et que le sujet ne rapporte aucun changement de la scène visuelle.

Le troisième problème est que la TRC est incapable de rendre compte de l'intuition selon laquelle nous avons un accès épistémique direct (non inférentiel) à nos propres états mentaux conscients. Quand un sujet sait ce que cela lui fait que de voir une tomate mûre sur le comptoir, il est tout à fait plausible que celui-ci n'est pas seulement dans une relation épistémique avec certains objets et propriétés de son environnement, mais qu'il est également dans une relation épistémique à son propre état mental conscient. Lorsque le sujet fait une telle expérience, dans des circonstances normales, il est certes justifié à accepter l'énoncé suivant :

(i) Il y a une tomate mûre sur le comptoir.

Mais il est également justifié à accepter :

(ii) Je vois qu'il y a une tomate mûre sur le comptoir.

Si, cependant, le caractère qualitatif de l'expérience est complètement capturé par le contenu représentationnel, qui porte uniquement sur le monde, il est alors difficile de comprendre comment le sujet serait justifié à accepter (ii). La seule réponse que la TRC peut donner est de dire que (ii) est indirectement justifié par une inférence à partir du contenu représentationnel de son expérience. Autrement dit, la TRC doit souscrire à une conception inférentialiste de la connaissance de soi⁸. Une telle conception de la connaissance de soi n'est cependant pas très plausible. Il est plutôt intuitif

6. Hill, 2009, a tenté de résoudre ce problème en soutenant que les apparences visuelles sont des propriétés de l'environnement qui sont représentées dans l'expérience; voir Bernier, 2013, pour un argument montrant que cette tentative de solution échoue.

7. Chalmers, 2010, p. 348-349.

8. Dretske, 2000, et Ludwig, 2005, ont tenté de défendre une telle conception inférentialiste; voir Aydede, 2002, et Bernier, sous presse, pour des arguments qui la réfutent.

de penser que nous avons un accès épistémique direct à nos états mentaux conscients. De plus, quelle forme pourrait prendre une telle inférence ? Il semble évident que ce ne peut pas être une inférence déductivement valide. Manifestement, (i) n'entraîne pas (ii). Il faudrait donc que ce soit une inférence non démonstrative. Mais là encore, le contenu de (i) est insuffisant pour appuyer (ii), d'autres prémisses sont requises. Si nous tentons de préciser ces prémisses additionnelles, alors les seules qui viennent à l'esprit sont des prémisses que nous utiliserions pour attribuer des états mentaux conscients à autrui, en troisième personne. De telles prémisses devraient inclure, par exemple, une description du système nerveux du sujet et une description des comportements publiquement observables que le sujet est disposé à avoir. Une telle stratégie revient à nier l'existence d'une importante asymétrie épistémique entre la connaissance de soi et la connaissance de l'esprit d'autrui, ce qui n'est pas très plausible⁹. Le fait que la TRC ait pour conséquence une conception inférentialiste de la connaissance de soi est donc un autre désavantage notoire de cette approche.

Finalement, comme Uriah Kriegel l'a souligné, ce que cela fait que d'avoir une expérience consciente comporte non seulement un caractère qualitatif mais également un caractère proprement subjectif¹⁰. Lorsqu'un sujet sait ce que cela fait que d'avoir une certaine expérience consciente, il est non seulement conscient de l'aspect proprement qualitatif du contenu de son expérience ou de certaines propriétés de l'environnement représentées par l'expérience, mais il est également conscient de ces propriétés comme lui étant présentées à lui-même, ou que le contenu de son expérience est un contenu « pour moi ». La TRC ne permet manifestement pas de rendre compte du caractère proprement subjectif de l'expérience puisqu'elle réduit ce que cela fait que d'avoir une expérience au seul caractère qualitatif, soit un contenu représentationnel visant des propriétés objectives du monde.

Ces divers problèmes suggèrent que ce que cela fait que d'avoir une expérience n'est pas complètement épuisé par le contenu représentationnel de l'expérience au sens de la TRC. Mais si le contenu représentationnel est entendu en un sens différent, le représentationnalisme pourrait être maintenu. L'idée qui vient naturellement à l'esprit est que ce qui est crucial pour le représentationnalisme de la conscience n'est pas tant la représentation d'objets ou de propriétés du monde, mais la représentation de l'état mental

9. Comme nous le savons, l'épistémologie contemporaine est marquée par un rejet de la conception cartésienne de la connaissance de soi, selon laquelle le sujet aurait une connaissance infallible de ses propres états mentaux. C'est une chose que de rejeter le cartésianisme épistémologique, et c'en est une autre que de soutenir qu'il n'existe pas d'asymétrie épistémique entre la connaissance de soi (ou l'accès épistémique en première personne) et la connaissance d'autrui (ou l'accès épistémique en troisième personne). Il est tout à fait possible d'accepter une telle asymétrie épistémique sans souscrire au cartésianisme épistémologique; voir, notamment, Davidson 1984 et 1987; Burge, 1988; C. Wright *et al.*, 1998.

10. Kriegel, 2009, p. 8; voir également Levine, 2006, p. 174.

conscient lui-même. Cela correspond justement à une thèse centrale de la théorie de la conscience d'ordre supérieur (TCOS).

3. La TCOS doit être préférée à la TRC

Une motivation fondamentale de la TCOS est de fournir un critère permettant de distinguer les états mentaux conscients des états mentaux inconscients. C'est là un attrait important de cette approche comparativement à la TRC. Alors que cette dernière vise essentiellement à réduire le caractère qualitatif de l'expérience consciente à un contenu représentationnel, elle demeure muette quant à ce qui différencie les états mentaux conscients et inconscients. En effet, il est tout à fait plausible que nous ayons des états mentaux inconscients qui représentent des objets et des propriétés de l'environnement. Pensons, par exemple, au cas de la vision aveugle (*blindsight*). Pourquoi n'y a-t-il pas quelque chose que cela fait que d'avoir de tels états mentaux inconscients bien que ceux-ci aient un contenu représentationnel ?

La TCOS repose sur le principe de transitivité selon lequel un état mental est conscient si, et seulement si, le sujet est conscient de cet état¹¹. Elle peut donc aisément rendre compte du caractère proprement subjectif de l'expérience consciente. Un état mental conscient est non seulement un état mental dont le sujet est conscient, mais c'est un état mental dont le sujet est conscient comme étant son propre état mental. C'est donc là un autre avantage de la TCOS sur la TRC.

Il est important de noter que même si la TCOS est la principale théorie rivale de la TRC, elle n'en est pas moins représentationnaliste. L'idée générale partagée par toutes les versions de cette approche est qu'un état mental est conscient si, et seulement s'il est l'objet d'un état mental d'ordre supérieur approprié. Comme je l'ai noté, selon les versions standards de la TCOS, l'état mental d'ordre supérieur et l'état mental d'ordre inférieur sont deux états mentaux numériquement distincts¹². L'état mental est l'objet d'un état mental d'ordre supérieur en ce sens que l'état mental d'ordre supérieur possède un contenu représentant l'état mental d'ordre inférieur. La TCOS est donc également représentationnaliste. Selon cette approche, le contenu représentationnel de l'état mental d'ordre supérieur ne représente donc pas des objets ou des propriétés de l'environnement physique externe, mais il représente certaines propriétés du sujet, à savoir toutes propriétés constitutives de l'état mental du sujet.

11. Rosenthal, 2005, p. 179.

12. Il existe trois versions standards de la TCOS ou approches à deux états : a) la théorie de la pensée d'ordre supérieur (*Higher-Order Thought Theory*) selon laquelle l'état mental d'ordre supérieur est une pensée, voir Rosenthal, 1997 et 2005 ; b) la version selon laquelle l'état mental d'ordre supérieur est de nature perceptuelle ou quasi-perceptuelle, voir Armstrong, 1968, 1978, et Lycan, 1996 ; c) la version dispositionnaliste, selon laquelle être « l'objet d'un état mental d'ordre supérieur approprié » doit être entendu au sens où l'objet est disponible pour un possible état mental d'ordre supérieur non actuel, voir Carruthers, 2000.

La TCOS n'est pas confrontée au problème des apparences et à celui de la saillance phénoménale. En effet, ces problèmes émergent en raison de la difficulté à identifier le caractère qualitatif d'un état mental conscient à un contenu représentant uniquement des propriétés physiques d'objets physiques externes. Mais, selon la TCOS, l'état mental est conscient dans la mesure où il est représenté par un état mental d'ordre supérieur. Dans l'exemple de Peacocke et celui de la saillance phénoménale, respectivement, deux états mentaux conscients distincts sont en jeu, c'est-à-dire deux états mentaux conscients qui ont des caractères qualitatifs différents. Ainsi, la TCOS a pour conséquence, de façon triviale, que ces deux états mentaux distincts sont conscients dans la mesure où ils sont représentés par deux états mentaux d'ordre supérieur qui ont des contenus représentationnels différents.

Finalement, la TCOS n'est pas engagée à souscrire à une conception inférentialiste de la connaissance de soi qui, comme je l'ai noté, est difficilement acceptable. Bien que la TCOS soit muette en ce qui concerne la connaissance de soi et l'asymétrie épistémique entre l'accès en première et l'accès en troisième personne, elle semble tout à fait compatible avec une telle asymétrie. La TCOS semble donc être préférable à la TRC puisqu'elle n'est pas confrontée aux principaux problèmes qui minent cette dernière et qu'elle offre un critère permettant de distinguer les états mentaux conscients des états mentaux inconscients. Cela étant dit, dans la mesure où la TCOS devrait permettre de rendre compte du caractère qualitatif des états mentaux conscients, les versions standards de cette théorie (ou approches à deux états) sont confrontées à un sérieux problème de régression à l'infini¹³.

4. La TCOS et le problème de la régression à l'infini

Il est notoire que les versions standards de la TCOS entraînent la possibilité d'états mentaux d'ordre supérieur vides, c'est-à-dire la possibilité qu'un sujet soit dans un état mental d'ordre supérieur représentant qu'il est dans un état sensoriel d'ordre inférieur alors qu'aucun état sensoriel ne se produit¹⁴. Les versions à deux états de la TCOS entraînent cette possibilité parce qu'un état mental de deuxième ordre est une entité dont l'existence est métaphysiquement indépendante de celle de l'état mental de premier ordre qu'il représente, et parce que par définition un contenu représentationnel peut être faux. Je peux être dans un état perceptuel dont le contenu est qu'il y a une dague en face de moi même si ce contenu est faux si, par exemple, j'hallucine. De la même manière, selon la TCOS je pourrais être dans un état mental de deuxième ordre dont le contenu est que je fais l'expérience d'une

13. Dans la section suivante, mon argument est formulé en discutant spécifiquement la théorie de la pensée d'ordre supérieur, mais cet argument général s'applique aux trois versions standards de la TCOS (voir la note 13).

14. Voir Block, 2011, Neander, 1998, Rosenthal, 2011 et Weisberg, 2011.

certaine sensation *S* même si ce contenu est faux parce que je ne suis dans aucun état sensoriel, auquel cas ce contenu serait faux parce qu'il est vide.

Ned Block a soutenu que la possibilité d'états mentaux de deuxième ordre vides rend la TCOS incohérente¹⁵. Le problème est que du point de vue du sujet les choses semblent être exactement comme s'il était effectivement dans un état sensoriel. C'est-à-dire qu'il y aurait quelque chose que cela fait que d'être dans l'état mental d'ordre supérieur alors qu'aucun état sensoriel ne se produit. David Rosenthal et Josh Weisberg ont tenté de résister à l'argument de Block en avançant qu'il n'y a aucune incohérence dans l'affirmation qu'il y a quelque chose que cela fait que d'être dans un état mental de deuxième ordre vide¹⁶. Par exemple, Rosenthal note: « Qu'il y ait quelque chose que cela fait que d'être dans un état mental est simplement l'apparence subjective d'être dans ce état¹⁷ », et il poursuit en notant qu'il n'y a aucune incohérence à dire qu'il peut apparaître au sujet qu'il est dans un état mental sensoriel alors qu'il ne l'est pas.

Lorsque nous réfléchissons à la nature des états mentaux conscients, nous sommes fortement enclins à nier la possibilité d'états mentaux d'ordre supérieur vides. Rosenthal est d'avis que ce sont nos intuitions à cet égard qui sont erronées. Cette erreur s'expliquerait parce que nous nous appuyons uniquement sur des rapports en première personne pour accepter l'occurrence d'états mentaux sensoriels. Nous acceptons ces rapports sans autres formes de procès et nous sommes justifiés de le faire parce qu'ils sont normalement corrects. Mais si nous imaginons un scénario dans lequel nous pourrions attribuer des états sensoriels autrement, alors l'incohérence se dissiperait. Supposons, pour les fins de l'argumentation, qu'un état sensoriel *S* est identique à un certain état cérébral *B*. Supposons que nous disposions d'une technologie nous permettant de scanner le cerveau du sujet et d'observer la présence de *B*. Dans un tel scénario, les rapports en première personne ne seraient plus l'unique source nous permettant d'accepter la présence ou l'absence de l'état *S*. Il ne serait alors aucunement incohérent d'imaginer un cas où le sujet dit faire l'expérience de *S* alors que notre scanner nous indique qu'aucun état sensoriel ne se produit. Comme Rosenthal l'indique, ce qui permet à un sujet de dire qu'il est dans un état sensoriel, d'un point de vue purement subjectif, est uniquement l'occurrence, dans le courant de sa conscience, d'une apparence subjective¹⁸. Rosenthal soutient que de telles apparences subjectives peuvent être trompeuses de la même manière que des apparences perceptuelles peuvent nous tromper sur ce qui se trouve dans notre environnement.

15. Block, 2011.

16. Rosenthal, 2011, et Weisberg, 2011.

17. Rosenthal, 2011, p. 433.

18. *Ibid.*, p. 432.

Cela dit, cette stratégie pour défendre la cohérence des états mentaux d'ordre supérieur vides est confrontée à un sérieux problème. En effet, elle entraîne une régression à l'infini d'états mentaux d'ordre supérieur vides. L'argument est le suivant :

- (1) Dans le cas d'un état mental d'ordre supérieur vide, il semble au sujet qu'il fait l'expérience d'un état sensoriel *S*.
 - (2) Si subjectivement il semble au sujet qu'il fait l'expérience de *S*, alors le sujet est dans un état mental conscient. Appelons cet état mental conscient *AS* (pour « apparence subjective »).
 - (3) *Ex hypothesis*, *AS* n'est pas *S*. (Pour les fins de l'argument, nous pouvons supposer que *AS* est un état doxastique.)
 - (4) Puisque *AS* est conscient alors, selon les versions standards de la TCOS, le sujet est conscient de *AS* en ce sens que *AS* est l'objet d'un état mental d'ordre supérieur approprié dont le contenu représentationnel est que le sujet est dans l'état *AS*. Appelons cet état mental d'ordre supérieur *AS-2* (où $AS-2 \neq AS$).
 - (5) Puisque les versions standards de la TCOS entraînent la possibilité d'états mentaux d'ordre supérieur vides, rien n'exclut la possibilité qu'*AS-2* soit lui-même un état mental d'ordre supérieur vide, c'est-à-dire qu'il se produise sans qu'*AS* ne se produise.
- (C) Cela entraîne une régression à l'infini d'états mentaux d'ordre supérieur vides. Si *AS-2* est vide, il semble bien au sujet qu'il ne l'est pas, et cette apparence subjective est à son tour un état mental conscient qui doit être l'objet d'un état mental d'ordre supérieur *AS-3* ($\neq AS-2$), et ainsi de suite à l'infini.

Cet argument est valide, mais est-il bien fondé ? Comme je l'ai indiqué, (1) est difficile à contester. Cette prémisse est d'ailleurs au cœur des argumentations de Rosenthal et de Weisberg pour défendre la cohérence des états mentaux d'ordre supérieur vides. Les prémisses (4) et (5) découlent directement des versions standards de la TCOS. Si on doit s'objecter à l'argument, on voudra sûrement tenter de rejeter la seconde prémisse en niant que le sujet est dans un état conscient *AS*. Par exemple, on pourrait dire que dans un état mental d'ordre supérieur vide, alors qu'il y a une apparence subjective dans le courant de la conscience du sujet, cette apparence subjective n'est pas elle-même un état mental conscient, contrairement à ce qu'affirme (2). L'apparence subjective pourrait être comprise comme étant une sorte d'état subliminal inconscient.

Cette objection à la seconde prémisse est pour le moins problématique. Comme Weisberg le note, dans un état mental d'ordre supérieur vide, « il semble au sujet que les choses sont exactement identiques à ce qu'elles

sont lorsque l'état mental d'ordre supérieur est accompagné par un état d'ordre inférieur¹⁹ ». Dans les deux cas, il y a dans le courant de la conscience du sujet la même apparence subjective. Autrement dit, ce que cela fait que d'être dans l'état mental d'ordre supérieur lorsque celui-ci est accompagné par l'état mental d'ordre inférieur est identique à ce que cela fait lorsque l'état d'ordre supérieur est vide. Dans le cas non vide cependant, il est difficile de nier que l'apparence subjective est consciente. Si on doit dire, comme le suggère l'objection, que dans le cas vide l'apparence subjective n'est pas consciente, alors nous devrions dire la même chose du cas non vide, ce qui n'est pas du tout plausible.

L'analogie avec l'hallucination visuelle est utile pour préciser cette réponse à l'objection précédente. On peut facilement imaginer que lorsque Macbeth hallucine une dague en face lui, les choses lui semblent être identiques à ce qu'elles seraient s'il voyait réellement une dague. L'apparence subjective dans le cas où Macbeth hallucine est la même que l'apparence subjective dans le cas où il verrait effectivement une dague. Mais dans le cas de l'hallucination, personne ne nierait que l'apparence subjective est consciente, ou ne nierait que l'apparence subjective est quelque chose dont Macbeth est conscient. Dans un tel cas, on ne comprendrait tout simplement pas que l'apparence subjective puisse être un état mental subliminal inconscient. La première conclusion à tirer est donc qu'il est tout à fait cohérent de parler de l'apparence subjective comme étant un état mental conscient, ou comme quelque chose dont le sujet est conscient. De plus, si un état mental d'ordre supérieur vide doit être compris comme une sorte d'hallucination d'un état mental d'ordre inférieur alors, tout comme dans le cas de l'hallucination, on doit comprendre l'apparence subjective d'un état mental d'ordre inférieur comme étant quelque chose dont le sujet est conscient, un état mental conscient comme la prémisse (2) l'énonce. Si on soutient que dans le cas de l'état mental d'ordre supérieur vide, l'apparence subjective n'est pas un état mental conscient, comme l'objection le propose, alors nous devrions également nier que dans le cas de l'hallucination l'apparence subjective est quelque chose dont le sujet est conscient, ce qui n'est pas plausible.

On pourrait tenter de rejeter la prémisse (2) d'une façon différente en admettant que, dans le cas d'un état mental d'ordre supérieur vide, lorsque le sujet rapporte que tout lui semble subjectivement comme s'il était dans un état sensoriel, ce dernier n'est en fait conscient que d'un contenu représentationnel qui est faux. Il est facile de voir que nous pouvons également répondre à cette objection à l'aide de l'analogie avec l'hallucination visuelle. Il y a un sens *non extensionnel* ou *intensionnel* de l'expression « conscient de » selon lequel Macbeth est conscient d'une dague en face de lui²⁰. Il est

19. Weisberg, 2011, p. 442.

20. Dans le sens *extensionnel*, « conscient de x » implique que x existe. Dans le sens *non extensionnel* ou *intensionnel*, cette expression n'a pas une telle implication.

tout à fait plausible d'interpréter ce sens non extensionnel de « conscient de » comme l'affirmation selon laquelle Macbeth est conscient uniquement du contenu représentationnel de son hallucination. Lorsque Macbeth est conscient du contenu représentationnel de son hallucination, les choses lui apparaissent exactement comme s'il voyait réellement une dague. Mais personne ne voudrait nier que lorsque Macbeth a une hallucination visuelle, il est dans un état mental. Appelons cet état mental H (pour hallucination). Personne ne nierait non plus que H est un état mental conscient. Selon les versions standards de la TCOS, H est conscient si et seulement s'il est accompagné par un état mental d'ordre supérieur approprié, disons $H-2$ ($\neq H$) qui représente que H se produit. Par conséquent, le sujet doit être conscient du contenu de $H-2$, c'est-à-dire que lui-même « voit » une dague. Considérons maintenant le cas d'une représentation d'ordre supérieur vide. Les choses apparaissent à Lucie exactement comme si elle était dans l'état sensoriel S alors qu'elle n'est dans aucun état sensoriel. Comment cela est-il possible ? Parce qu'elle est consciente d'un contenu représentationnel qui est identique au contenu représentationnel dont elle serait consciente si S se produisait. Mais si dans le cas de Macbeth nous acceptons que l'état mental H se produit et que H est conscient, comme je viens de le soutenir, alors nous devons également accepter que dans le cas où Lucie « hallucine » S , pour ainsi dire, elle est dans un état mental AS et, tout comme H est conscient, AS est également un état mental conscient. Par conséquent, cette deuxième façon de s'opposer à la prémisse (2) échoue également. L'argument montrant que les versions standards de la TCOS entraînent une régression à l'infini est donc bien fondé.

5. Une interprétation non standard de la TCOS: la théorie auto-représentationnelle

L'idée générale sur laquelle la théorie auto-représentationnelle de la conscience (TARC) repose n'est pas nouvelle²¹. La TARC est tout à fait dans l'esprit des vues de Franz Brentano sur la conscience. Le passage suivant de Brentano exprime bien l'idée centrale sur laquelle cette approche repose :

[Chaque acte de conscience] comprend une conscience de lui-même. Par conséquent, chaque acte [de conscience], aussi simple soit-il, possède un double objet, un objet primaire et un objet secondaire. L'acte le plus simple, par exemple celui d'entendre, a pour objet primaire le son, et pour objet secondaire, lui-même, le phénomène mental dans lequel le son est entendu²².

21. Caston, 2002, soutient qu'une version de la TARC est endossée par Aristote dans le *De Anima* (III.2, 425b12-25). Bernier, sous presse, défend une interprétation de la théorie réflexive de la conscience du philosophe bouddhiste Dignaga (v^e siècle) qui est également une version de la TARC.

22. Brentano, 1874/1973, p. 153-154.

La TARC peut être interprétée de diverses façons. Comme première approximation, nous pouvons la formuler de la façon suivante :

(AR) Un état mental M du sujet S est conscient si et seulement s'il existe un état mental M^* qui représente M , et $M^* = M$.

(AR) est manifestement une version de la TCOS puisque, comme je l'ai indiqué, selon cette dernière un état mental est conscient si et seulement s'il est l'objet d'un état mental d'ordre supérieur approprié. La différence cruciale entre (AR) et les versions standards de la TCOS est que, selon ces dernières, l'état mental d'ordre supérieur et celui d'ordre inférieur sont deux états numériquement distincts, alors que selon (AR) un seul et même état joue le rôle d'état à la fois d'ordre inférieur et d'ordre supérieur. L'usage d'« ordre supérieur » et d'« ordre inférieur » dans le contexte de la TARC sert uniquement à distinguer deux aspects, ou deux fonctions, d'un seul et même état mental, soit la fonction de représenter un objet primaire (le son) et celle de représenter l'objet secondaire (l'état mental lui-même). La TARC satisfait également le principe de transitivité selon lequel un état mental est conscient si et seulement si le sujet est conscient de cet état. Selon cette approche, en ayant un état conscient, le sujet est non seulement conscient d'un objet primaire, mais il est également conscient de son état mental. Par conséquent, à l'instar de la TCOS, et contrairement à la TRC, la TARC fournit un critère permettant de distinguer les états mentaux conscients de ceux qui ne le sont pas. Elle permet également de rendre compte du caractère proprement subjectif de l'expérience consciente, du fait que l'expérience est ressentie comme étant « pour moi » puisque, encore une fois, l'état mental est conscient non seulement parce que c'est un état dont le sujet est conscient, mais parce que c'est un état dont le sujet est conscient comme étant son propre état mental. Et, tout comme la TCOS, la TARC n'est pas engagée à une conception inférentialiste de la connaissance de soi. Finalement, il est évident que cette approche n'est pas confrontée au problème de la régression à l'infini puisqu'elle exclut la possibilité d'états mentaux d'ordre supérieur vides²³.

Dans certaines œuvres posthumes, Brentano a modifié sa théorie de la conscience de façon à tenir compte de l'idée voulant que, dans un état mental conscient, le sujet n'est pas tant conscient de son état mental en tant que tel que plutôt conscient de lui-même comme sujet engagé dans l'activité de représenter l'objet primaire²⁴. Mon but ici n'est pas de m'engager dans une exégèse de ces modifications et des motivations de Brentano qui les ont sus-

23. On pourrait penser que cela exclut la possibilité d'états mentaux inconscients. Mais ce n'est pas le cas car, selon la TARC, seuls sont conscients les états mentaux auto-représentationnels. Cela n'exclut en rien la possibilité qu'il y ait des états mentaux qui ne sont pas auto-représentationnels et qui sont, par conséquent, inconscients.

24. Voir, notamment, Brentano, 1954.

citées, mais deux choses importantes en ressortent clairement²⁵ : a) l'agent est entendu comme un sujet en un sens substantiel ; et b) l'état mental conscient est entendu comme une activité mentale. Cette conception modifiée de la conscience suggère la définition suivante :

(B) Un état mental *M* d'un sujet *S* est conscient si et seulement si *M* est une activité de *S* tel que en *M*-ant *S* représente un objet primaire *O*, et *S* est directement conscient de lui-même et de son activité de *M*-er.

Il est crucial de souligner que la clause « *S* est directement conscient de lui-même et de son activité de *M*-er » doit être comprise comme stipulant que le sujet n'est pas dans une relation représentationnelle avec lui-même et avec son activité mentale. Selon cette interprétation, le seul contenu représentationnel ne suffit pas à définir le fait que l'état mental est conscient. Une relation additionnelle est requise. La nature d'une telle relation peut paraître mystérieuse. On pourrait proposer, par exemple, qu'il s'agit d'une relation métaphysique primitive irréductible. Ainsi, bien que (B) soit dans l'esprit général de la TARC, cette interprétation n'est pas à strictement parler une approche auto-représentationnelle mais bien une approche auto-présentationnelle. Quoi qu'il soit, mon but n'est pas d'endorser (B), puisque cette approche me semble douteuse, particulièrement parce qu'elle repose sur un engagement ontologique à l'existence d'un sujet en un sens substantiel²⁶. La théorie modifiée de Brentano suggère cependant certaines interprétations intéressantes. Nous pouvons retenir les idées de Brentano voulant qu'un état mental conscient ait un objet primaire et un objet secondaire, et qu'un état mental conscient soit une activité, mais en faisant l'économie d'un engagement ontologique à l'existence d'un sujet en un sens substantiel.

La première interprétation est explicitement représentationnaliste. Elle a l'avantage de faire l'économie d'un sujet en un sens substantiel et de faire l'économie d'une relation métaphysique primitive. Il s'agit donc d'une version de la TARC au sens strict :

(AR-1) Un état mental *M* d'un sujet *S* est conscient si, et seulement si *M* est une activité de *S* qui représente à la fois un objet primaire *O* et le fait que *M* se produit.

Contrairement à (B), le fait que l'état mental soit conscient est complètement défini relativement au contenu représentationnel de l'état mental. Aucune relation additionnelle n'est requise.

25. Voir Fisette, manuscrit, pour une interprétation de ces développements des vues de Brentano.

26. Voir Metzinger, 2003, qui soutient que le soi n'existe pas et qu'il n'existe que des processus qui créent l'impression illusoire de l'existence d'un sujet. Voir également Bernier, 2010 et 2011, qui défendent la cohérence d'une ontologie dépourvue de sujets substantiels.

Selon l'approche générale de la TARC, un état mental conscient a un objet primaire et un objet secondaire. Certains défenseurs de cette approche ont suggéré qu'un état mental conscient aurait une structure complexe constituée de parties qui ont les fonctions respectives de représenter l'objet primaire et l'objet secondaire²⁷. Par contre, selon (AR-1), ce qui est complexe n'est pas l'état mental mais bien le contenu représenté. Selon (AR-1), *M* est une activité en vertu de laquelle plusieurs choses sont représentées.

On pourrait s'opposer à la TARC, comme certains l'ont fait, sur la base de l'idée suivant laquelle la relation entre la conscience d'un de nos états mentaux et l'état mental lui-même ne peut pas être une relation représentationnelle, puisqu'intuitivement une telle relation semble être beaucoup trop intime²⁸. Appelons cette intuition l'intuition de l'intimité. Cette relation ne pourrait pas être représentationnelle parce qu'elle n'impliquerait pas l'existence d'une brèche entre le véhicule de la représentation et son contenu. L'intuition de l'intimité suggère une autre interprétation de la TARC similaire à (B) qui, comme je l'ai noté, est une théorie auto-présentationnelle plutôt qu'auto-représentationnelle :

(AP) Un état mental *M* d'un sujet *S* est conscient si et seulement si *M* est une activité de *S* qui : a) représente un objet primaire *O* ; b) représente *S* ; et c) où *S* est directement conscient de son activité de *M*-er.

Tout comme dans (B), dire que *S* est directement conscient de son activité de *M*-er c'est dire que cette relation n'est pas une relation représentationnelle, ce qui permet de tenir compte de l'intuition de l'intimité. Encore une fois, contrairement à (AR-1), le seul contenu représentationnel ne suffit pas à définir le fait que l'état mental est conscient ; une relation additionnelle est requise. La clause c) est manifestement requise pour tenir compte de la relation à l'objet secondaire, mais pourquoi a-t-on besoin de la clause b) ? On en a besoin parce qu'il n'y aurait pas de sens à dire que *S* est conscient de son activité de *M*-er, comme le stipule la clause c), sans supposer que *S* est conscient de cette activité comme étant sa propre activité. C'est d'ailleurs ce qui permet à (AP) de rendre compte du caractère proprement subjectif des états mentaux conscients. Pour que le sujet soit conscient de son activité de *M*-er comme étant sa propre activité, il doit ou bien être directement conscient de lui-même, comme le stipule (B), ou bien être indirectement conscient de lui-même par une représentation de lui-même, comme le stipule la clause (b) dans (AP). La première option correspond à l'interprétation (B) qui, comme je l'ai indiqué, implique l'existence d'un sujet en un sens substantiel. La seconde option, soit (AP), permet de préserver une neutralité ontologique quant à l'existence d'un tel sujet en un sens substantiel, puisqu'elle est compatible avec l'idée selon laquelle il n'existe pas de sujet à

27. Voir Kriegel, 2009, p. 228, et Gennaro, 2006.

28. Voir Levine, 2001, 2006, et Hellie, 2007.

strictement parler, mais seulement une représentation qui renvoie à une fiction utile.

Cela étant dit, cette interprétation repose également sur une relation métaphysique primitive dont la nature reste mystérieuse. On pourrait penser, par exemple, qu'il s'agit d'une relation de connaissance directe (*acquaintance*) au sens de Russell²⁹. Mais il y a des raisons de douter de la plausibilité de cette façon de préciser la clause c) de (AP)³⁰.

Il y a cependant une autre interprétation qui permet de rendre compte de l'intuition de l'intimité dans un cadre purement représentationnaliste, c'est-à-dire sans avoir à invoquer une relation métaphysique primitive comme la connaissance directe russellienne. Cette interprétation est une version plus spécifique de (AR-1). (AR-1) laisse indéterminée la façon dont l'activité de *M-er* se représente elle-même. Elle pourrait se représenter à l'aide du concept de *M-er* ou à l'aide d'une description caractérisant un tel concept. Une interprétation plus spécifique de (AR-1), disons (AR-2), spécifie que l'activité de *M-er* se représente elle-même de façon indexicale.

(AR-2) Un état mental *M* d'un sujet *S* est conscient si et seulement si *M* est une activité de *S* qui : a) représente un objet primaire *O*; b) représente *S*; et c) se représente elle-même de façon indexicale.

Selon cette interprétation, le véhicule représentant le fait que l'activité de *M-er* se produit est constitué par un élément indexical référant directement à l'activité de *M-er*, de telle sorte que le contenu représentationnel de l'état mental est une proposition *de re* constituée, en partie, par l'activité même de représenter l'objet primaire et l'objet secondaire. Autrement dit, la référence à l'activité de *M-er* ne dépendrait pas d'un élément descriptif, ou conceptuel, représentant cette activité. Encore une fois, à l'instar de (AR-1) et contrairement à (B) et (AP), le fait que l'état mental est conscient est complètement caractérisé par son contenu représentationnel, aucune relation additionnelle n'est requise. Il est facile de voir que (AR-2) occupe une position mitoyenne entre (AR-1) et (AP). Il s'agit bien d'une version de (AR-1), mais une version qui préserve l'intuition de (AP) voulant que le sujet soit en contact direct avec son activité mentale consciente. Par conséquent, (AR-2) permet également de rendre compte de l'intuition de l'intimité. Mais puisque que cette interprétation est pleinement représentationnelle, elle fait l'économie d'une relation métaphysique primitive, comme par exemple la notion russellienne de connaissance directe (*acquaintance*) qui est requise par les interprétations (B) et (AP).

29. Voir Russell, 1911.

30. Voir Crane, 2012, pour un argument montrant que la stratégie qui consiste à utiliser la notion russellienne de connaissance directe (*acquaintance*) pour rendre compte de la relation entre un sujet et ses propres états mentaux conscients doit être rejetée.

On pourrait peut-être objecter que (AR-2) ne permet pas réellement de rendre compte de l'intuition de l'intimité parce qu'elle implique l'existence d'une brèche entre le véhicule de la représentation et son contenu. Cette remarque est sans doute juste, mais cette brèche n'est pas pertinente dans la mesure où le véhicule représentationnel, soit l'activité de *M-er* elle-même, fait partie du contenu représentationnel. Plus précisément, selon (AR-2) alors qu'il existe bien une brèche entre le véhicule représentationnel et son contenu — dans la mesure où le véhicule représentationnel est constitué par un élément indexical et par des éléments descriptifs, ou conceptuels, comme par exemple le concept *je* et une certaine structure compositionnelle — il n'existe pas de brèche entre le véhicule représentationnel et la façon dont celui-ci contribue à former le contenu représentationnel complet, puisque ce véhicule n'est pas représenté par un élément descriptif et que l'élément indexical permet que l'activité même de *M-er* soit une partie constitutive du contenu représenté. (AR-2) est donc une interprétation attrayante puisqu'elle permet de préserver les vertus de (AP) et de (AR-1) en évitant leurs défauts. (AR-2) me semble donc être la version la plus prometteuse de la TARC.

Notons deux objections qu'on pourrait être tenté d'adresser à la TARC, au sens général de cette approche. La première repose sur la thèse de la transparence mooréenne qui est typiquement invoquée dans le contexte de la TRC³¹. Selon cette objection, lorsque je considère par introspection un de mes états mentaux conscients, la seule chose dont je suis conscient serait l'objet visé par l'état mental. Comment pourrais-je alors être conscient de moi-même et de mon activité représentationnelle? La seconde objection est que la relation auto-représentationnelle serait mystérieuse et qu'elle ferait de la conscience une propriété intrinsèque des états mentaux conscients. Par conséquent, la TARC exclurait a priori toute possibilité de naturalisation de la conscience³². Considérons ces objections tour à tour.

Pour les philosophes qui endossent la TRC, la transparence mooréenne joue le rôle d'un axiome fondamental non négociable. Selon cette thèse, lorsque je considère par introspection un de mes états mentaux conscients, disons ma perception visuelle d'un arbre, je ne suis conscient que du contenu de ma perception, ou de ce que je vois, soit l'arbre et ses diverses propriétés. Mais cette thèse n'est pas parfaitement claire. Ce qui n'est pas clair, c'est ce qu'on entend par la notion d'introspection sur laquelle elle repose. Il semble que la notion d'introspection invoquée corresponde en fait

31. Voici une citation de Gilbert Harman qui résume bien la thèse de la transparence mooréenne: «Regardez un arbre et essayez de diriger votre attention sur les aspects intrinsèques de votre expérience visuelle. Je prédis que les seuls aspects sur lesquels vous pourrez diriger votre attention seront des aspects de l'arbre qui vous est présenté» (Harman, 1990, réédité dans Block *et al.*, 1997, p. 667).

32. Une objection similaire a été formulée par David Rosenthal qui propose qu'une motivation cruciale de la TCOS est d'être une théorie qui permet de réduire la conscience à une relation contingente entre deux états mentaux numériquement distincts; voir Rosenthal, 1997.

à une activité intellectuelle très abstraite, car dire qu'on peut être dans une relation introspective à un état mental particulier, comme la perception visuelle d'un arbre, en faisant abstraction de l'activité mentale consciente complexe où elle s'insère, est une idéalisation simplificatrice. En effet, on peut difficilement nier que lorsque nous *pensons* à diverses expériences conscientes, nous sommes frappés par le fait que la plupart d'entre elles sont normalement tournées uniquement vers le monde extérieur. Cela n'a rien de très surprenant puisque nos expériences conscientes ont comme fonction principale de nous informer sur notre environnement, ce qui est crucial à notre survie. L'introspection est une activité par laquelle nous tournons volontairement notre attention vers notre propre activité mentale, vers ce qui se passe dans le courant de notre conscience, pour reprendre l'expression de William James. L'introspection n'est pas quelque chose que nous faisons habituellement. Si, par contre, nous faisons l'effort volontaire soutenu de tourner notre attention vers notre vie mentale consciente, nous réalisons rapidement qu'il est relativement difficile de focaliser notre attention sur des expériences conscientes particulières. Plus spécifiquement, il n'est pas facile de focaliser notre attention sur une expérience consciente, comme par exemple la perception visuelle d'un arbre, en faisant abstraction du contexte de l'activité consciente complexe où elle s'inscrit. Pour cette raison, il semble un peu trompeur de dire que, dans l'introspection d'une telle expérience visuelle, le sujet serait uniquement conscient de l'arbre et de ses propriétés. Comme je l'ai noté, si nous faisons l'effort de nous engager dans une activité d'introspection, nous nous rendons compte que lorsque nous avons, par exemple, une expérience visuelle, il y a beaucoup de choses qui se produisent en arrière-plan dans le courant de notre conscience. Plus précisément, ce dont nous nous rendons compte est que notre attention se déplace constamment; elle peut par exemple se déplacer sur notre humeur. Essayez de focaliser votre attention sur un arbre pendant une assez longue période de temps, je prédis que ce que cela vous fera que de faire une telle expérience variera en fonction de la durée. Finalement, comme je l'ai déjà indiqué dans ma discussion de la TRC, si l'introspection de ma perception visuelle d'un arbre ne révélait que l'arbre et ses propriétés, alors je ne pourrais uniquement savoir que je fais cette expérience qu'à l'aide d'une inférence à partir du contenu de l'expérience. Et, comme je l'ai soutenu, une telle conception inférentialiste de la connaissance de soi est inacceptable.

Selon la seconde objection, la relation auto-représentationnelle serait une relation mystérieuse, ce qui suggère que la TARC serait incapable de produire une explication naturaliste de la conscience. Comme nous le savons, bien que le projet de naturalisation de la conscience soit tout à fait souhaitable, il n'y a pas encore de consensus philosophique quant à savoir quelle approche permet de mener à bien ce projet, ou même à savoir si ce projet pourra effectivement être mené à bien. Les versions standards de la TCOS ont certes l'avantage de proposer une théorie qui va dans le sens de

la naturalisation de la conscience, mais ce projet demeure une question ouverte. Cette seconde objection serait sérieuse seulement si on pouvait démontrer que la TARC entraîne une forme de dualisme ou qu'elle est incompatible avec la naturalisation de la conscience. Mais pourquoi devrions croire qu'il en est ainsi ? À ce que je sache, aucun argument n'a été proposé pour montrer cela. Mon but ici n'est pas de montrer que la TARC peut être développée dans le cadre du projet de naturalisation de la conscience, mais rien ne semble exclure cette possibilité a priori³³.

Le représentationnalisme de la conscience semble offrir la meilleure stratégie générale pour développer une théorie de la conscience phénoménale. J'ai souligné quatre problèmes importants auxquelles la TRC est confrontée : a) le problème des apparences, b) le problème de la saillance phénoménale, c) le problème du caractère proprement subjectif de l'expérience consciente, et d) le problème de la connaissance de soi. J'ai soutenu que la version à un seul état de la TCOS, soit la TARC, doit être préférée aux versions standards, en raison d'un problème de régression à l'infini que celles-ci entraînent. J'ai proposé diverses interprétations de la TARC, qui reposent sur une suggestion de Brentano, et j'ai soutenu que la plus plausible de ces interprétations est celle suivant laquelle le contenu représentationnel d'un état mental conscient est une proposition *de re* constituée en partie par l'activité mentale consciente³⁴.

Bibliographie

- Armstrong, David M. « What is Consciousness ? », *Proceedings of the Russellian Society*, vol. 3, 1978, p. 65-76.
- . *A Materialist Theory of Mind*, London, Routledge and Kegan Paul, 1968.
- Aydede, Murat. « Is Introspection Inferential ? », dans B. Gertler, dir., *Privileged Access*, Aldershot, Ashgate Publishing, 2002, p. 55-64.
- Bernier, Paul, sous presse. « Dignaga on Reflexive Awareness », *Philosophy East & West*.
- . « La théorie représentationnelle de la conscience phénoménale et le problème des apparences visuelles », *Dialogue*, 52, 2013, p. 1-23.
- . « Is the Buddhist Doctrine of Non-Self Conceptually Incoherent ? », *Buddhist Studies Review*, 28, 2011, p. 187-202.
- . « La pensée sans sujet pensant », *Dialogue*, 49, 2010, p. 1-14.
- Block, Ned. « The Higher Order Approach to Consciousness is Defunct », *Analysis*, 71, 2011, p. 419-431.
- Block, Ned, O. Flanagan et G. Güzeldere, dir. *The Nature of Consciousness*, Cambridge (Mass.), MIT Press, 1997.

33. Voir Kriegel, 2009 qui soutient que la TARC s'inscrit effectivement dans le projet de naturalisation de la conscience ; voir particulièrement les chapitres VI à VIII.

34. Je remercie Rocco Gennaro, David Rosenthal, Daniel Shargel et Josh Weisberg pour leurs commentaires sur une version antérieure de l'argument de la régression à l'infini.

- Brentano, Franz. *Religion und Philosophie*, édition de F. Mayer Hillebrand, Bern, Francke, 1954.
- Brentano, Franz. 1874. *Psychology from an Empirical Standpoint*, trad. par A. C. Rancurello, D. B. Terrell et L. L. McAlister, Londres, Routledge & Kegan Paul, 1973.
- Burge, Tyler. « Individualism and Self-Knowledge », *The Journal of Philosophy*, 85, 1988, p. 649-663.
- Carruthers, Peter. *Phenomenal Consciousness*, Cambridge, Cambridge University Press, 2000.
- Caston, Victor. « Aristotle on Consciousness », *Mind*, 111, 2002, p. 751-815.
- Chalmers, David J. « The Representational Character of Consciousness », dans *The Character of Consciousness*, Oxford, Oxford University Press, 2010, p. 339-379.
- Crane, Tim. « Tye on Acquaintance and the Problem of Consciousness », *Philosophy and Phenomenological Research*, 84, 2012, p. 190-198.
- Davidson, Donald. « Knowing One's Own Mind », *Proceedings and Addresses of the American Philosophical Association*, 60, 1987, p. 441-458.
- . « First-Person Authority », *Dialectica*, 38, 1984, p. 101-112.
- Dretske, Fred. « Experience as Representation », *Philosophical Issues*, 13, 2003, p. 67-82.
- . « The Mind's Awareness of Itself », dans *Perception, Knowledge and Belief*, Cambridge, Cambridge University Press, 2000, p. 158-177.
- . *Naturalizing the Mind*, Cambridge (Mass.), MIT Press, 1995.
- Fisette, Denis, manuscrit. « Franz Brentano and Higher-Order Theories of Consciousness ».
- Gennaro, Rocco J. « Between Pure Self-Referentialism and the Extrinsic HOT Theory of Consciousness », dans U. Kriegel et K. Williford, dir., *Self-Representational Approaches to Consciousness*, Cambridge (Mass.), MIT Press, 2006, p. 221-248.
- Harman, Gilbert. « The Intrinsic Quality of Experience », *Philosophical Perspectives*, vol. 4, Oxford, Blackwell, 1990, p. 31-42. Réédité dans N. Block, O. Flanagan et G. Güzeldere, dir., *The Nature of Consciousness*, Cambridge (Mass.), MIT Press, 1997, p. 663-675.
- Hellie, Benj. « Higher-Order Intentionality and Higher-Order Acquaintance », *Philosophical Studies*, 134, 2007, p. 289-324.
- Hill, Christopher. *Consciousness*, Cambridge, Cambridge University Press, 2009.
- Kriegel, Uriah. *Subjective Consciousness*, Oxford, Oxford University Press, 2009.
- Kriegel, Uriah et K. Williford, dir. *Self-Representational Approaches to Consciousness*, Cambridge (Mass.), MIT Press, 2006.
- Levine, Joseph. « Conscious Awareness and (Self) Representation », dans U. Kriegel et K. Williford, dir., *Self-Representational Approaches to Consciousness*, Cambridge (Mass.), MIT Press, 2006, p. 173-197.
- Levine, Joseph. *Purple Haze*, Oxford, Oxford University Press, 2001.
- Ludwig, Pascal. « Une défense hétérodoxe de la conception inférentialiste de l'introspection », *Dialogue*, 44, 2005, p. 123-144.
- Lycan, William. *Consciousness and Experience*, Cambridge (Mass.), MIT Press, 1996.
- Metzinger, Thomas. *Being No One*, Cambridge (Mass.), MIT Press, 2003.

- Nagel, Thomas. « What Is It Like to Be a Bat? », *Philosophical Review*, 83, 1974, p. 435-450.
- Neander, Karen. « The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness », *Philosophical Perspectives*, 12, 1998, p. 411-434.
- Peacocke, Christopher. *Sense and Content*, Oxford, Clarendon Press, 1983. Réédité dans N. Block, O. Flanagan et G. Güzeldere, dir., *The Nature of Consciousness*, Cambridge (Mass.), MIT Press, 1997, p. 341-354.
- Rosenthal, David M. « Exaggerated Reports: Reply to Block », *Analysis*, 71, 2011, p. 431-37.
- . « Sensory Qualities, Consciousness and Perception », dans *Consciousness and Mind*, Oxford, Oxford University Press, 2005, p. 175-226.
- . « A Theory of Consciousness », dans N. Block, O. Flanagan et G. Güzeldere, dir., *The Nature of Consciousness*, Cambridge (Mass.), MIT Press, 1997, p. 729-753.
- Russell, Bertrand. *The Problems of Philosophy*, Oxford, Oxford University Press, 1912.
- . « Knowledge by Acquaintance and Knowledge by Description », *Proceedings of the Aristotelian Society (New Series)*, Vol. XI, 1911, p. 108-128.
- Tye, Michael. *Consciousness, Color and Content*, Cambridge (Mass.), MIT Press, 2000.
- Van Gulick, Robert. « Mirror Mirror — Is that All? », dans U. Kriegel et K. Williford, dir., *Self-Representational Approaches to Consciousness*, Cambridge (Mass.), MIT Press, 2006, p. 11-39.
- Weisberg, Josh. « Abusing the Notion of What-it's-like-ness: A Response to Block », *Analysis*, 71, 2011, p. 438-443.
- Wright, Crispin, B. C. Smith et C. Macdonald, dir. *Knowing our Own Minds*, Oxford, Clarendon Press, 1998.