

Sens et neurosciences

Étude critique de Paul Thagard, *The Brain and the Meaning of Life*, Princeton University Press, 2010, 274 pages

Jean-Frédéric de Pasquale

Volume 38, numéro 2, automne 2011

URI : <https://id.erudit.org/iderudit/1007465ar>

DOI : <https://doi.org/10.7202/1007465ar>

[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

de Pasquale, J.-F. (2011). Sens et neurosciences / Étude critique de Paul Thagard, *The Brain and the Meaning of Life*, Princeton University Press, 2010, 274 pages. *Philosophiques*, 38(2), 579–595. <https://doi.org/10.7202/1007465ar>

# Étude critique

## Sens et neurosciences

Étude critique de Paul Thagard, *The Brain and the Meaning of Life*, Princeton University Press, 2010, 274 pages.

JEAN-FRÉDÉRIC DE PASQUALE

Université du Québec à Montréal

jfdepasquale@yahoo.com

*La troisième objection est la suivante : vous recevez d'une main ce que vous donnez de l'autre ; c'est-à-dire qu'au fond les valeurs ne sont pas sérieuses, puisque vous les choisissez. À cela je réponds que je suis bien fâché qu'il en soit ainsi ; mais si j'ai supprimé Dieu le père, il faut bien quelqu'un pour inventer les valeurs. Il faut prendre les choses comme elles sont. Et par ailleurs, dire que nous inventons les valeurs ne signifie pas autre chose que ceci : la vie n'a pas de sens a priori. Avant que vous ne viviez, la vie, elle, n'est rien, mais c'est à vous de lui donner un sens et la valeur n'est pas autre chose que ce sens que vous choisissez. (J.-P. Sartre, *L'existentialisme est un humanisme*, dans Nagel, 1970, p. 89-90)*

Sartre se trompait : nous ne créons pas le sens, et pas non plus les valeurs. Le sens et les valeurs trouvent leur source dans notre neurobiologie. Une philosophie informée par les neurosciences permet de répondre aux questions fondamentales de la philosophie traditionnelle : Quel est le sens de la vie ? Qu'est-ce que la réalité ? Comment la connaissons-nous ? Qu'est-ce qui rend une action bonne ou mauvaise ?

C'est la thèse que défend Paul Thagard dans *The Brain and The Meaning of Life*<sup>1</sup>. S'il n'est pas le premier à approcher la question du sens de la vie dans un cadre analytique ou naturaliste, il nous offre quelque chose qui est relativement rare de nos jours : un livre qui résume l'essentiel de la pensée philosophique de l'auteur sur les principaux problèmes de la philosophie en un *système*, ou quelque chose qui y ressemble beaucoup. « Les systèmes complets de philosophie ne sont plus à la mode, mais j'essaie de montrer comment s'agencent, les unes avec les autres, mes conclusions à propos du réalisme, de la cohérence, des conséquences morales, et des multiples dimensions du sens de la vie » (p. 111). La vision du monde développée dans le livre de Thagard constitue un système au moins au sens suivant : 1) elle touche à la plupart des domaines importants de la philosophie ; 2) elle est raisonnée et utilise une méthodologie unifiée (ici un naturalisme cohérentiste et neural) ; 3) ses parties « s'agencent les unes avec les autres ». Le tout forme selon Thagard un « argument étendu » (*extended argument*, p. xii)

---

1. Thagard, cependant, ne discute pas Sartre, mais Camus.

pour démontrer la pertinence des neurosciences dans les domaines traditionnels de la philosophie<sup>2</sup>.

En conséquence de ce qui précède, les trois principaux buts de la présente étude critique seront les suivants. Premièrement, nous évaluerons les solutions, souvent neuroscientifiques, aux problèmes traditionnels abordés par Thagard. Deuxièmement, nous essaierons de faire ressortir cet aspect (quasi-) systématique dans la description de l'ouvrage; et troisièmement, nous évaluerons à quel point le développement du système démontre la pertinence de la méthode du naturalisme neural.

Ceci est l'évaluation d'un système de philosophie dont la méthodologie est un naturalisme *neural*; c'est pourquoi la priorité ira ici à la discussion de l'aspect neuroscientifique des arguments, même si d'autres niveaux d'explication sont examinés dans le livre (en accord avec le pluralisme de l'auteur). De plus, c'est un *système* de philosophie naturaliste: se concentrer sur l'analyse d'un seul argument serait manquer le but de l'ouvrage. Ne vous surprenez donc pas si des facettes variées du système sont abordées. La tâche de l'analyste est évidemment plus facile que celle du penseur synthétique, et c'est en toute modestie qu'elle est entreprise ici.

Le livre de Thagard touche à presque tous les grands domaines de la philosophie, de l'épistémologie à la philosophie politique en passant par la philosophie de l'esprit, la morale et la théorie de la décision. Nous essaierons de donner une idée des sujets traités dans chaque chapitre, en nous concentrant sur ceux qui nous semblent les plus importants pour son projet d'un système basé sur le naturalisme neural.

Le premier chapitre s'ouvre sur le constat posé par Camus selon lequel le point de départ de la philosophie est la question du suicide, qui à son tour peut être vue comme la question de savoir si la vie a un sens et si oui, lequel. Thagard se propose de répondre à cette question ainsi qu'à d'autres questions essentielles: Qu'est-ce que la réalité? Comment la connaissons-nous? Qu'est-ce qui fait qu'une action est bonne ou mauvaise? La connaissance des choses importantes (Quelles sont-elles? Pourquoi sont-elles importantes? Et comment les réaliser?) est ce qu'il appelle la sagesse. La philosophie peut apporter de la sagesse, mais elle doit suivre une méthode particulière, le naturalisme neural.

Contrairement à Camus, Thagard considère que les problèmes liés à la connaissance doivent être résolus avant qu'on puisse répondre à la question du sens de la vie. Par conséquent, le second chapitre traite de la nature de la connaissance, en particulier de l'opposition entre la science et la foi; l'arbitraire de la foi y est contrasté avec les contrôles rigoureux de la méthode scientifique, et celle-ci est illustrée par une intéressante discussion de la

---

2. Thagard n'est évidemment pas le seul à espérer que les neurosciences renouvellent la philosophie, quoique le caractère systématique de l'entreprise et l'intérêt pour les questions traditionnelles soient un trait remarquable de son ouvrage.

médecine basée sur les données (*evidence-based medicine*). La plupart des lecteurs familiers avec son œuvre ne seront pas surpris de voir Thagard rejeter la foi. Mais sans la religion (et sans l'âme, rejetée au chapitre suivant), il faudra se demander ce qui donnera sens à la vie, puisque ce ne peut être la promesse d'une vie éternelle. La position que Thagard adopte sur la connaissance est peut-être plus controversée à cause de son réalisme : la connaissance est basée sur l'inférence à la meilleure explication, qui maximise la cohérence explicative avec les données (cela entre en contradiction avec, par exemple, l'empirisme constructif de van Fraassen). C'est entre autres parce que les entités de la foi ne sont pas nécessaires pour expliquer les données du monde observable que la religion doit être rejetée.

Il est important de décrire la notion de cohérence développée par Thagard. C'est important parce que c'est une partie de la méthodologie philosophique utilisée, mais aussi parce que le « cerveau » du titre est celui des neurosciences *théoriques* autant que celui des neurosciences *empiriques*<sup>3</sup>. Les premières fournissent des modèles de la manière dont les processus cognitifs, perceptuels ou émotionnels peuvent être produits par des réseaux de neurones ; les secondes fournissent les données sur les structures du cerveau impliquées dans ces processus. Le cerveau, pour Thagard, est une machine qui produit de la satisfaction parallèle de contraintes, c'est-à-dire une machine à résoudre des problèmes de cohérence — ce qui est en même temps la meilleure façon de connaître la réalité ou de prendre des décisions. Pour mieux comprendre, voyons un exemple (tiré du chapitre II).

Soit un patient qui se présente chez son médecin avec deux symptômes : le symptôme  $E_1$ , un mal de ventre, et le symptôme  $E_2$ , une fièvre. Le patient admet avoir mangé de la nourriture exotique la veille. Le médecin propose deux hypothèses, l'hypothèse  $H_1$  selon laquelle le patient a une infection bactérienne, hypothèse elle-même expliquée par l'hypothèse  $H_3$  selon laquelle le patient a mangé de la nourriture contaminée ; et l'hypothèse  $H_2$  selon laquelle le patient a un ulcère<sup>4</sup>. L'hypothèse  $H_1$  et l'hypothèse  $H_2$  étant en conflit, il s'agit de déterminer laquelle est la meilleure explication. L'hypothèse  $H_1$  explique la présence de fièvre et de mal de ventre, alors que l'hypothèse  $H_2$  n'explique que le mal de ventre. De plus l'hypothèse  $H_1$  a elle-même une explication. Dans ces conditions, on peut penser que le médecin sera porté à privilégier l'hypothèse  $H_1$ .

Notons qu'on peut formaliser complètement le problème sous la forme d'un graphe.  $H_1$  a des liens explicatifs positifs avec  $E_1$  et  $E_2$ , alors que  $H_2$

---

3. On fait parfois une distinction entre connexionnisme et neurosciences théoriques ; il suffira ici de dire que les récents travaux de Thagard sur la cohérence font le pont entre les deux.

4. L'exemple du chapitre II est évidemment délibérément simplifié. Comme Thagard a déjà écrit sur le sujet, on ne peut lui reprocher d'ignorer que les ulcères sont causés par une bactérie.

n'en a qu'avec E1. H3 a un lien explicatif positif avec H2. Et H1 et H2 étant en conflit pour l'explication des mêmes données, elles ont un lien explicatif négatif. Pour résoudre le problème, il faut *maximiser la cohérence*, c'est-à-dire que si deux éléments ont un lien positif — une contrainte positive —, ils doivent autant que possible être acceptés ou refusés en même temps, et inversement pour les contraintes négatives. Les données ont dans ce cadre une priorité que n'ont pas les hypothèses. Faute d'espace suffisant pour présenter la théorie formelle de la cohérence (consulter Thagard 2000 à ce sujet), nous dirons simplement que le formalisme des graphes se traduit directement dans celui des réseaux de neurones connexionnistes récurrents — les propositions devenant des neurones, les contraintes des connexions — et que ceux-ci forment une heuristique efficace pour solutionner le problème de cohérence associé. Thagard en a fait la démonstration en appliquant un tel réseau à des problèmes de cohérence explicative tirés de l'histoire des sciences (modèle ECHO; Thagard 1992) et a montré qu'on pouvait étendre le formalisme à d'autres types de problèmes comme la décision (Thagard 2000). Comme on le verra, Thagard lui a apporté certaines améliorations cruciales dernièrement.

Revenons à l'évaluation des différentes méthodes faite par Thagard au chapitre II. Une des propositions radicales de ce chapitre est le rejet des méthodes philosophiques traditionnelles. Les vieilles méthodes de l'apriorisme ou de l'analyse conceptuelle sont impuissantes; la nécessité supposément révélée par les expériences de pensée est une illusion. Ce qui est important, c'est de produire de meilleurs concepts, pas de les analyser. On peut penser que le rejet des expériences de pensée et de l'analyse conceptuelle est exagéré. Certes, on ne peut qu'être embarrassé devant certains abus des expériences de pensée; qu'on songe seulement à celle de la chambre chinoise, où une expérience de pensée faisant appel à nos préjugés est supposée nous révéler des faits importants sur le monde ou sur les concepts appropriés de représentation et de compréhension que doivent utiliser les sciences cognitives<sup>5</sup>. Cette expérience de pensée a malheureusement été prise très au sérieux par les psychologues, ce qui montre que le philosophe doit être prudent. Cependant, il n'est pas nécessaire de souscrire à l'idée que les

---

5. On peut trouver que cette caractérisation de l'expérience de pensée de Searle est radicale; mais l'auteur de ces lignes doit avouer qu'il ne voit aucune raison de la rétracter. Selon les deux possibilités viables et intéressantes d'applications du résultat (un fait du monde ou le concept de compréhension approprié aux sciences cognitives) les intuitions en question ne sont que des préjugés (on pourrait limiter l'application de l'expérience au concept populaire de compréhension, mais il n'y a aucune raison pour que ce concept soit celui qui soit approprié à une science de l'esprit). Bien que la position de Thagard lui-même soit moins définitive sur ce cas précis, les motifs de sa condamnation des expériences de pensée semblent parfaitement s'y appliquer. Voir (Poirier et Payette 2007) pour la question de la révision des concepts naïfs par le processus scientifique et (Ladymann et Ross 2007, chapitre I) pour un rejet du rôle joué par les intuitions en philosophie.

expériences de pensée nous donnent accès à des vérités nécessaires pour penser qu'elles jouent un rôle important dans des domaines scientifiques. En biologie, par exemple, la théorie de la sélection naturelle valide les raisonnements d'optimalité et l'utilisation du formalisme de la génétique des populations contraint fortement les raisonnements sur les situations hypothétiques. Les expériences de pensée permettent aussi parfois de clarifier certaines distinctions conceptuelles — par exemple entre hérédité et déterminisme génétique, voir (Lewontin 1974) — sans préjuger de leur nécessité. On accordera sans problème à Thagard que les philosophes se doivent surtout de proposer de meilleurs concepts et non simplement de décrire l'usage courant; mais si les concepts sont des conventions, et les analyses conceptuelles des tentatives d'établir des conventions (Brandon, 2005), il faut pour qu'elles soient adoptées, qu'elles soient au moins aussi cohérentes que possible avec l'usage, c'est-à-dire qu'elles soient, comme disait Carnap, des « explications » des anciens concepts.

On a vu que Thagard rejetait les entités de la religion parce que la science permettait d'en faire l'économie dans l'explication des données empiriques. Une de ces entités est l'âme; si on veut pouvoir résoudre empiriquement la question du sens de la vie, il faut que l'esprit puisse être étudié scientifiquement, ce qui est impossible s'il est un objet surnaturel. Le troisième chapitre se penche sur la relation entre les esprits et les cerveaux, relation d'identité pour Thagard. Ici, le lecteur déjà naturaliste ne découvrira rien qui le convaincra davantage dans les arguments contre le dualisme — et le lecteur wittgensteinien, s'il n'a pas déjà jeté le livre au bout de ses bras, devra suspendre son scepticisme<sup>6</sup>. Thagard reconnaît un seul argument convainquant pour le dualisme, la difficulté d'expliquer la conscience, une faiblesse qu'il tentera de surmonter au chapitre V.

L'identification des esprits aux cerveaux pourrait causer un problème pour la philosophie naturaliste de Thagard: en effet, et si nos connaissances du cerveau étaient incompatibles avec la possibilité pour le cerveau d'atteindre une connaissance objective de la nature? Le chapitre IV présente la manière dont les cerveaux parviennent à connaître la réalité. Thagard défend ici un réalisme constructif. Réalisme, car la réalité existe indépendamment de l'esprit — même les couleurs, selon une idée que Thagard reprend à Churchland, ont une existence objective, comme propriétés de réflexion de la lumière —, mais réalisme constructif, car nos connaissances neuroscientifiques ne permettent pas un réalisme naïf. En effet, de la perception à la production de théorie, notre connaissance de la réalité est basée sur un grand nombre d'inférences — et l'inférence est encore ici décrite comme un processus de satisfaction de contraintes, sur le plan tant perceptuel que propositionnel. Les représentations de l'esprit sont encodées dans des popula-

---

6. Il peut aussi lire Poirier et Payette (2007), qui le convaincraient peut-être que ses doutes sont infondés.

tions de neurones distribuées à travers différentes modalités, incluant des composantes perceptuelles, linguistiques ou émotionnelles.

Tout en décrivant les processus cognitifs à l'œuvre, l'auteur donne une meilleure idée du type de conception de l'esprit qu'il défend et montre que les réseaux de neurones connexionnistes qui servaient d'heuristiques dans les premières versions de la théorie de la cohérence explicative ont fait du chemin. Par le passé, au moins trois grands problèmes avaient été notés en ce concerne ces réseaux : l'absence de structure compositionnelle des propositions-neurones du réseau, l'absence de plausibilité biologique et la dépendance à une formulation particulière du problème en termes de réseau de cohérence, en l'absence de règles d'encodage unifiées. Il ne semble pas y avoir de progrès en ce qui a trait au troisième de ces problèmes, mais cela n'est guère surprenant : c'est une triviale en intelligence artificielle de dire que la résolution correcte d'un problème dépend de la représentation de ce problème. Quant aux deux autres, de grands progrès ont été faits. Le problème de la compositionnalité est réglé par la méthode d'Eliasmith utilisant les représentations holographiques réduites (HRR) de Plate. Ces représentations peuvent encoder des propositions structurées du type « A cause B, et C cause D, etc. » dans un vecteur unique. On peut encoder de tels vecteurs dans une population de neurones et construire un réseau de neurones qui, confronté à « B », appliquera à ces représentations une opération équivalant à poser la question « Qu'est-ce qui cause B ? », laquelle produit comme réponse la solution « A ». De cette manière, un modèle de l'abduction a pu être développé (Thagard et Litt, 2008). Thagard présente aussi une solution au problème du réalisme biologique. Celui-ci est évident dans les réseaux tels qu'ECHO : connexions bidirectionnelles entre neurones, représentations locales, connexions à la fois excitatrices et inhibitrices. En plus de la rapide présentation qu'il fait dans ce chapitre, Thagard a développé (Thagard et Aubie, 2008) un modèle nommé NECO où une population d'un millier de neurones encode collectivement la totalité des valeurs d'acceptation des propositions par leurs patterns d'impulsions, et où une connexion du modèle initial se transforme en de multiples connexions entre les neurones de la population, éliminant la nécessité de connexions bidirectionnelles directes entre deux neurones, ou de connexions à la fois excitatrices et inhibitrices partant d'un neurone. On peut dire que, même si Thagard ne dispose pas d'une théorie complète de la cohérence neurale, les pièces essentielles sont là, et il ne reste qu'à les assembler.

Thagard applique sa méthodologie de manière peut-être trop impérialiste lorsqu'il cherche à lier inextricablement notre concept de causalité et nos représentations sensorimotrices. Il se fonde sur des expériences montrant que notre traitement de l'information causale est probablement lié en partie à nos représentations sensorimotrices de manipulation ou à des mécanismes sensoriels innés de détection de la causalité, tels que démontrés par la surprise de très jeunes enfants lorsqu'on leur montre des séquences causa-

lement « impossibles » (par exemple : un objet roulant cogne sur un autre objet roulant, mais le second ne bouge pas). Mais une théorie de la causalité comme manipulation peut être développée sans appel à ces bases sensorimotrices (voir Craver 2007). De plus, les résultats des expériences sur les enfants et les singes pourraient aisément se voir donner l'interprétation suivante : nous avons des détecteurs, probablement innés, de la conservation du momentum. Ils sont imparfaits, et peuvent sans doute être déjoués, mais on pourrait en dire autant de nos détecteurs de lumière. Thagard soutiendrait-il que toute théorie de la lumière est « inéliminablement » sensorimotrice ? Le comportement de la lumière dans la théorie de la relativité ou dans la mécanique quantique est particulièrement étrange — si des résidus de nos concepts naturels jouent encore un rôle dans nos calculs, ils y font sans doute plus de mal que de bien. On voit difficilement pourquoi on devrait soutenir que les représentations sensori-motrices sont un trait inéliminable de l'un de ces concepts et non de l'autre. On pourrait accommoder une notion de causalité qui ne soit pas sensorimotrice en acceptant des concepts amodaux, mais il n'est pas sûr que Thagard, qui fait grand cas de la multimodalité, leur laisse une place dans sa théorie des concepts (p. 76-81). Beaucoup de philosophes pensent que les données neuroscientifiques sont incompatibles avec l'existence de concepts amodaux ; ce n'est pas le cas, bien que la théorie multimodale ait effectivement été favorisée récemment (voir Mahon et Caramazza 2008 pour une critique).

En utilisant son réalisme constructif, Thagard rejette l'idée que la réalité (et/ou la science) est une construction sociale. La démonstration est peut-être convaincante si on accepte une définition extrême de cette position telle que défendue autrefois par les tenants du « programme fort », mais le réalisme constructif pourrait être cohérent avec une position un peu plus forte que la seule qu'il accepte, à savoir la trivialité que la science est une entreprise sociale. Il est possible à la fois de considérer que la science nous permet de connaître une réalité indépendante de l'esprit — les connaissances qu'elle nous apporte ne sont pas une illusion — *et* que la science contribue à créer, dans certains cas, sa propre réalité. Considérons la position nuancée de Cordelia Fine sur les différences entre les sexes dans les tâches cognitives, émotionnelles, et dans les choix de carrière. Les femmes, en tant que groupe, ont souvent de moins bonnes performances sur les tests d'aptitudes mathématiques, ce qui peut causer des choix de carrières différents. Ces différences sont ensuite rapportées par les scientifiques et diffusées dans la culture populaire, confirmant le stéréotype courant. Le problème, c'est que ce stéréotype a une efficacité causale. En effet, on sait (grâce notamment aux travaux du psychologue Claude Steele sur la « menace du stéréotype ») qu'on peut faire disparaître ces différences dans des tests d'aptitudes mathématiques en manipulant la manière dont les tests sont présentés (Fine 2010, p. 30-31). Si on prend un groupe tel que les hommes testés ont, en moyenne, reçu les mêmes notes en classe que les femmes testées, et si on leur dit qu'il

n'y a pas de différence entre les sexes pour ce test, les femmes réussirent en moyenne mieux que les hommes, confirmant que le stéréotype agissait hors du laboratoire (elles réussirent aussi mieux que les femmes à qui on a dit que le test servait à comprendre pourquoi certaines personnes étaient meilleures que d'autres en mathématiques, ce qui est une situation de « menace du stéréotype »). Voilà pourquoi Fine soutient que les scientifiques œuvrant dans ce domaine doivent être prudents lorsqu'ils rapportent leurs conclusions. Un tel lien est très indirect, mais il existe une boucle causale où la science joue un rôle, pas du tout mystérieux ou « postmoderne », dans la construction de la réalité qu'elle découvre.

Nous avons vu que, pour Thagard, les représentations de l'esprit sont multimodales et comprennent parfois une composante d'évaluation émotionnelle. Le cinquième chapitre développe une théorie des émotions. C'est celui où les deux faces des neurosciences, théoriques et empiriques, sont le plus intimement liées. Comment le cerveau peut-il ressentir les émotions? La théorie de la conscience émotionnelle de Thagard (EMOCON) combine à la fois la théorie de l'émotion comme sensation corporelle et celle de l'émotion comme évaluation cognitive, posant une influence bidirectionnelle du cerveau et du corps. L'évaluation émotionnelle d'une situation se fait par un processus de cohérence impliquant à la fois des zones sensorielles et cognitives, avec de multiples boucles de rétroactions entre la plupart des zones du modèle — par exemple entre le thalamus, qui joue un rôle dans la perception et même dans la production des réactions corporelles (par l'intermédiaire de l'hypothalamus), et le cortex préfrontal, qui produit une évaluation cognitive. Grâce à cette théorie des émotions Thagard complète sa démonstration de l'identité esprit-cerveau, en montrant que sa théorie permet d'expliquer des caractéristiques importantes de la conscience des émotions, et ancre la notion de but dans les neurosciences, ce qui sera important pour la théorie de la décision.

En effet, un but est une représentation neurale d'une situation possible qui comprend une composante émotionnelle. La décision, étudiée au chapitre VI, est un processus d'inférence au meilleur plan, qui sélectionne les plans satisfaisant le mieux nos buts — toujours par un processus de cohérence, cette fois entre les données du problème, les actions possibles et les buts fixés. Parmi les systèmes en jeu, le système dopaminergique de récompense joue un rôle important, entre autres comme un mécanisme d'apprentissage par renforcement. Thagard explique le processus de révision de buts et décrit les causes des mauvaises décisions selon la recherche courante. Il reconnaît que le libre arbitre est rendu impossible par l'identité esprit-cerveau et la présence de processus inconscients dans la prise de décision, mais considère néanmoins que cette vision de la décision nous laisse, pour reprendre une expression de Dennett, « les sortes de liberté qui valent la peine qu'on les veuille » : « Nous pouvons encore avoir beaucoup des attributs désirables du libre-arbitre, y compris le contrôle de soi, l'expression de

soi, l'individualité, la capacité à réagir aux raisons, la délibération rationnelle, la responsabilité, et une sorte importante d'autonomie discutée au chapitre VII » (p. 139). Malgré le défi que pose le rejet de la notion de libre-arbitre à celle de responsabilité, nos vies peuvent être morales, comme l'auteur le montrera au chapitre IX, et même être douées de sens, en poursuivant une certaine catégorie de buts, comme on le verra à l'instant.

Car les buts que se fixe le cerveau humain ne sont pas tous égaux : certains sont de simples désirs arbitraires, et d'autres dotent la vie de signification en répondant à des besoins vitaux. Les chapitres VII et VIII forment le cœur de l'ouvrage et abordent le thème qui lui donne son titre. Dans le chapitre VII, Thagard avance que le sens de la vie a plusieurs dimensions, dont l'amour (incluant l'amitié et la famille), le travail et le jeu (incluant les sports et les arts). Il refuse le nihilisme et promet de montrer comment la plupart des gens peuvent trouver un sens à leur vie. Il refuse également d'associer le sens et le bonheur : on peut avoir l'un sans l'autre, comme le montre l'exemple des drogués (à qui la drogue fournit un bonheur éphémère mais dénué de sens) et des jeunes parents (moins heureux que ceux qui n'ont plus à prendre soin de leurs enfants, mais dont l'expérience est riche de sens). Cependant, le fait que des gens trouvent important de poursuivre certains buts qui par ailleurs ne leur apportent pas nécessairement du bonheur n'est pas suffisant pour dire que leur vie est vraiment dotée de sens. Fournissant une analyse conceptuelle, Thagard dit qu'une vie est sensée si : 1) « vous avez des buts [...] ; 2) certains de vos buts ont été atteints à un certain degré ; 3) vous avez d'autres buts, pas encore atteints, mais que vous avez une chance raisonnable d'atteindre ; 4) vos buts sont cohérents les uns avec les autres ; et 5) vos buts ont une valeur objective » (p. 151). Ce sont là des conditions souples, qui peuvent être plus ou moins satisfaites. Le chapitre VII, dans lequel Thagard fait un travail surtout descriptif, se concentre surtout sur les genres de buts que les gens considèrent comme importants ; il décrit par exemple à quel point l'amour au sens large est important, et pour quelles raisons neuroscientifiques (entre autres l'implication du système dopaminergique chez les personnes amoureuses, la représentation neurale de la douleur lorsque nos sentiments sont heurtés, l'augmentation du taux d'ocytocine menant à une plus grande confiance entre pairs). Il traite aussi du travail et du jeu, mais plus superficiellement, du moins sur le plan neuroscientifique. Le chapitre VIII montre que ces domaines satisfont de plus des besoins psychologiques, au sens technique de ce qui nous permet de prospérer et dont la privation nous fait du tort. Ces besoins, suivant la théorie de l'auto-détermination de Deci et Ryan, sont le besoin d'entrer en relation (*relatedness*), le besoin de compétence, et le besoin d'autonomie ; on se passera ici des définitions, les noms étant suffisamment descriptifs. Ces besoins ont été postulés, puis identifiés et validés par de multiples tests empiriques (Deci et Ryan 2000). À l'appui de ces besoins, Thagard apporte des données psychologiques mais aussi neuroscientifiques. Par exemple, les données pour le

besoin d'entrer en relation sont en partie les mêmes que celles décrites au sujet de l'amour au chapitre précédent, et le besoin de compétence trouve une base neurale dans l'existence d'un système de récompense dépendant de l'effort: un plus grand effort mène à une plus grande récompense sous la forme d'une activité accrue du système dopaminergique (p. 173).

Puisque le meilleur moyen de satisfaire ces besoins vitaux est de poursuivre des buts dans les domaines de l'amour, du travail et du jeu, les buts dans ces trois domaines ont une valeur objective — ce qui satisfait la condition (5) ci-dessus pour une vie sensée. Thagard, cependant, considère cette justification comme spéculative, car les données en faveur des domaines du sens sont plus fortes que celles en faveur des besoins. Et il semble ambivalent sur l'innéisme des besoins: d'une part il écrit que leur caractère inné n'est pas suffisamment prouvé et qu'il suffit qu'ils soient des besoins au sens technique (p. 171); d'autre part il écrit qu'ils sont indépendants de la culture et de l'expérience (p. 172) et les qualifie de « biologiques » (p. 182). Une interprétation charitable de ces hésitations serait que les trois besoins sont hautement canalisés par les contraintes environnementales communes et la biologie humaine.

Des éléments dans la liste des besoins vitaux et dans celle des domaines du sens semblent manquer. Étrangement, malgré le soin apporté dans le livre à fournir une vision du monde naturaliste compréhensive et l'importance évidente que l'auteur donne à la connaissance, il faut attendre la fin du livre pour que celui-ci accorde, dans un court passage de quatre lignes, le statut de besoin objectif à la vérité et à l'explication (p. 210). Une autre omission est également assez grave: bien que l'auteur ne fasse pas mystère de ses sympathies politiques, l'appartenance politique, culturelle ou nationale ne joue aucun rôle dans sa théorie du sens. Le traitement de l'art comme jeu est peut-être le signe d'un biais méthodologique individualiste dans le naturalisme neural; si l'art est important, c'est au moins en partie en tant que vecteur de valeurs et d'appartenance (notre réaction à un hymne patriotique ou aux *Sorcières de Salem* n'est pas basée uniquement sur leurs qualités formelles). Thagard fondera son approche des valeurs morales sur la théorie des besoins, mais Deci et Ryan reconnaissent que la forme que prendront ces besoins dans chaque culture peut varier, et que le rôle des appartenances identitaires, nationales, politiques ou autre est d'autant plus positif qu'elles sont sincèrement intériorisées (Deci et Ryan 2000, Ryan et Deci 2002). L'appartenance nationale répond également à un, et peut-être même deux, des trois besoins vitaux: le besoin d'entrer en relation, sous l'espèce du besoin d'appartenance, et (peut-être) le besoin d'autonomie. Kymlicka (1995) a en effet soutenu qu'une culture nationale met les individus en présence d'un contexte de choix (un ensemble de valeurs et de manières de vivre leur vie) qui serait un élément essentiel à l'autonomie. Cela semble en harmonie avec ce que disent Deci et Ryan de la variation de l'implémentation des besoins; inversement, une

culture nationale en déclin mettra l'individu à risque de devoir adopter une identité culturelle pour des raisons extrinsèques<sup>7</sup>.

Par ailleurs, bien que Thagard semble au moins démontrer la fausseté de l'existentialisme extrême de Sartre exprimé dans la citation au début de ce texte, il reste que l'essentialisme et la biologie ne font pas bon ménage. En effet, même si on accepte que les besoins vitaux soient innés (ce que Thagard lui-même hésite à faire), les domaines du sens sont d'ores et déjà modulables par des moyens pharmacologiques (par exemple l'amour [p. 155], ou le jeu [p. 162]). De plus, le neuroscientifique Read Montague a présenté un mécanisme par lequel les idées abstraites deviennent elles-mêmes des signaux de récompense évalués par le système dopaminergique, qui nous permettent de contourner nos besoins vitaux, pour le bien — un individu faisant la grève de la faim pour la libération des prisonniers politiques d'une dictature —, ou pour le mal — des membres d'une secte faisant un pacte de suicide pour « avancer au niveau suivant » (Montague, 2006).

La notion de besoin vital fait non seulement que le sens est possible dans une vision du monde naturaliste, mais que l'éthique peut recevoir un fondement. Le chapitre IX traite des décisions morales. Ce qui fonde le jugement moral, ce n'est pas une grammaire morale innée ou l'équilibre réflexif, mais deux ingrédients : une composante cognitive et normative, la considération des besoins vitaux des autres et des conséquences de nos actions sur ceux-ci, et une composante émotive et motivationnelle. Cette dernière composante fait une place aux intuitions, conçues comme des réactions émotives telles que décrites au chapitre V (bien que Thagard reconnaisse qu'elles n'ont pas la robustesse des perceptions, et ne peuvent donc pas être traitées comme des données pour une théorie morale). Mais le mécanisme émotif le plus important, donnant à la morale une force motivationnelle, est le système des neurones miroirs, qui nous permettrait de ressentir de l'empathie. Thagard adopte un conséquentialisme modéré (évaluant les règles et non seulement les actes individuels) ayant comme source les besoins fondamentaux. Les actions et les règles morales sont choisies au moyen du processus cohérentiste d'inférence au meilleur plan présenté au chapitre VI. La question de la responsabilité est abordée d'un point de vue conséquentialiste : malgré le rejet du libre arbitre, il est permis de tenir un individu pour responsable de ses actes, et même de le punir d'un crime, si cela évite que des actes semblables soient commis dans l'avenir. Thagard accepte le verdict de Hume selon lequel on ne peut dériver un devoir d'un fait, mais on peut toutefois décider dans quelle mesure une théorie morale est plus ou moins cohérente avec les faits, et c'est tout ce qui est nécessaire.

---

7. Un reportage de Radio-Canada sur la communauté francophone de Saskatchewan souligne à quel point le contexte de choix des individus de cette communauté s'est appauvri durant la période récente. « Le déclin de la culture francophone en Saskatchewan », 20 mars 2008, en ligne : [http://www.radio-canada.ca/actualite/v2/tj22h/archive113\\_200803.shtml#](http://www.radio-canada.ca/actualite/v2/tj22h/archive113_200803.shtml#).

Il faut malheureusement mettre un bémol à l'enthousiasme de Thagard pour la théorie des neurones miroirs, censés jouer un rôle — important dans sa théorie éthique —, dans la compréhension de l'action et dans l'empathie. Cette théorie est sous attaque. De nombreuses données rendent douteuse la théorie selon laquelle les neurones miroirs forment le mécanisme neural produisant la compréhension de l'action. Entre autres données rendant moins probable que les neurones miroirs remplissent cette fonction, il existe des résultats de dissociation double entre production et compréhension de l'action, ainsi qu'une démonstration de la plasticité des neurones dits miroirs (Hickok, 2009). Mais c'est leur rôle dans l'empathie qui est le principal problème pour la théorie de Thagard. Or ce rôle est aussi discutable. Un résultat dont Thagard fait grand cas est la réponse « miroir » à la douleur d'autrui : les mêmes neurones s'activent lorsqu'un individu voit de la douleur et la ressent, ce qui permettrait de donner une force « de motivation » à la morale. Malheureusement de telles études corrélatives chez les personnes saines sont limitées. Une étude neuropsychologique fascinante implique des patients souffrant d'insensibilité congénitale à la douleur. Ces patients n'ont jamais ressenti de douleur de leur vie. Pourtant, la réponse de leurs neurones supposément « miroirs » est identique à celle des individus non atteints, ce qui laisse supposer que les circuits en question sont impliqués dans une autre fonction, comme la reconnaissance des situations à éviter et la préparation motrice à l'évitement ; le fait que chez les médecins cette zone ne soit pas activée lorsqu'ils sont piqués avec une aiguille — une situation peu menaçante pour eux — conforte cette hypothèse. Quant aux résultats reliant l'habileté à l'empathie et la réponse des neurones miroirs, ils sont peu convaincants : certaines études en trouvent, mais d'autres non, en particulier les études neuropsychologiques. Ces données et beaucoup d'autres militent contre un rôle des neurones miroirs dans l'empathie sont recensées par Decety (2010). Si ces critiques s'avéraient fondées, le lien si simple et si utile que les philosophes pensent pouvoir faire entre les neurones miroirs et la moralité s'effondrerait. Et pour ce qui est de Thagard, il devrait se retrancher dans la position moins « neurale » selon laquelle l'empathie motive le comportement moral (ce qu'elle fait sans aucun doute), peut-être à travers le mécanisme de contagion émotive décrit p. 192.

Un autre problème de la théorie de Thagard, qui pourrait être plus sévère, c'est que dans son traitement des intuitions il tient pour acquis la perméabilité des jugements moraux aux évaluations cognitives et au raisonnement sur les besoins. C'est justifié du point de vue de sa théorie des émotions, mais Haidt (2001) a soutenu, sur la base de données troublantes, que le rôle des raisonnements était de justifier aux yeux d'autrui les jugements auxquels nos réactions viscérales, nos intuitions, nous poussent. Penser que le raisonnement moral joue un rôle dans l'action et le jugement est, selon sa métaphore, penser que la queue du chien fait bouger le chien, alors que c'est évidemment l'inverse qui se produit. Dans cette optique, l'entreprise

de Thagard de construire une théorie éthique objective est totalement vaine, puisqu'elle n'aura aucune influence sur le comportement moral. De plus cela supprime deux formes de liberté que Thagard considère encore valides : la capacité à réagir aux raisons et la délibération. La théorie de Haidt a été remise en question, avec des arguments empiriques, par Fine (2006) et Kenneth et Fine (2009), qui montrent qu'à un certain degré le contrôle conscient peut supprimer l'influence de nos intuitions et préjugés, mais cela ajoute une complexité à l'application d'une théorie éthique qui manque au chapitre de Thagard. Ce débat aurait aussi pu enrichir son traitement de la notion de responsabilité.

Les différentes conclusions des chapitres précédents sont remises en contexte dans le chapitre X. On y trouvera aussi une procédure naturaliste pour l'adoption de normes (où les normes sont des moyens pour les buts de chaque domaine, par exemple la vérité : « l'épistémologie normative est la technologie de la recherche de vérité », dirait Quine). Thagard applique cette méthode à la détermination du meilleur type de gouvernement et à la production des changements sociaux. Par exemple, il détermine que le meilleur régime politique est une démocratie libérale dotée d'un large filet social sur la base du critère objectif donné par l'indice de développement humain. Il aborde ensuite deux problèmes plus traditionnels : l'existence des entités mathématiques (il se prononce en faveur du fictionnalisme) et celle de quelque chose plutôt que rien (de manière cohérente avec son emphase sur l'explication, plutôt que de considérer la question de la cause du Big Bang comme dénuée de sens, il fait appel à la théorie des cordes).

On a pu observer dans les pages qui précèdent à quel point l'ouvrage de Thagard est systématique : il aborde les grands domaines de la philosophie, les chapitres se soutiennent les uns les autres, et il utilise une méthode unifiée. Les thèses et arguments les plus importants ont été évalués. Reste à voir si le livre est réellement un « argument étendu » pour le naturalisme neural, comme il le prétend.

Avant de le déterminer, certains périls de l'usage des neurosciences doivent être mentionnés. Supposons qu'il existe un biais cognitif, la « malédiction de la connaissance », selon laquelle les gens sont plus susceptibles de penser que les autres connaissent un fait qu'ils savent déjà eux-mêmes qu'un fait qu'ils ne connaissent pas. De ces deux explications, laquelle préférez-vous ?

- 1) « Les chercheurs pensent que cette « malédiction » existe parce que les gens font plus d'erreurs quand ils ont à juger la connaissance des autres. Les gens sont bien meilleurs à juger de ce qu'ils savent eux-mêmes. »
- 2) « L'imagerie cérébrale montre que cette “malédiction” existe à cause des circuits du lobe frontal dont on sait qu'ils sont impliqués dans la connaissance de soi. Les gens font plus d'erreurs lorsqu'ils ont à juger la connaissance des autres. Les gens sont bien meilleurs à juger de ce qu'ils savent eux-mêmes. »

Weinsberg *et al.* (2008) ont démontré que les gens préfèrent la seconde explication à la première, alors que les deux sont tout aussi circulaires. De manière intéressante, cet effet trompeur de l'ajout d'information neuroscientifique n'est pas présent lorsque l'explication est une bonne explication. Autrement dit, nous devons nous méfier de « l'aspect séducteur » des neurosciences.

Racine *et al.* (2005) ont montré sur la base d'une analyse d'articles de vulgarisation que l'imagerie cérébrale fonctionnelle a également l'effet de faire considérer des phénomènes psychologiques comme plus réels, voire innés, alors qu'elle ne devrait pas avoir cet effet. Que la nourriture grasse apporte du plaisir n'est pas, comme tendait à le faire croire un des articles à l'étude, plus vrai sous prétexte que nous avons un IRM-f qui détecte une différence d'activation dans des zones du cerveau selon le taux de gras consommé<sup>8</sup>. Et une différence apprise (par exemple, entre joueurs d'échecs novices ou expérimentés) sera reflétée par cette technique autant qu'une réponse innée.

Un aspect de l'évaluation du naturalisme neural doit porter sur l'intégration des neurosciences au projet. Thagard ne peut se contenter de simplement énumérer les zones impliquées dans une tâche s'il veut défendre la pertinence du naturalisme neural comme *méthodologie*: il ne peut même pas se contenter de montrer qu'un fait cognitif est *expliqué* par les neurosciences (ce qui éviterait tout de même l'écueil mentionné par Weinsberg *et al.*). Il doit montrer que ses thèses sont *justifiées* par elles, bien que parfois une explication puisse être une justification lorsque, par exemple, en expliquant un comportement, on montre pourquoi les causes neurales donnent de bonnes *raisons* d'effectuer ce comportement. Et, bien qu'il soit anti-réductionniste, Thagard ne peut non plus se contenter, s'il veut démontrer l'importance du naturalisme *neural*, de fournir des justifications à ses thèses philosophiques qui proviennent exclusivement des sciences cognitives ou de la psychologie expérimentale: au moins une partie de la justification doit être neuroscientifique. Peut-on, à partir du traitement des quatre questions traditionnelles ciblées par Thagard, tirer une conclusion sur la pertinence du naturalisme neural?

*Qu'est-ce que la réalité?* Il n'est pas certain que le réalisme constructif proposé par Thagard nous apporte de grandes révélations sur la réalité, mais il a au moins le mérite de poser le problème en des termes différents (étant donné ce qu'on sait du cerveau, un réalisme naïf n'est pas possible; par quoi le remplacera-t-on?) En insistant sur le caractère inférentiel de la perception, Thagard offre également une solution de remplacement à la vision simpliste de l'encapsulation de la perception qui a été le legs de la théorie de la modularité de Fodor à l'épistémologie naturalisée<sup>9</sup>.

8. Mais la zone impliquée pourrait nous donner une indication du degré de plasticité de ce trait, ou sa force de motivation.

9. L'encapsulation de la perception est réfutée par la variabilité culturelle de la susceptibilité aux illusions.

*Comment connaît-on la réalité?* L'épistémologie cohérentiste de Thagard — sur le plan à la fois individuel et social — reste incomplète, mais montre des signes encourageants d'amélioration d'un livre à l'autre. Cependant, un préjugé envers les concepts abstraits, moins étudiés par les neurosciences, semble une des racines de la portée indue qu'il accorde à l'aspect sensorimoteur de la notion de causalité. Et on a vu qu'il sous-estimait la capacité de la science de construire, d'une certaine façon, sa propre réalité, en interagissant avec ses objets; ici, la psychologie sociale semble apporter une contribution qui risque d'échapper à une approche exclusivement neurale, du moins pour le moment.

*Quel est le sens de la vie?* La solution au problème du sens intègre bien neurosciences, psychologie et philosophie: ici, et particulièrement dans le cas de l'amour, l'explication neuroscientifique est réellement une justification: les relations humaines impliquent les structures de récompense, ce qui montre leur aspect positif, et, lorsque ces relations sont brisées, les structures traitant la douleur sont activées, ce qui montre que leur absence nous cause un tort. C'est aussi le cas pour le besoin de compétence, où l'existence de circuits de récompense dépendants de l'effort justifie l'importance de la compétence. L'écueil mentionné par Racine *et al.* est souvent bien évité — lorsque l'auteur examine le rôle du système dopaminergique dans l'amour, par exemple, nous savons déjà que ce système est une source de motivation grâce aux chapitres précédents. Il existe un risque de biais individualiste inhérent au naturalisme neural, on l'a vu, mais toute méthode a des biais, qui peuvent être contrés (Wimsatt 2007), par exemple par l'usage du nouveau champ des neurosciences culturelles (Losin *et al.* 2010). Un certain nombre de descriptions des processus neuraux impliqués dans tel ou tel phénomène sont partielles et ne rencontrent pas toujours le critère de justification, il est vrai; cela pourrait soulever la suspicion de la présence d'explications circulaires. Mais une interprétation charitable serait que Thagard fournit de cette façon les pièces des mécanismes (Craver 2007) pour les explications et justifications à venir.

*Qu'est-ce qui fait qu'une action est bonne ou mauvaise?* Finalement, pour ce qui est de l'action morale, Thagard accorde une confiance trop grande à la théorie des neurones miroirs. C'est inévitable dans une philosophie naturaliste: les sciences, contrairement à la foi — comme Thagard ne manquerait sans doute pas de répondre — révisent leurs croyances sur la base des nouvelles données.

Il faut donc reconnaître que le naturalisme neural est au moins partiellement soutenu par cet ouvrage. Certains commentateurs ont reproché à Thagard sa distance face à la psychologie évolutionniste: il pense que l'esprit est largement constitué de mécanismes généralistes (p. 212-213), et cela aurait rendu son ouvrage moins sombre qu'il devrait l'être. Mais cette distance lui permet d'éviter l'inférence indue des observations neurales à l'innéisme, et d'offrir une philosophie naturaliste systématique avec somme

toute assez peu de spéculation, restant proche des données empiriques. Et c'est peut-être une raison de saluer le naturalisme neural, et d'espérer que cette méthodologie aura davantage à livrer dans les années futures.

## Bibliographie

- Brandon, R.N. *Evolutionary Modules: Conceptual Analyses and Empirical Hypotheses*. In W. Callebaut et D. Rasskin-Gutman (eds), *Modularity*, Cambridge, Mass., The MIT Press, 2005, p. 51-60.
- Craver, Carl F. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, New York, Oxford University Press, 2007.
- Decety, J. « To What Extent is the Experience of Empathy Mediated by Shared Neural Circuits? », *Emotion Review*, Vol. 2, N° 3, 2010, p. 204-207.
- Deci, E.L. et R.M. Ryan. « The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior », *Psychological Inquiry*, N° 11, 2000, p. 227-268.
- Fine C. « Is the Emotional Dog Wagging its Rational Tail, or Chasing it? Philosophical Explorations », N° 9, 2006, p. 83-98.
- . *Delusions of Gender: How our Minds, Society and Neurosexism Create Difference*, New York, W. W. Norton, 2010.
- Haidt J. « The Emotional Dog and its Rational Tail: a Social Intuitionist Approach to Moral Judgment », *Psychological Review* 108, 2001, p. 814-834.
- Hickok G. « Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans », *Journal of Cognitive Neuroscience*, Vol. 21 N° 7, 2009, 1229-43.
- Kennett, J. et C. Fine. « Will the Real Moral Judgment Please Stand Up? The Implications of Social Intuitionist Models of Cognition for Meta-ethics and Moral Psychology, Ethical Theory and Moral Practice », Vol. 12, N° 1, 2009, p. 77-96.
- Kymlicka, W. *Multicultural Citizenship: a Liberal Theory of Minority Rights*, Oxford University Press, 1995.
- Ladyman, J. et D. Ross. *Every Thing Must Go. Metaphysics Naturalized*, Oxford University Press, 2007.
- Lewontin, R.C. « The Analysis of Variance and the Analysis of Causes », *American Journal of Human Genetics*, N° 26, 1974, p. 400-411.
- Losin, E.A., M. Dapretto et M. Iacoboni. « Culture and Neuroscience: Additive or Synergistic? », *Social, Cognitive and Affective Neuroscience*, 5, 2010, p. 148-158.
- Mahon, B. Z. et A. Caramazza. « A Critical Look at the Embodied Cognition Hypothesis and a New Proposal for Grounding Conceptual Content », *Journal of Physiology*, Paris, 102, 2008, 59-70.
- Montague, R. *Why Choose This Book?* New York, NY., Penguin Group, 2006.
- Poirier, P. et N. Payette. « Les gardiens du bon usage: Étude critique de “Philosophical Foundations of Neuroscience”, de P. M. S Hacker et M. R. Bennett ». *Philosophiques*, Vol. 34, N° 1, 2007, p. 183-200.
- Racine, E., O. Bar-Ilan et J. Illes. « fMRI in the Public Eye », *Nature Reviews Neuroscience*, February, Vol. 6, N° 2, 2005, p. 159-164.
- Ryan, R. M. et E. L. Deci. *On Assimilating Identities to the Self: a Self-Determination Theory Perspective on Internalization and Integrity within Cultures*. In *Handbook of Self and Identity*, M. R. Leary et J. P. Tangney (eds), New York, Guilford, 2002, p. 255-273).

- Ryan, R.M., V. Huta et E.L. Deci. « Living Well: A Self-Determination Theory Perspective on Eudaimonia », *Journal of Happiness Studies*, N° 9, 2008, p. 139-170.
- Thagard P. et A. Litt. *Models of Scientific Explanation*. In *The Cambridge Handbook of Computational Cognitive Modeling*, R. Sun (ed.), Cambridge, England, Cambridge University Press, 2008.
- Thagard, P. *Conceptual Revolutions*, Princeton, Princeton University Press, 1992.
- . *Coherence in Thought and Action*, MIT Press, 2000.
- Thagard, P. et B. Aubie. « Emotional Consciousness: A Neural Model of How Cognitive Appraisal and Somatic Perception Interact to Produce Qualitative Experience ». *Consciousness and Cognition*, N° 17, 2008, p. 811-834.
- Weisberg, D.S., F.C. Keil, J. Goodstein, E. Rawson, et J.R. Gray. « The Seductive Allure of Neuroscience Explanations », *Journal of Cognitive Neuroscience*, Vol. 20, N° 3, 2008, p. 470-477.
- Wimsatt, W.C. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*, Harvard University Press, Cambridge, MA, 2007.