

Université de Montréal

L'intuitionnisme de Huemer : une solution au scepticisme moral?

Par
Karl-Antoine Pelchat

Département de philosophie, Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de Maîtrise ès arts (M.A.) en philosophie,
option recherche

Août 2022

© Karl-Antoine Pelchat, 2022

Ce mémoire intitulé

L'intuitionnisme de Huemer : une solution au scepticisme moral?

Présenté par

Karl-Antoine Pelchat

A été évalué par un jury composé des personnes suivantes

Aude Bandini

Présidente-rapporteuse

Christine Tappolet

Directrice de recherche

Denise Celentano

Membre du jury

Résumé

Situé à l'intersection de grands mouvements en philosophie morale, comme la renaissance du réalisme non naturaliste et l'affirmation d'un scepticisme épistémologique moral, ce mémoire se propose d'interroger la fiabilité des intuitions morales à la lumière des récentes découvertes effectuées en neuroscience, psychologie morale et théorie de l'évolution. Ne pouvant recouvrir l'intégralité des conceptions de l'intuition et de la justification morales, ce mémoire se penchera prioritairement sur l'intuitionnisme éthique prôné par Michael Huemer. Solution alléguée au scepticisme épistémologique moral, l'intuitionnisme de Huemer garantit une justification *prima facie* aux énoncés moraux, une propriété conférée par le principe du conservatisme phénoménal. Profitant aux intuitions morales de tous niveaux de généralité, cette justification *prima facie* peut toutefois être défaite, comme le supposent les arguments étiologiques proximaux et distaux. Au cours de ce mémoire, je défendrai l'idée qu'une majorité d'intuitions morales, telles que conçues par Huemer, sont effectivement défaites par la voie étiologique. D'une part, je tenterai de montrer que l'hypothèse du double système, bénéficiant désormais d'un vaste soutien empirique, n'est pas compatible avec la conception huémérienne des intuitions morales. En l'occurrence, la « trolleyologie » de Joshua Greene, le « Modèle intuitionniste social » de Jonathan Haidt ainsi que les procédés heuristiques seront mobilisés à des fins argumentaires. D'autre part, je tenterai de montrer que la critique généalogique évolutionnaire de Street jette le doute sur une grande partie des intuitions morales, principalement les intuitions morales substantives (non formelles). Comme le soutient Street, l'incapacité du réaliste à rendre compte de la relation entre les vérités évaluatives postulées et les pressions évolutives constitue une raison convaincante d'abandonner le réalisme moral. Ayant présenté deux types d'arguments étiologiques, l'un psychologique et l'autre évolutionnaire, je conclus que seules les intuitions morales formelles et métadiscursives sont réellement susceptibles d'éviter la charge du sceptique épistémologique.

Mots clés

Épistémologie morale, justification des énoncés moraux, intuitionnisme éthique, scepticisme épistémologique moral, conservatisme phénoménal, intuitions morales, arguments étiologiques, défaiseurs discréditants, hypothèse du double système, critique généalogique évolutionnaire

Abstract

Located at the intersection of a great many philosophical movements, such as the rebirth of non-naturalist realism and the affirmation of moral epistemological skepticism, this master's thesis interrogates the reliability of moral intuitions in light of recent advancements in neuroscience, moral psychology and evolutionary theory. While there are various ways to conceive of moral intuition and justification, this thesis will prioritize a version of ethical intuitionism defended by Michael Huemer. Put forward as a solution to moral epistemological skepticism, Huemer's intuitionism guarantees *prima facie* justification to moral propositions, a property conferred by the principle of phenomenal conservatism. While intuitions of every level of generality can benefit from such *prima facie* justification, this very type of justification is susceptible to defeat, as suggested by proximal and distal debunking arguments. Throughout this master's thesis, I will defend the claim that a majority of moral intuitions, as conceived by Huemer, are effectively defeated by means of debunking arguments. On the one hand, I will try to demonstrate that the dual-process hypothesis, which is widely supported, is not compatible with Huemer's conception of moral intuitions. In order to substantiate this view, I will draw from Joshua Greene's "trolleyology", Jonathan Haidt's "Social Intuitionist Model" and recent work on moral heuristics. On the other hand, I will attempt to show that Street's evolutionary debunking argument throws significant doubt on a large portion of moral intuitions, mostly the ones which are substantial (non-formal). As Street purports, the realist's incapacity to explain the relation between posited evaluative truths and evolutionary pressures constitutes a strong reason to reject moral realism. Having presented two types of debunking arguments, one psychological and the other evolutionary, I conclude that only formal and metadiscursive moral intuitions are really susceptible to avoid the charge of the epistemological sceptic.

Keywords

Moral epistemology, justification of moral propositions, ethical intuitionism, moral epistemological skepticism, phenomenal conservatism, moral intuitions, debunking arguments, undercutting defeaters, dual-process hypothesis, evolutionary debunking

Table des matières

| | |
|--|-----|
| Résumé | iii |
| Abstract | iv |
| Remerciements | vii |
| Introduction | 1 |
| Chapitre 1 : L'intuitionnisme éthique de Huemer..... | 3 |
| 1.1 Pourquoi l'intuitionnisme de Huemer plutôt qu'un autre? | 3 |
| 1.2 Le positionnement métaéthique de Huemer | 4 |
| 1.3 La thèse du conservatisme phénoménal | 5 |
| 1.3.1 La formulation du principe | 5 |
| 1.3.2 Internalisme, externalisme et théories rivales | 6 |
| 1.3.3 L'argument de l'auto-réfutation..... | 10 |
| 1.4 Les intuitions selon Huemer | 13 |
| 1.4.1 Ce qu'elles sont..... | 13 |
| 1.4.2 Ce qu'elles ne sont pas..... | 14 |
| 1.4.3 Les quatre types d'intuitions..... | 15 |
| 1.4.4 Les intuitions intellectuelles morales..... | 15 |
| 1.4.5 Les types d'intuitions morales | 16 |
| 1.5 Les défauteurs envisagés par Huemer | 17 |
| 1.6 Les arguments étiologiques | 21 |
| 1.6.1 Structure..... | 21 |
| 1.6.2 Les types de défauteurs..... | 22 |
| 1.6.3 Une typologie des arguments étiologiques | 22 |
| 1.7 Présentation de ma critique..... | 23 |
| Chapitre 2 : L'hypothèse du double processus et ses implications..... | 26 |
| 2.1 « Révolution affective » et philosophie morale..... | 27 |
| 2.1.1 La « trolleyologie » de Joshua Greene..... | 27 |
| 2.1.2 Le « Modèle intuitionniste social » de Jonathan Haidt..... | 38 |
| 2.2 « Rationalité limitée » et règles heuristiques..... | 46 |
| 2.2.1 Heuristiques et moralité : un projet au stade embryonnaire | 46 |
| 2.3 Intuitionnisme huémérien et modèle du double processus : une épreuve | 51 |
| Chapitre 3 : La critique généalogique évolutionnaire de Street | 54 |
| 3.1 Le dilemme darwinien de Street : contre les théories réalistes de la valeur..... | 55 |

| | |
|--|----|
| 3.1.1 Branche I : négation d'une relation..... | 57 |
| 3.1.2 Branche II : affirmation d'une relation | 59 |
| 3.1.3 Objections envisagées par Street | 61 |
| 3.1.4 Schéma de l'argument..... | 63 |
| 3.2 L'explication par un « troisième facteur » : une solution tangible? | 64 |
| 3.2.1 Enoch : la survie comme « troisième facteur »..... | 64 |
| 3.2.2 Wielenberg : certaines facultés cognitives comme « troisième facteur »..... | 66 |
| 3.2.3 Skarsaune : le plaisir et la douleur comme « troisième facteur » | 68 |
| 3.2.4 Discussion sur les explications par un « troisième facteur » | 70 |
| 3.3 Intuitionnisme huémérien et critique généalogique évolutionnaire : quelles implications? | 76 |
| Conclusion..... | 79 |
| Bibliographie..... | 83 |

Remerciements

J'aimerais, dans un premier temps, remercier ma directrice, Christine Tappolet, pour ses précieux conseils et commentaires tout au long de mon parcours à la maîtrise. En plus de m'avoir fait découvrir le très excellent Michael Huemer, elle m'a aidé focaliser ma recherche et mon propos, de manière à rédiger un mémoire philosophiquement pertinent et conceptuellement précis. Dans un second temps, je souhaite remercier mes proches et amis, lesquels m'ont appuyé tout au long de mes études, et de ma rédaction. Plus particulièrement, je voudrais remercier mes chers parents, dont la patience et le soutien furent essentiels à la réalisation de mes desseins académiques.

Introduction

Des nombreuses problématiques traitées en épistémologie morale, la justification est probablement celle qui accable le plus les épistémologues. Faisant partie intégrante de la connaissance, comprise au sens de « croyance, vraie, justifiée et non défaite », la justification est un desideratum pour toute théorie métaéthique qui affirme la possibilité d'une connaissance morale. Pour répondre à cette exigence de justification, plusieurs approches comme le fondationnalisme, le cohérentisme ou le contextualisme ont été proposées. Or, celles-ci sont exposées à des problèmes rédhibitoires comme l'argument de la régression, de la circularité ou des désaccords intersubjectifs.

S'inscrivant dans le renouveau du réalisme non naturaliste, Michael Huemer prétend, par son intuitionnisme rationaliste, dénouer l'aporie sceptique et ainsi, procurer un socle sûr à la morale. Reposant sur le principe du conservatisme phénoménal, l'intuitionnisme huémérien conçoit la justification d'une manière non inférentielle et internaliste. Selon l'auteur, le recours au conservatisme phénoménal serait nécessaire, car toute réfutation de ce même principe serait auto-réfutante. Cet argument fait toute la force de sa proposition épistémologique.

Concevant les intuitions comme des attitudes propositionnelles pré-doxastiques, Huemer distingue quatre genres d'intuitions : perceptives, mnémoniques, introspectives et intellectuelles. Sous-genre des intuitions intellectuelles, les intuitions morales correspondent à des états cognitifs affichant un contenu propositionnel. Étant d'un niveau de généralité variable, les intuitions morales sont elles-mêmes divisibles en trois sous-groupes : les intuitions concrètes, intermédiaires et abstraites (non formelles et formelles). En l'occurrence, on pourrait suggérer que la généralité des intuitions permet de prédire leur degré de défaisabilité, ou de susceptibilité à des défaiteurs épistémiques. Similairement, il serait sensé de croire que les intuitions formelles sont plus résistantes que les intuitions substantives auxdits défaiteurs.

Au cours de ce mémoire, je me propose justement de soumettre les intuitions morales à la critique. Plus précisément, il s'agira de dresser deux arguments étiologiques (*debunking*

arguments), l'un psychologique et l'autre évolutionnaire, afin de montrer que les processus ayant permis la formation de ces intuitions ne sont pas « rationnels¹ ».

Dans un premier temps, je formulerai un argument étiologique axé sur l'hypothèse du double processus. Interrogeant les causes proximales de nos intuitions morales, cet argument mobilisera la « trolleyologie » des Joshua Greene, le « Modèle intuitionniste social » de Jonathan Haidt ainsi que de récents travaux portant sur les règles heuristiques. Les résultats abordés dans ce chapitre, qu'ils soient en neuroscience ou en psychologie sociale, suggèrent fortement que les émotions, et non la raison désincarnée, seraient à l'origine d'une grande partie de nos jugements moraux. Ceci suffirait pour défaire la justification *prima facie* dont bénéficient nos intuitions morales, en vertu du conservatisme phénoménal.

Dans un second temps, je présenterai la critique généalogique évolutionnaire de Sharon Street. Interrogeant les causes distales de nos intuitions morales, cet argument jette un doute considérable sur la fiabilité desdites intuitions. Adoptant la forme d'un dilemme, l'argument de Street suggère que le réaliste ne peut expliquer, d'une manière plausible, la relation entre les vérités évaluatives postulées et les pressions darwiniennes observées. Si le réaliste devait nier l'existence d'une relation entre vérités évaluatives et pressions évolutives, il serait réduit au scepticisme normatif. À l'inverse, si le réaliste devait affirmer l'existence d'une relation entre vérités évaluatives et pressions évolutives, il serait contraint d'adopter une hypothèse scientifiquement improbable. Ne pouvant accepter ni l'une, ni l'autre des branches du dilemme, le réaliste serait forcé d'admettre que sa position est intenable. Comme le veut la bonne pratique philosophique, cet argument, aussi convaincant soit-il, sera soumis à une objection fort répandue, l'explication par un « troisième facteur ». Une fois cette objection abordée, il s'agira d'explicitier les implications du dilemme darwinien pour les intuitions morales, au sens où les conçoit Huemer. Je montrerai, encore une fois, que les intuitions substantives sont particulièrement susceptibles à la critique de Street.

Au terme de mon analyse, je conclurai que l'épistémologie intuitionniste de Huemer, et son conservatisme phénoménal, ne peuvent rien pour une majorité d'intuitions morales, lesquelles sont laissées vulnérables au scepticisme épistémologique. Étant soumises aux arguments étiologiques

¹ Par processus « rationnel », j'entends tout mécanisme cognitif qui permettrait la perception de vérités morales *aprioriques*.

proximaux et distaux, les intuitions morales substantives (non formelles) perdent la justification *prima facie* dont elles pouvaient bénéficier, les rendant ainsi caduques. Un tel constat signifie, non pas qu'une majorité d'intuitions morales sont fausses, mais qu'elles ne sont pas justifiées. On se gardera néanmoins d'étendre ce constat aux intuitions formelles et métadiscursives, lesquelles résisteraient mieux aux arguments étiologiques présentés dans ce mémoire. Ceci consolera malheureusement fort peu de réalistes, sachant que la morale résultante est dépourvue de tout contenu normatif.

Chapitre 1 : L'intuitionnisme éthique de Huemer

Au cours de ce chapitre, nous tâcherons d'abord d'expliquer le positionnement de Michael Huemer sur le plan métaéthique. Une attention particulière sera portée à son épistémologie et, plus précisément, au conservatisme phénoménal. Ensuite, il s'agira de donner une définition des intuitions morales, tout en respectant le cadre perceptualiste préféré par Huemer. Au même moment, il sera possible de dresser une typologie des intuitions morales. Finalement, il s'agira d'identifier une gamme de défaiteurs pouvant jeter un doute sur la fiabilité de nos intuitions. En l'occurrence, un souci sera accordé, spécifiquement, aux arguments étiologiques. Cela permettra de préparer le terrain pour la critique subséquente.

1.1 Pourquoi l'intuitionnisme de Huemer plutôt qu'un autre?

La raison pour laquelle l'intuitionnisme huémérien a été sélectionné pour la rédaction de ce mémoire est qu'il s'agit d'un intuitionnisme plus modeste que celui prôné par un Robert Audi ou un Russ Shafer-Landau². En effet, Huemer ne prétend pas que les intuitions morales soient des propositions « évidentes par elles-mêmes³ », mais simplement des propositions qui justifient nos croyances morales de manière non inférentielle (Stratton-Lake 2020). Pour cette même raison, il semblerait que la version huémérienne de l'intuitionnisme éthique soit la plus plausible. De même,

² À des fins dialectiques, on tâchera d'attaquer la version la plus convaincante de l'intuitionnisme éthique.

³ Pour Audi et Shafer-Landau, une proposition est « évidente par elle-même » dans la mesure où la « compréhension adéquate » de cette proposition suffit pour justifier notre croyance en celle-ci. Si cette notion d'évidence n'exclut pas la possibilité de défaite épistémique, elle demeure relativement exigeante (Bedke 2010, 1071–72).

Huemer propose plusieurs innovations épistémologiques intéressantes, comme le principe du conservatisme phénoménal, lesquelles servent à contrer l'offensive sceptique⁴.

1.2 Le positionnement métaéthique de Huemer

L'intuitionnisme éthique comporte deux versants majeurs : l'un métaphysique et l'autre épistémologique (Kirchin 2005, 582; Huemer 2009, 222). Sur le plan métaphysique (ou ontologique), Huemer adopte un réalisme non naturaliste, posture selon laquelle il existerait des propriétés morales objectives et *sui generis*. D'une part, on suppose que ces entités morales sont parfaitement indépendantes de nos attitudes évaluatives particulières et contingentes⁵. D'autre part, on affirme que ces entités morales ne seraient en aucun cas réductibles à des entités naturelles^{6,7}. Ce positionnement métaphysique est doublé d'un cognitivisme (Huemer 2008, 370). Cela signifie que le langage moral vise à décrire une réalité objective. En l'occurrence, on dira que les énoncés normatifs contenant des termes tels que « bien » et « mal » peuvent être vrais ou faux. Outre ces considérations sémantiques, l'éthique huémérienne repose sur une épistémologie intuitionniste, laquelle permettrait une connaissance morale de la forme synthétique *a priori* (Huemer 2016, 1987, 2003)⁸. En ce sens, l'agent épistémique intellige (*cognizes*) des propositions morales, lesquelles jouissent d'une justification non inférentielle (qui n'est pas dérivée de propositions autres) (Huemer 2008, 211). Cet intuitionnisme rationaliste s'appuie sur le conservatisme phénoménal, théorie développée et défendue par Huemer⁹. Le versant épistémologique de l'intuitionnisme huémérien sera décrit en détail dans les prochaines sections.

⁴ Aucune des versions classiques de l'intuitionnisme ne sera traitée ici (comme celles de R. Cudworth, S. Clarke, R. Price, T. Reid, D. Stewart, W. Whewell, H. Sidgwick, H.A. Prichard, W.D. Ross, G.E. Moore, etc.) (dans Brink 2014, 666).

⁵ Les faits moraux ne peuvent pas non plus constituer la volonté d'une subjectivité idéalisée, comme celle de Jésus ou de Dieu.

⁶ La métaphysique de Huemer est d'allégeance platonicienne ou frégréenne, dans la mesure où les propriétés morales ne possèdent ni *spatialité*, *temporalité* ou *efficience causale*. Sur le plan ontologique, elles sont similaires à des propriétés mathématiques ou logiques, lesquelles existeraient dans un « troisième domaine » (Bengson 2015, 1–38).

⁷ Toutefois, cela n'empêche pas que la qualité morale d'une action survient à la totalité des propriétés non morales de cette action (Huemer 2008, 203, 207).

⁸ Par synthétique *a priori*, on entend deux choses. D'une part, la connaissance morale ne découle pas simplement de l'analyse terminologique (clause sémantique). D'autre part, la connaissance morale proviendrait de la raison. Elle serait donc indépendante de l'expérience (clause épistémologique).

⁹ Les implications psychologiques de l'intuitionnisme de Huemer ne sont pas directement abordées dans ses textes. Or, comme on le verra plus loin, il s'agit d'un angle mort important à son épistémologie morale.

1.3 La thèse du conservatisme phénoménal

1.3.1 La formulation du principe

Dans son œuvre intitulée *Ethical Intuitionism*, Michael Huemer défend un réalisme moral robuste basé sur le principe du conservatisme phénoménal (CP) (Huemer 2008, 99–127). Proposée pour la première fois dans *Skepticism and the Veil of Perception* (Huemer 2001, 99), cette théorie épistémologique peut être énoncée ainsi :

CP : « S’il semble à S que P, donc, en l’absence de défaiteurs, S possède au moins une certaine justification pour croire que P [ma traduction]¹⁰. » (Huemer 2007, 30)

Pierre angulaire de l’éthique huémérienne, le conservatisme phénoménal (CP) suppose que l’intuition¹¹ de X justifie directement la croyance de X, mais seulement de manière *prima facie* défaisable. En ce sens, la justification initialement procurée par les intuitions peut être déboulonnée par des facteurs invalidants, que l’on nomme « défaiteurs ». C’est pour cette raison que CP garantit seulement une justification *prima facie*¹².

À ce titre, Huemer distingue entre la justification pour *adopter* une croyance et la justification pour *retenir* une croyance. *Primo*, le sujet épistémique est justifié d’*adopter* la croyance de X s’il lui semble que X, et qu’il n’a aucune raison de douter de X. *Secundo*, le sujet est justifié de *retenir* la croyance de X si, et seulement si, il se rappelle que X, et il n’acquiert aucun défaiteur pour X. Et donc, CP ne garantit pas que le sujet n’obtiendra pas de défaiteur pour sa croyance, mais seulement que la croyance de X est *prima facie* justifiée par le fait d’avoir l’apparence de X, sans plus. Comme on le verra plus loin, les défaiteurs peuvent être de deux sortes, *réfutants* ou *discréditants*. Dans le premier cas, il s’agit de prouver non-X, et dans le second, de montrer que la croyance de X n’a pas été formée par un processus fiable (Huemer 2009, 226).

¹⁰ Pour la suite du texte, on assumera que les citations en français sont des traductions de ma part. La mention « ma traduction » sera donc omise à partir de ce point.

¹¹ Dans la pensée de Huemer, les termes « intuition » (*intuition*) et « apparence » (*appearance/seeming*) sont généralement considérés comme des synonymes. Ceci dit, le terme intuition est plus spécifiquement employé pour les apparences de type intellectuel (Huemer 2001, 99).

¹² Le principe du conservatisme phénoménal se rapporte également à des expressions plus communes comme « les apparences sont innocentes jusqu’à preuve du contraire » et « les apparences sont trompeuses ».

Il est aussi utile de souligner que CP n'offre aucune justification *prima facie* aux croyances issues de l'aveuglement, du déni ou de la foi (Huemer 2001, 109). Dans les cas énumérés ici, les croyances ne sont pas basées sur des apparences. L'aveugle nie sciemment ses apparences, et le fidèle est le seul artiste de ses fictions (Huemer 2006, 156). Des agents épistémiques souffrant de dommages cérébraux ou de troubles mentaux invalidants ne pourraient pas non plus se prémunir de la justification fournie par CP (Huemer 2007, 39).

Agissant comme une source de justification non inférentielle¹³ et défaisable, les intuitions¹⁴ répondent à l'exigence de fondement (*basing requirement*) pour les croyances. C'est-à-dire qu'elles confèrent une justification épistémique aux croyances correspondantes. Plus spécifiquement, l'exigence de fondement (chez Huemer) comporte que certaines croyances sont *doxastiquement* justifiées précisément parce qu'elles sont basées sur des apparences, lesquelles confèrent une justification *propositionnelle* (Huemer 2009, 227). Et donc, la condition nécessaire et suffisante pour obtenir la justification *propositionnelle* pour X est d'avoir une expérience de X (Huemer 2009, 231). Là réside l'essence du conservatisme phénoménal¹⁵.

1.3.2 Internalisme, externalisme et théories rivales

CP est une théorie *internaliste* de la justification (Huemer 2001, 104). En l'occurrence, toutes les conditions qui confèrent la justification surviennent à la manière dont les choses apparaissent au sujet (Huemer 2006, 147–48). CP est une variante de *l'internalisme phénoménal*¹⁶, car elle survient spécifiquement aux apparences, et les apparences sont des états mentaux dont le sujet peut

¹³ La justification dont jouit ces croyances est fondationnelle, ou non inférentielle, car elle ne dérive d'aucune autre croyance justifiée (Huemer 2001, 100–3; Markie 2005, 351).

¹⁴ Notons que, pour Huemer, les intuitions elles-mêmes ne peuvent pas être justifiées :

« But, it will be objected, one belief can only justify another belief if the first belief is, itself, justified. Similarly, therefore, shouldn't we say that a perceptual experience can only justify a perceptual belief if the experience, itself, is justified? But the latter condition makes no sense. It does not make sense – it is a category error – to say that an experience is justified or unjustified. » (Huemer 2001, 97)

¹⁵ À l'occasion, un éclaircissement terminologique s'impose : Généralement, les épistémologues distinguent entre la justification *propositionnelle* et la justification *doxastique*. La justification propositionnelle requiert seulement qu'un agent *possède* de bonnes raisons pour croire une proposition. La justification doxastique, quant à elle, requiert non seulement que l'agent *possède* de bonnes raisons pour croire une proposition, mais aussi qu'il *croit* cette proposition en vertu de bonnes raisons (Hasan 2018).

¹⁶ Dans « Phenomenal Conservatism and the Internalist Intuition », Huemer dénombre cinq versions de l'internalisme : la version des « états internes » (*internal state account*), la version de la « conscience actuelle » (*actual awareness account*), la version de la « conscience potentielle » (*potential awareness account*), la version de « l'introspectabilité » (*introspectability account*) et la version « phénoménale » (*appearance account*) (Huemer 2006, 147–48).

être conscient de manière introspective. À cet égard, CP s'oppose au fiabilisme, une théorie épistémologique voulant que la justification soit fonction de la fiabilité d'une méthode donnée. La fiabilité est un mode de l'*externalisme*, car elle ne survient pas aux états mentaux du sujet. En quelques mots, le sujet peut être justifié de croire une proposition quelconque, sans même connaître les conditions justification. Il suffit que la méthode soit fiable. Rien d'autre n'est exigé (Huemer 2006, 148–50).

Huemer défend une théorie internaliste, car il considère que l'option externaliste comporte des conséquences contre-intuitives. Notamment, l'interprétation externaliste suggère que, devant deux apparences [X et Y] équivalentes sur le plan phénoménal, un sujet pourrait être contraint d'adopter deux attitudes doxastiques opposées. En ce sens, il serait forcé d'affirmer l'un [X] et de nier l'autre [Y], même si les deux cas sont phénoménologiquement indistinguables (Huemer 2006, 151). Ce genre de situation peut se produire lorsque le processus qui a produit X est fiable, tandis que le processus qui a produit Y ne l'est pas. Dans ce cas, le sujet doit accorder son assentiment doxastique à l'un plutôt qu'à l'autre, sans avoir identifié de facteur discriminant d'un point de vue phénoménal.

Afin d'illustrer cette problématique, il est pertinent de rapporter et de résumer l'exemple formulé par Huemer. Dans le cas du *Cerveau clairvoyant*, nous avons que Susanne forme des croyances de deux manières : par la perception sensible apparente et par la perception extrasensorielle (clairvoyance). Un jour, Susanne a la perception sensible apparente d'un chien et la perception extrasensorielle d'une licorne dans un champ. On sait que la perception sensible apparente de Susanne n'est pas fiable et qu'il n'y a pas de chien devant elle. On sait aussi que la perception extrasensorielle de Susanne est fiable et qu'il y a une licorne quelque part dans un champ. (Mais Susanne ne sait pas quel mécanisme est fiable et quel mécanisme ne l'est pas.) Phénoménologiquement, Susanne possède la même justification pour croire qu'il y a effectivement un chien devant ses yeux et une licorne quelque part. Or, le fiabilisme veut que Susanne retienne seulement la perception extrasensorielle de la licorne, car il s'agit du seul mécanisme fiable qu'elle possède. Dans ce cas précis, il serait raisonnable pour Susanne de rejeter la perception du chien et d'accepter la perception de la licorne, même si elle ne possède aucune raison consciente pour en faire ainsi (Huemer 2006, 151–52).

Selon Huemer, le fiabilisme et l'externalisme, plus généralement, ont des implications fortement contre-intuitives (Huemer 2007, 48). Comme le suggère l'exemple précédent, l'externalisation du critère de justification contraint le sujet épistémique à effectuer des choix doxastiques curieux. Possédant deux apparences similaires en tous aspects pertinents, le sujet serait censé rejeter l'une des apparences en raison d'un facteur extra-mental, d'une raison étrangère. De toute évidence, une telle méthode ne peut être raisonnablement entérinée. Il en résulterait une épistémologie de l'arbitraire et de l'inexplicable (Huemer 2006, 155). Au contraire, CP accommode cette intuition robuste selon laquelle la justification survient entièrement aux apparences (internes) du sujet.

Aussi, Huemer exprime son désaccord envers la doctrine du *fiabilisme internaliste* de Steup. Cette théorie suppose que « le sujet est justifié de croire que P lorsqu'il a une apparence de P, n'a pas de défauteurs pour P, et possède des données mnémoniques soutenant la fiabilité de l'apparence » (Huemer 2013, 348–49). L'auteur rejette la composante fiabiliste de la théorie, car il juge que la mémoire est elle-même une sorte d'apparence. Si la mémoire est une forme d'apparence, il ne sera jamais possible d'obtenir la justification initiale requise pour intégrer la mémoire au processus justificatif. Pour cette raison, il ne sera jamais possible de satisfaire la clause de fiabilité et de démarrer le raisonnement. Maintenant, on pourrait considérer que la mémoire n'est pas une apparence, mais cela impliquerait qu'aucune donnée mnémonique ne pourrait jamais être justifiée et justifiante (car le sujet ne pourrait avoir une telle apparence). Pour ces deux raisons, Huemer juge que le fiabilisme internaliste de Steup n'est pas viable et que le conservatisme phénoménal est la meilleure alternative (Huemer 2013, 348–49).

Huemer rejette également la théorie de l'*accointance simple* (Huemer 2013, 347–48), selon laquelle la justification réside dans la conscience directe (*direct awareness*) d'un fait. En l'occurrence, le sujet est justifié de croire une proposition, si et seulement si le sujet est accointé avec la vérité d'une proposition. À cet égard, on dira que l'accointance est une relation factive. C'est-à-dire que l'accointance de P implique la vérité de P¹⁷. Étant une rivale du conservatisme phénoménal, cette théorie internaliste souffre toutefois d'un vice rédhibitoire, le problème de la faillibilité. En ce sens, la théorie de l'accointance simple serait irrecevable, car elle exclut la

¹⁷ Comme vous l'aurez deviné, les apparences sont des attitudes propositionnelles *non factives* (Huemer 2007, 44).

possibilité que nos intuitions soient défaites. Comme l'explique Huemer, une situation pourrait surgir où l'agent épistémique pense être accointé avec un fait particulier, mais ne l'est pas réellement, puisqu'un défauteur insoupçonné s'est immiscé dans le processus. Pensons à la théorie naïve des ensembles qui, malgré son apparence de vérité, fut éventuellement réfutée par le paradoxe de Russell (Huemer 2007, 45). Simplement dit, la théorie de l'accointance simple serait viciée, car elle serait incapable d'accommoder les croyances (non inférentielles) justifiées et *défaites*.

Certains, comme Richard Fumerton, ont tenté de contourner ce problème en proposant des amendements à la théorie de l'accointance. En l'occurrence, l'auteur propose un affaiblissement de la condition de justification. Selon le principe d'accointance avec similarité, est justifié celui qui est accointé avec un fait *suffisamment similaire* au fait pertinent. Si une telle modification permet d'accommoder certains cas limites, où deux apparences sont sensiblement similaires, elle ouvre aussi la porte à une variété de cas où deux apparences *semblent* être similaires, mais ne le sont pas réellement. Et nous voilà de retour à la case départ (Huemer 2006, 153).

Mais encore, on pourrait tenter de dissoudre le problème en affirmant qu'il faille seulement qu'un fait *semble* être *similaire* [mais pas *suffisamment similaire*] au fait pertinent. Or, cet affaiblissement subséquent n'ajoute rien sur le plan de la justification, comparativement à la théorie défendue par Huemer. En effet, s'il semble au sujet que X et il semble au sujet que X est *seulement* similaire à Y, il doit forcément sembler au sujet que X n'est pas Y. D'un point de vue épistémique, il est rationnel de retenir uniquement l'apparence de X. De toute évidence, la relation « sembler être similaire » est redondante, superflue, et donc stérile. Pour cette raison précise, il suffit de conserver seulement la notion d'apparence, laquelle nous renvoie à l'internalisme phénoménal prôné par Huemer (2006, 154).

Ensuite, l'auteur aborde l'*évidentialisme* prôné par Earl Conee, théorie selon laquelle le sujet S a un quelconque degré de justification défaitable pour X, si la croyance de X est basée sur de l'évidence justifiant de manière propositionnelle (Huemer 2013, 346). Huemer s'accorde avec cette affirmation, mis à part un élément. Il souligne que l'évidence qui justifie X de manière propositionnelle est toujours constituée d'apparences. Ceci signifie que l'évidentialisme serait réductible au conservatisme phénoménal.

Jusqu'à présent, ont été présentés des arguments ciblés contre le *fiabilisme externaliste*, le *fiabilisme internaliste*, la théorie de *l'accointance simple*, la théorie de *l'accointance avec similarité*, la théorie de *l'accointance avec apparence de similarité* et *l'évidentialisme*¹⁸. Or, une telle exposition serait incomplète sans l'argument de l'auto-réfutation, lequel réunit une pluralité d'arguments sous une forme commune.

1.3.3 L'argument de l'auto-réfutation

Élément central de la démonstration huémérienne, l'argument de l'auto-réfutation (*self-defeat argument*) consiste à montrer que le rejet du conservatisme phénoménal est absurde (Huemer 2001, 104–7; Huemer 2007, 39–41, 54; Huemer 2008, 101). En d'autres mots, ceux qui considèrent que CP est faux ne seraient pas justifiés de croire une telle chose. L'argument est fort simple :

P1 : Toutes les croyances (qui sont des candidats plausibles à la justification) sont basées sur les apparences.

P2 : Une croyance est doxastiquement justifiée seulement si elle est basée sur une source de justification propositionnelle.

C1 : Pour qu'une croyance quelconque soit justifiée, les apparences doivent être la source de justification propositionnelle. (Huemer 2013, 341)

La prémisse #1 repose sur un constat *empirique*. En l'occurrence, on remarque que nos croyances sont causalement déterminées par nos apparences. Cela signifie que les contenus de nos croyances s'inscrivent dans une relation de survenance avec nos apparences. À titre d'exemple, si un agent forme la croyance perceptuelle qu'un chien se trouve devant lui, cette croyance perceptuelle doit être basée sur l'expérience sensorielle correspondante. Similairement, si un sceptique forme la croyance que les apparences ne sont pas fiables, cette croyance doit être basée sur l'intuition intellectuelle correspondante (Huemer 2007, 40).

La prémisse #2 constitue une clause *épistémique*. Cette clause stipule que la justification doxastique ne peut obtenir sans la justification propositionnelle. À cet égard, Huemer distingue

¹⁸ L'auteur explore d'autres options, comme le conservatisme phénoménal restreint, mais il semble raisonnable de croire que les alternatives principales ont été traitées ici (Huemer 2007, 42–44).

entre le simple fait *d'avoir* la source de justification propositionnelle pertinente et *d'avoir une croyance basée sur* la source de justification propositionnelle pertinente. Afin d'illustrer la chose, l'auteur donne l'exemple d'un particulier, nommé Henri, et de ses pratiques doxastiques douteuses. Notre protagoniste, Henri, possède toutes les raisons appuyant la théorie de l'évolution. Toutefois, il rejette ces raisons, en bloc, étant déterminé à ne jamais accepter ladite théorie (pour aucune raison épistémiquement valable). L'attitude de Henri change, un jour, lorsque sa lectrice de tarot lui annonce que la théorie de l'évolution est vraie. Désormais, Henri est convaincu de la vérité de la théorie. Il est nouvellement converti. Dans ce cas précis, il apparaît qu'Henri *possède* la justification propositionnelle pertinente pour croire la théorie concernée. Or, le protagoniste ne *base* pas sa croyance sur cette évidence, ou source de justification. Sa croyance est basée sur une lecture de cartes, laquelle ne constitue *pas* une source de justification propositionnelle. Et donc, un agent épistémique peut entretenir une *croyance vraie*, posséder la *justification* pour cette croyance, sans toutefois mobiliser cette source de justification. Il en découle que l'agent entretient une croyance vraie, non justifiée. À la lumière de cet exemple, il est possible de saisir la différence entre la *possession* et l'*usage* d'une source de justification. C'est précisément l'usage de cette justification qui rend une croyance justifiée. Sommairement, il ne suffit pas d'avoir une croyance vraie, encore faut-il que l'on ait cette croyance *pour les bonnes raisons* (Huemer 2007, 41).

En combinant P1 & P2, nous avons que, pour toute croyance justifiée, les apparences sont la source de justification propositionnelle (C1). Il en découle que les apparences doivent fournir, à tout le moins, une justification défaisable pour nos croyances. Faute de quoi, nos croyances ne peuvent être justifiées. Ceci signifie qu'il est impossible de justifier la croyance que le conservatisme phénoménal est faux. *Ergo*, toute théorie épistémologique qui affirme la négation de CP est auto-réfutante (car chacune d'entre elles implique tacitement CP) (Huemer 2007, 41).

Plus encore, Huemer rajoute que CP reflète la pratique épistémique courante. C'est-à-dire que notre pratique doxastique ordinaire nous engage implicitement à recourir aux apparences (Huemer 2007, 39, 50; Huemer 2008, 101). En ce sens, la manière par laquelle nous procédons normalement est la manière par laquelle nous devons procéder. Ici, le passage du descriptif au normatif est justifié par le simple fait que la pratique épistémique serait injustifiable sans CP. C'est le constat de la nécessité concrète de CP. Mais encore, on remarque que l'intuition de CP justifie

automatiquement la croyance de CP¹⁹. C'est un principe qui se justifie par lui-même. Comme le dirait Bedke, il s'agit d'un processus justificatif par lequel l'agent épistémique se soulève en tirant sur les languettes de ses propres bottes (*epistemic bootstrapping*) (Bedke 2010, 1075)²⁰.

Quiconque accepte ce résultat logique, et reconnaît la plausibilité de CP, sera aussi en mesure de répondre convenablement au défi sceptique, affirme Huemer. En effet, il semblerait que le scepticisme épistémologique devienne indéfendable si l'on admet CP. Pourquoi?

La thèse du sceptique (épistémologique), dans son expression la plus forte, comprend que les intuitions ne sont pas une source de justification pour nos croyances, qu'elles ne peuvent mener à la connaissance. Le problème avec cette perspective est patent : le sceptique ne peut exercer son doute sans faire appel aux intuitions et, par extension, au principe du conservatisme phénoménal. Spécifiquement, le sceptique ne peut justifier son scepticisme, car il ne possède pas l'outil pour le faire. Résultat : le scepticisme épistémologique, au sens radical, est intenable (Huemer 2007, 49–50). Mais encore, certaines apories sceptiques comme le problème du malin génie (cerveau dans une cuve), le problème de la régression à l'infini et, plus généralement, le problème du critère sont essentiellement dénouées par CP. N'en déplaise à Agrippa ou Sextus Empiricus, il est impossible de formuler un argument sceptique sans se fier aux intuitions intellectuelles. Or, comme Huemer le rappelle, la justification procurée par CP est défaisable, et les croyances qui en découlent, révisables. CP est un point de départ et non une ligne d'arrivée.

Afin de couronner sa défense du conservatisme phénoménal, Huemer lance un défi à ses détracteurs :

Si les apparences non défaites ne sont pas une source de croyance justifiée, alors, comment peut-on éviter le scepticisme sur le monde extérieur, le passé, les valeurs, les objets abstraits et encore? À moins que ce problème puisse être résolu [par une théorie concurrente], il serait sage de faire confiance aux apparences, et puis d'entériner un principe comme celui du conservatisme phénoménal. (Huemer 2013, 349)

¹⁹ Certains, comme Bonjour, pourraient clamer que Huemer ne répond pas entièrement ou correctement à la demande de méta-justification, sachant que la justification employée ici est circulaire. Or, Huemer considère qu'une telle demande est injuste (Huemer 2007, 50–54).

²⁰ Dans le cas présent, je ne me prononcerai pas sur la légitimité d'un tel processus et de son emploi.

Comme en témoigne l'extrait ci-haut, Huemer prétend que le conservatisme phénoménal est la théorie épistémologique la mieux placée pour répondre au desideratum de justification. Selon l'auteur, il s'agirait de la solution la plus convaincante au scepticisme épistémologique moral, devant une variété de théories concurrentes.

1.4 Les intuitions selon Huemer

1.4.1 Ce qu'elles sont

Chez Huemer, les intuitions constituent des états phénoménaux (*seeming states*). On peut les décrire ainsi :

« Pour que S ait une intuition que P, S doit avoir un état mental selon lequel il lui semble que P. » (Bedke 2010, 1075)

Ces intuitions constituent des attitudes propositionnelles pré-doxastiques (Huemer 2007, 30). Autrement dit, ce sont des paires attitude-contenu, dont le contenu propositionnel n'a pas obtenu l'assentiment doxastique de l'agent épistémique. Comme le souligne Huemer, l'apparence n'implique pas nécessairement la croyance, parce qu'il est possible d'avoir des apparences perceptuelles sans nécessairement les ériger au rang de croyance (Huemer 2007, 31). De cette manière, un agent quelconque peut s'abstenir d'accorder son assentiment doxastique à certaines apparences, s'il considère que leur statut épistémique est incertain²¹.

Par exemple, notre agent peut être confronté à une illusion d'optique, comme la déformation d'une paille dans un verre d'eau, sans toutefois croire que la paille est effectivement courbée, ou même brisée. En effet, si l'agent a fait l'expérience antérieure de cette illusion, ou s'il comprend que l'apparence est causée par la réfraction, il n'accordera pas son assentiment doxastique aux contenus appréhendés. Avec l'entraînement, le sujet est en mesure de distinguer entre les apparences véridiques et les apparences trompeuses (Huemer 2001, 99–101).

²¹ À ce titre, Huemer est partisan du *perceptualisme* et opposant du *doxasticisme*. Le perceptualisme comporte que les intuitions sont des expériences perceptuelles pré-doxastiques, lesquelles représentent des contenus abstraits. À l'inverse, le doxasticisme suppose que les intuitions sont des attitudes doxastiques, comme des jugements ou des inclinations à juger (Bealer 2000, 1–10; Huemer 2001, 93–117; Chudnoff 2011, 626). On en dira davantage là-dessus plus loin.

Au grand chagrin des partisans de l'analyse linguistique, Huemer reconnaît qu'il est incapable produire une réduction, en termes primitifs, du concept d'apparences (Huemer 2013, 328–29). À défaut de fournir une analyse philosophique rigoureuse, il réitère simplement que les apparences sont des états mentaux conscients, avec contenu propositionnel, et que tous les individus normaux sont familiers avec de tels états. Pour Huemer, il n'est pas nécessaire de connaître les contours exacts de la notion, il suffit que les agents épistémiques reconnaissent l'existence d'une chose semblable.

Brièvement, on se rappellera que les apparences sont antérieures aux croyances, et qu'elles constituent leur soubassement. Comme le veut CP, ces apparences confèrent une justification défaisable aux croyances correspondantes. Cette justification est non inférentielle.

1.4.2 Ce qu'elles ne sont pas

Selon Huemer, nos intuitions ne sont pas des états doxastiques. Sur ce point, il s'oppose à David Lewis, philosophe affirmant que nos intuitions se réduisent à de simples opinions (Pust 2019). De même, les apparences huemériennes ne sont pas des dispositions à croire (Huemer 2013, 329), contrairement à ce que dirait Peter van Inwagen, ni des inclinations conscientes à juger, au sens de Timothy Williamson (Huemer 2007, 31; Pust 2019). Similairement, ces intuitions ne se rapportent pas à des contenus représentationnels auxquels nous serions portés à assentir, comme le voudrait Ernest Sosa (Sosa 2006, 101; Huemer 2007, 31). Si la conception des intuitions parrainée par Huemer ressemble à celle d'un Elijah Chudnoff ou d'un John Bengson, pour lesquels il s'agit de propositions intellectuelles présentées au sujet comme étant vraies, il serait probablement trop généreux de poser une identité entre ces conceptions (Chudnoff 2011, 313–33; Chudnoff 2011, 625–54).

Jusqu'à présent, il a seulement été question de conceptions philosophiques de l'intuition. Or, il serait tout aussi pertinent d'aborder les intuitions avec la lunette des psychologues, philosophes expérimentaux et neuroscientifiques. Dans ces domaines, les intuitions sont généralement décrites comme des jugements spontanés et immédiats, lesquels concernent la vérité d'une proposition (Hannon 2018, 4149). Ces intuitions doivent être contrastées avec les jugements réflexifs, lesquels sont le produit d'une réflexion plus approfondie. Est-ce que la conception huemérienne des intuitions cadre avec celle des philosophes expérimentaux? À première vue, on pourrait croire qu'il

en est ainsi. Toutefois, le prochain chapitre montrera que le fonctionnement réel du cerveau n'est pas parfaitement compatible avec le modèle huémérien, surtout pour les intuitions morales substantives.

1.4.3 Les quatre types d'intuitions

Ensuite, l'auteur recense quatre types d'apparences : perceptives, mnémoniques, introspectives et intellectuelles (Huemer 2008, 102). Les apparences perceptives concernent le donné sensoriel, c'est-à-dire la présentation de contenus propositionnels par les sens. Elles sont du genre : « Il semble qu'une tasse se trouve sur la table ». Les apparences mnémoniques se concentrent sur la mémoire. Elles sont du genre : « Il m'apparaît que je me suis brossé les dents ce matin ». Les apparences introspectives, quant à elles, regroupent les apparences de désir et autres phénomènes psychologiques relatifs à la conscience de soi (*self-awareness*) (Huemer 2007, 33; Bedke 2010, 1075). Elles sont du genre : « Il me semble que je désire manger des biscuits », « Il me semble que j'ai mal au doigt » ou « Il me semble que j'ai espoir de ... » (Huemer 2007, 32–33). Et puis, les apparences intellectuelles concernent la compréhension de normes, de relations, de règles, et la perception de vérités conceptuelles. Elles sont du genre : « Il me semble que le *Modus Ponens* est une règle d'inférence valide », « Il me semble que le trajet le plus court entre deux points est une ligne droite » ou encore « Il me semble qu'aucun objet ne puisse être à la fois entièrement rouge et entièrement bleu » (Huemer 2008, 100–1). Aussi, les apparences intellectuelles comptent deux sous-catégories : évaluatives et non évaluatives. Les apparences évaluatives peuvent, à leur tour, être morales et non morales.

1.4.4 Les intuitions intellectuelles morales

Les intuitions morales sont des apparences intellectuelles initiales *sui generis*, lesquelles agissent comme des sources d'une justification non inférentielle (Huemer 2008, 370–71). Perçues par la faculté *intellectuelle*, ces intuitions désignent des états cognitifs, lesquels possèdent un contenu propositionnel. On dira qu'il s'agit d'apparences intellectuelles *initiales*, car elles surgissent avant toute activité de raisonnement, ou de réflexion (Huemer 2008, 101). Elles sont *sui generis*, ou de leur propre genre, car elles ne peuvent être réduites à d'autres types d'intuitions. Plus exactement, elles se distinguent des intuitions intellectuelles non morales, ou de toute autre intuition non intellectuelle. (On remarquera aussi que les *intuitions* morales pointent vers des

propriétés morales *sui generis*, lesquelles jouissent d'un statut métaphysique privilégié.) Ces intuitions agissent comme des sources d'une justification non inférentielle, car cette justification ne dérive pas d'une autre croyance préalablement admise.

En pratique, les intuitions morales surgissent lorsque le sujet est confronté à un dilemme moral, ou lorsqu'il doit répondre à une question morale. De même, l'appel aux intuitions sera nécessaire pour justifier des règles d'action générales, ou pour formuler des principes plus abstraits.

1.4.5 Les types d'intuitions morales

S'inspirant de la littérature intuitionniste, Huemer distingue trois types d'intuitions morales, lesquelles peuvent être classées selon leur niveau de généralité. En allant du moins général au plus général, il y a les intuitions concrètes, intermédiaires et abstraites²². Les intuitions concrètes sont celles qui concernent des cas particuliers, comme la noyade de Singer (1972), l'explorateur de Williams (1973), le problème du trolley (Foot 1967) et ses variantes (Thomson 1985; Huemer 2008, 104). Ces intuitions se rapportent le plus souvent à des expériences de pensée, des dilemmes moraux ou des enjeux relevant de l'éthique appliquée.

Les intuitions intermédiaires, comportant un niveau de généralité moyen, regroupent des dictats comme celui de tenir nos promesses ou de montrer de la gratitude. Celles-ci concernent, le plus souvent, des énoncés normatifs comme les devoirs *prima facie* de Ross (1930). En ce sens, les intuitions intermédiaires sont des énoncés éthiques déontologiques, lesquels résultent d'une généralisation opérée à partir d'intuitions concrètes (Huemer 2008, 383–84). Cela dit, on aurait tort de réduire les intuitions intermédiaires aux énoncés déontologiques. En effet, on pourrait raisonnablement affirmer que des énoncés évaluatifs comme « le plaisir est bien », « la douleur est mal » et « le plaisir est mieux que la douleur » figurent aussi parmi les intuitions de niveau moyen (Street 2006, Huemer 2008, 102).

Exprimant un degré de généralité élevé, sinon maximal, les intuitions abstraites se divisent en deux camps. Il existe les intuitions abstraites non formelles (substantives) et abstraites formelles. Du côté non formel, Huemer identifie deux intuitions morales prégnantes : « l'idée que la meilleure

²² Il serait justifié de croire que cette distinction tripartite est tirée de Sidgwick, lequel distinguait entre les intuitions *perceptuelles*, *dogmatiques* et *philosophiques* (dans Brink 2014, 669–70).

action est toujours celle qui a les meilleures conséquences » et « l'idée qu'il est mal de traiter les autres comme des moyens ». À ce titre, les intuitions abstraites non formelles désignent des axiomes de systèmes moraux, comme le principe de maximisation de l'utilité et l'impératif catégorique (Huemer 2008, 383). Parmi ces intuitions abstraites, on recensera aussi la règle d'or chrétienne, la doctrine contractualiste et la théorie aristotélicienne du « juste milieu » (Hooker 2002, 170, 182). On comprendra que ces intuitions substantives désignent de possibles « critères du bien ». Du côté formel cette fois-ci, Huemer considère la transitivité et l'additivité des propriétés évaluatives (mais pas nécessairement morales). La transitivité peut être exprimée ainsi : « Si A est mieux que B, B est mieux que C, alors A est mieux que C. » (Huemer 2008, 383) L'additivité axiologique, quant à elle, peut être énoncée ainsi : « La valeur d'un événement est égale à la valeur combinée de la première moitié et de la seconde moitié de l'évènement. » (Huemer 2008, 384) À ces deux éléments, se rajoutent plusieurs autres exemples d'intuitions abstraites formelles (Huemer 2008, 386), lesquels seront toutefois passés sous silence ici.

Si la chose n'est pas explicitement soulignée par Huemer, il serait pertinent d'indiquer que les intuitions morales peuvent aussi être classées selon leur positionnement dans le discours moral. En ce sens, on peut mentionner qu'il existe des intuitions de premier ordre (normatifs) et de second ordre (descriptifs ou métadiscursifs). Les intuitions de premier ordre reflètent ce que l'on doit faire tandis que les intuitions de second ordre décrivent le discours ou le phénomène moral. Encore une fois, les intuitions métadiscursives peuvent être divisées selon leur contenu formel. D'un côté, figurent les intuitions non formelles, comme « la morale porte sur des propriétés telles que le bien et le mal » (Vavova 2021, 724). De l'autre, figurent les intuitions formelles, comme « les propriétés morales surviennent aux propriétés non morales » (Huemer 2008, 386). Ce sont là des intuitions qui décrivent le phénomène moral d'un point de vue externe.

1.5 Les défaites envisagés par Huemer

Si le principe du conservatisme phénoménal élimine certaines difficultés sceptiques générales, notamment le problème de la régression *ad infinitum*, il n'en demeure pas moins que les intuitions morales particulières demeurent faillibles, chose que Huemer reconnaît lucidement. Dans son article intitulé « Revisionary Intuitionism », Huemer identifie quatre grands types de défaites, lesquels mettraient en doute la fiabilité de nos intuitions morales. Rapidement, nous

avons l'incohérence (1), l'endoctrinement culturel (2), la programmation biologique (3) et les biais personnels (4).

En premier lieu, l'incohérence recouvre les situations où plusieurs intuitions morales entreraient en conflit. Puisant dans les travaux de Peter Singer et de Peter Unger, Huemer formule un exemple où trois intuitions raisonnables donnent lieu à une contradiction :

- 1) Refuser de sauver l'enfant dans l'étang de Singer est sérieusement mal.
- 2) Refuser de donner pour réduire la famine dans le tiers-monde n'est pas sérieusement mal.
- 3) La proximité et la visibilité de la souffrance d'un individu ne sont pas des facteurs moralement pertinents pour évaluer notre obligation d'aider cet individu. (Huemer 2008, 372–74)

Si l'on accepte la proposition 3), il en découle que nous ne pouvons simultanément retenir les propositions 1) et 2) sans commettre une faute logique. Afin de résoudre le paradoxe, certains pourraient exiger l'abandon de 2), de 3), ou encore de 1). Quelle que soit la solution préférée, on peut supposer que celle-ci sera un compromis fort malheureux (Huemer 2008, 372–74). Pour les sceptiques, ce malaise logique montre que l'intuition morale n'est pas une source de connaissance fiable. À la fois fréquents et profonds, ces conflits *intra-systèmes* mineraient la thèse intuitionniste et, plus généralement, celle du réalisme moral.

En second lieu, Huemer aborde l'influence de la culture sur les intuitions morales. L'auteur souligne que plusieurs comportements et pratiques désormais condamnés dans le monde occidental, comme le cannibalisme, l'esclavage et l'infanticide, étaient autrefois entérinés par des groupes culturels, comme les Grecs ou les Romains. En l'occurrence, on constate que les croyances morales de ces divers groupes sont mieux expliquées par des contingences historiques et culturelles que par la perception de vérités morales supposées (Huemer 2008, 374–75). Ces contingences seraient d'ailleurs à l'origine des désaccords et divergences qui existent entre les systèmes moraux (comme conflits *extra-systèmes*).

En troisième lieu, l'auteur se penche sur le rôle des facteurs biologiques et évolutionnaires dans le façonnement de nos jugements moraux. Dans le cas présent, il s'agit de montrer que les pressions évolutionnaires auraient favorisé les individus formant des jugements qui maximisent leur survie, adaptabilité et reproductivité. En ce sens, il s'agirait d'une coïncidence massive si la

sélection naturelle avait porté les humains à appréhender les faits moraux objectifs. Et donc, il serait plus parcimonieux de tenir que nos intuitions morales reflètent des impératifs biologiques plutôt qu'une réalité objective, quelle qu'elle soit. Voilà un autre argument sceptique qui joue contre la thèse réaliste (Huemer 2008, 376–77).

En quatrième lieu, Huemer décrit comment les intérêts personnels et les émotions pourraient corrompre les jugements moraux. D'une part, un agent moral pourrait être porté à défendre des principes moraux qui lui confèrent des avantages, mais à refuser d'autres principes qui contrarient ses intérêts. Par exemple, un éleveur de porcs pourrait exiger la fidélité, la loyauté et l'honnêteté de ses proches et de ses employés, mais refuser toute forme de droits aux animaux non humains, comme les porcs (mon exemple). D'autre part, nous avons que les émotions, surtout les émotions fortes, pourraient aisément nuire au jugement moral. Pour s'en convaincre, il suffit de songer à la fougue qui anime les militants pro/anti-avortement. Aux yeux du sceptique, les intérêts personnels et les mouvements émotionnels sont autant de facteurs qui viennent biaiser le jugement moral. C'est pourquoi les intuitions morales ne seraient pas dignes de confiance.

Conscient de la faillibilité des intuitions morales, Huemer propose une « méthodologie critique », laquelle permettrait d'identifier et d'éliminer les intuitions affectées par des biais non rationnels. Cette théorie comporte cinq grands axes :

- 1) Critère de cohérence interne : L'agent moral doit assembler un corps de croyances morales qui s'accordent mutuellement. Il est préférable que ce système de croyances soit de format élargi.
- 2) Critère de non-controverse : L'agent moral doit prioriser des intuitions qui sont largement adoptées et défendues, à travers les peuples et les âges.
- 3) Critère évolutionnaire : L'agent moral doit favoriser des intuitions qui ne semblent pas être corrélées avec le succès de l'espèce humaine.
- 4) Critère de préférence : L'agent moral doit faire preuve de scepticisme devant les intuitions qui lui procurent un bénéfice ou qui confirment son statut.
- 5) Critère émotionnel : L'agent moral doit se distancier des intuitions qui suscitent des émotions vives, qu'elles soient favorables ou défavorables. (Huemer 2008, 381–82)

Élément central de l'intuitionnisme révisionniste de Huemer, cette « méthodologie critique » permettrait de répondre au défi soulevé par le sceptique. Comportant cinq filtres épistémiques, cette méthode est conçue pour assurer la fiabilité des intuitions morales érigées au rang de croyance. Tout d'abord, le critère de cohérence garantit que les intuitions retenues coexistent

harmonieusement au sein d'un vaste réseau de dépendance. Comme le voudrait le cohérentisme, les intuitions sont justifiées si elles se soutiennent et s'expliquent mutuellement (Huemer 2008, 379–80). Sont injustifiées les intuitions qui provoquent des incohérences, ou qui n'ont aucun pouvoir explicatif. En l'occurrence, il s'agit d'éviter tout conflit *interne au système*²³. Ensuite, le critère de non-controverse sélectionne les intuitions qui sont largement partagées, qui suscitent l'assentiment de la majorité. Ce critère permet de transcender tout désaccord culturel et historique qui pourrait surgir. En d'autres mots, il s'agit d'éviter tout conflit *externe au système*. Le critère évolutionnaire, quant à lui, évince les intuitions qui favorisent la survie et la reproduction de l'espèce. Étant façonnées par les pressions évolutives, ces intuitions sont d'étiologie douteuse et, pour cette même raison, doivent être écartées. Finalement, les critères préférentiel et émotionnel éliminent les intuitions influencées par des biais non rationnels. Les uns concernent l'intérêt personnel et les autres, l'émotion. En appliquant ces cinq grands filtres, suppose Huemer, il est possible de former un corps de croyances morales qui est plus résistant aux assauts sceptiques. Brièvement, on comprendra que l'intuitionnisme révisionniste de Huemer implique une démarche autoréflexive.

Exerçant sa propre méthodologie critique, Huemer constate que les intuitions concrètes et intermédiaires sont plus susceptibles d'être influencées par l'intérêt personnel, les émotions, les biais culturels et les influences évolutionnaires/biologiques (Huemer 2008, 383–84). Par contre, les intuitions abstraites souffriraient du problème de la généralisation abusive (*overgeneralization*), selon lequel nous jugeons improprement de la validité d'une généralisation, celle-ci étant susceptible à des contre-exemples (Huemer 2008, 384–85). (Les intuitions intermédiaires seraient aussi porteuses de ce vice, affirme-t-on.)

Selon Huemer, les intuitions abstraites formelles possèdent un statut particulier, car elles seraient spécialement résistantes aux biais mentionnés ci-haut. Si ces intuitions ne peuvent produire

²³ Ici, il ne faudrait pas confondre cohérence et équilibre réflexif. Huemer considère que la méthodologie de l'équilibre réflexif est simplifiée à outrance (*oversimplified*) et, surtout, qu'elle contribue à perpétuer des erreurs. Pour Huemer, la méthodologie adéquate ne doit pas recycler nos intuitions, mais critiquer ces mêmes intuitions. C'est pourquoi il milite en faveur d'un intuitionnisme révisionniste, lequel rejette les intuitions communes qui ne sont pas le produit de la rationalité (Huemer 2008, 219–23; Huemer 2008, 369, 391).

des croyances substantives, elles demeurent néanmoins robustes aux offensives sceptiques (Huemer 2008, 386–87, 391).

Pour les adeptes du scepticisme, il existe différentes manières d’attaquer l’intuitionnisme éthique de Huemer. Plus généralement, il est possible de questionner la viabilité de l’épistémologie intuitionniste, en tant que mode d’accès aux vérités morales. Plus spécifiquement, il est possible de remettre en question la thèse du conservatisme phénoménal et la conception huémérienne des intuitions. De même, on peut se pencher sur l’étiologie des intuitions morales, et questionner la fiabilité de celles-ci. C’est cette stratégie qui sera privilégiée ici. Une telle offensive se présente sous la forme d’arguments étiologiques.

1.6 Les arguments étiologiques

1.6.1 Structure

Les arguments étiologiques²⁴ (*debunking explanations/arguments*) sont des raisonnements qui montrent que l’origine causale d’une croyance, d’un jugement ou d’une intuition constitue un *défaiteur* pour ladite croyance. En l’occurrence, il s’agit de montrer qu’une croyance est formée par un processus qui ne piste pas la vérité ou qui n’est pas fiable épistémiquement (Kahane 2010, 105–6; Mason 2011, 449–52; Sauer 2018). Les arguments étiologiques adoptent la structure suivante :

P1. Prémisse doxastique / explanandum : S croit que P.

P2. Prémisse explicative / explanans : La croyance de S que P est causalement expliquée par X, où X est la meilleure explication pour la croyance de S que P.

P3. Prémisse épistémique / normative : X est un procédé qui ne piste pas la vérité, X est un *défaiteur* pour P, ou X n’est pas une méthode épistémiquement fiable.

C1 : Conclusion : La croyance de S que P n’est pas justifiée.

Ici, on notera le rôle pivotale des prémisses *explicative* et *épistémique*. La première permet d’identifier la « meilleure explication » pour la croyance de S que P, tandis que la seconde permet

²⁴ Dans la langue française, on pourrait aussi employer les expressions: « arguments généalogiques », « déboulonnage épistémique », « explications causales », etc.

de disqualifier cette même explication sur des bases épistémiques. La prémisse explicative fait appel à l'abduction tandis que la prémisse épistémique fait référence aux normes justificatives^{25,26}.

1.6.2 Les types de défaiteurs

Soulignons ensuite que les arguments étiologiques constituent des défaiteurs discréditants (*undercutting/undermining defeaters*) et non des défaiteurs réfutants (*rebutting/overriding defeaters*). Cela signifie que les arguments étiologiques minent la justification dont jouit une proposition (P), sans toutefois fournir des preuves pour appuyer une proposition contraire (non-P). C'est donc dire que les arguments étiologiques ont uniquement une valeur négative sur le plan dialectique (Kahane 2010, 108). Bien sûr, on reconnaîtra qu'une position alternative deviendra relativement plus plausible, mais seulement d'une manière indirecte, ou accidentelle (Huemer 2001, 111).

Pour ceux qui acceptent la *Loi de Hume*, laquelle interdit l'inférence d'un être à un devoir-être, les défaiteurs réfutants doivent provenir de prémisses évaluatives, tandis que les défaiteurs discréditants peuvent provenir de prémisses non évaluatives. À ce titre, on dira que les arguments étiologiques sont des arguments discréditants de nature non évaluative (Huemer 2008, 380–81).

1.6.3 Une typologie des arguments étiologiques

Maintenant, les arguments étiologiques sont de plusieurs types : historiques, psychologiques et évolutionnaires (Kahane 2010, 106–8; Mason 2010, 771–73). Sur le plan historique, il est bien connu que Friedrich Nietzsche ou Karl Marx ont fait usage de ces types d'arguments. Il suffit de penser à la généalogie de la morale nietzschéenne ou au matérialisme historique de Marx pour s'en convaincre. Sur le plan psychologique et neuroscientifique, des chercheurs, comme Jonathan Haidt (2001; 2007; 2013) et Joshua Greene (2014; 2017), ont identifié une myriade de facteurs, souvent des biais cognitifs, qui viendraient déformer (*distort*) nos croyances. Sur le plan évolutionnaire, Michael Ruse (1986), Richard Joyce (2001; 2007; 2016) et Sharon Street (2006; 2015; 2016) ont travaillé sur différentes formes d'arguments étiologiques ancrés dans les travaux de Darwin sur

²⁵ Il est possible que plusieurs explications équivalentes pour la croyance de S que P existent. Dans ce cas, il s'agit de montrer que chacune des explications en jeu est épistémiquement invalidée et invalidante.

²⁶ À des fins d'exhaustivité, on rajoutera que les arguments étiologiques visent des théories morales cognitivistes et réalistes (Mason 2010, 770–76).

l'évolution. Ces arguments prétendent que l'évolution serait un facteur prédominant dans la formation de nos tendances évaluatives et croyances morales.

1.7 Présentation de ma critique

Ma critique, suivant le patron du *défi explicatif* de Lillehammer, se fera en deux temps. Dans un premier temps, il s'agira d'interroger certaines causes *proximales* (psychologiques) de nos intuitions et, dans un second temps, de questionner certaines causes *distales* (évolutives) de ces mêmes intuitions (Lillehammer 2011, 176–77; Zimmerman 2013; Railton 2014, 832). En étudiant l'étiologie de nos intuitions, il sera possible de déterminer leur degré de fiabilité. Mon hypothèse est qu'une majorité d'intuitions substantives vont révéler un caractère étiologique douteux. Il est toutefois plausible que les intuitions abstraites formelles et métadiscursives résisteront à cette offensive sceptique.

Quelles sont les motivations pour ce genre de critique? D'une part, de nombreux travaux et découvertes ont été effectués récemment dans le domaine de la neuroscience et de la psychologie sociale. Ceux-ci illuminent l'étiologie des intuitions morales. D'autre part, de nouveaux arguments concernant l'évolution ont été développés contre le réalisme non naturaliste. Notons, parmi ceux-ci, la critique généalogique évolutionnaire de Street (2006), laquelle sera explorée plus bas.

A fortiori, il s'agit de remettre en question l'idée selon laquelle les intuitions intellectuelles morales n'entretiendraient pas de différences épistémologiquement significatives avec les autres types d'intuitions. S'adressant à des critiques comme DePaul²⁷, ou des adversaires philosophiques comme Mackie, Huemer soutient (à tort) que les intuitions intellectuelles forment un groupe épistémologiquement homogène et que les intuitions éthiques ne possèdent aucune caractéristique saillante qui devrait miner notre confiance dans leur fiabilité :

Given the reality of intuition in general, ethical intuition is not very different at all from other kinds of intuition. *The only difference between ethical intuitions and non-ethical intuitions is in what they are about* – and that cannot be taken as grounds for the

²⁷ Huemer répond à DePaul en ces termes: « His point seems to be that the several kinds of appearances nevertheless have salient differences, some of which may be epistemologically significant. While I agree that there are obvious differences among kinds of appearances, *I do not believe these differences are epistemologically significant* [mes italiques]. » (Huemer 2009, 229)

queerness Mackie sees, unless we are to reject ethical knowledge merely for being ethical [mes italiques]. (Huemer 2008, 111)

Plus spécifiquement, Huemer souligne que les intuitions intellectuelles non évaluatives et intellectuelles évaluatives sont phénoménologiquement identiques :

[The above evaluative] intuitions seem to be products of rational reflection²⁸, deriving from a grasp of relevant concepts, in just the same way as the intuitions that two is less than three and that no object can be completely green and completely red at the same time. Phenomenologically, *I find no difference between the above evaluative intuitions and mathematical or metaphysical intuitions* [mes italiques]. (Huemer 2009, 233)

Si les intuitions éthiques et métaphysiques sont phénoménologiquement indistinguables, ou presque, cela ne suffit d'aucune manière pour affirmer qu'elles sont indistinguables en tous aspects ontologiquement et épistémologiquement significatifs²⁹. De toute vraisemblance, il existe des différences significatives entre les intuitions intellectuelles non évaluatives et les intuitions intellectuelles évaluatives. Ces différences sont à la fois métaphysiques et épistémologiques. Comme le soulignent DePaul (2009, 210), Bedke (2008, 264) et Kirchin (2005, 606), il est raisonnable de croire que les intuitions éthiques sont particulièrement susceptibles à certains types de défaiteurs, notamment psychologiques et évolutionnaires. Ces défaiteurs n'affectent pas les intuitions métaphysiques et mathématiques de la même manière. En raison de ces vices étiologiques, les intuitions intellectuelles *morales* ne peuvent être considérées comme aussi certaines que les intuitions intellectuelles *non morales*. Mais encore, il importe de souligner que le processus de formation des intuitions intellectuelles morales jette un doute sérieux sur le statut ontologique privilégié que leur accorde Huemer. Si l'on prend en compte les récentes avancées en psychologie expérimentale et en neuroscience, il semble tout à fait raisonnable d'affirmer que les

²⁸ À travers les textes de Huemer, il semble y avoir une incohérence ou, à tout le moins, une ambiguïté logique en ce qui concerne la statut rationnel et réflexif des intuitions. Dans *Ethical Intuitionism*, Huemer nous dit que les apparences intellectuelles initiales sont ce genre de choses qui surgissent *avant* l'exercice de notre raison et que les intuitions éthiques participent des intuitions intellectuelles initiales : « Reasoning sometimes changes how things seem to us. But there is also a way things seem to us prior to reasoning; otherwise, reasoning could not get started. The way things seem prior to reasoning we may call an "initial appearance". » (Huemer 2008, 102) Or, plus loin, Huemer affirme que l'intuition est fonction de la raison : « Intuition is a function of reason. Moral intuition differs from mathematical intuition in the way that perception of cars differ from perceptions of trees-that is, merely in having different objects. » (Huemer 2008, 215-16) Et, dans la citation présentée ci-haut, Huemer affirme que les intuitions évaluatives sont des produits de la réflexion rationnelle. Clairement, les intuitions ne peuvent être à la fois pré-rationnelles et rationnelles, à moins que Huemer utilise les termes de manière délibérément différente.

²⁹ L'existence de similarités sur le plan phénoménal (surface) n'implique pas nécessairement l'existence de similarités sur le plan étiologique (profond). Huemer rétorquerait que la justification par CP est garantie lorsque certaines propriétés de surface sont présentes. Or, l'épistémologie ne se limite pas aux propriétés phénoménales.

énoncés moraux n'ont pas le même statut métaphysique que les énoncés mathématiques ou logiques. Conséquemment, il ne serait pas justifié d'affirmer que les propositions morales sont analogues aux propositions mathématiques (en tant que propriétés réelles, non naturelles, *sui generis*), et ce, même si elles sont quasi-identiques d'un point de vue phénoménal.

Nonobstant ces quelques éléments de controverse, il est raisonnable de soutenir que les intuitions morales sont compatibles avec le conservatisme phénoménal, et que celles-ci ne doivent pas être écartées pour la simple et bonne raison qu'elles sont morales. Le problème réside dans la classification de ces apparences en tant que perceptions intellectuelles, au même titre que les vérités mathématiques ou logiques. La thèse que nos intuitions évaluatives sont formées d'une manière analogue aux intuitions non évaluatives peut difficilement être défendue compte tenu des avancées dans le domaine des sciences empiriques. Or, cela ne revient pas à dire qu'il n'existe pas de faits évaluatifs, il s'agit simplement de montrer que *la perception de vérités morales aprioriques supposées n'est pas à l'origine de nos croyances morales factuelles*³⁰. En ce sens, on pourrait affirmer que Huemer se trompe sur l'étiologie des intuitions morales. On précisera toutefois que cette méprise ne recouvre pas toutes les intuitions. Comme on le verra plus bas, certaines d'entre elles pourraient résister aux arguments étiologiques compte tenu de leur caractère plus abstrait. En effet, on remarquera que certaines intuitions morales ressemblent particulièrement aux vérités métaphysiques, ce qui porte à croire qu'elles trouveraient réellement leur origine dans la faculté rationnelle³¹. Or, il ne s'agirait que d'une classe restreinte.

Maintenant, si Huemer reconnaît que les intuitions morales sont susceptibles à une gamme élargie de défaiteurs, comme en témoigne son article intitulé « Revisionary Intuitionism », l'auteur ne va pas suffisamment loin dans sa critique des intuitions morales. Bien qu'il reconnaisse que la majorité de nos intuitions morales soient défaisables, il ne révisé pas sa conception de la morale sur le plan métaphysique et sur le plan épistémologique pour autant. C'est véritablement curieux!

³⁰ Huemer ne prétend toutefois pas que l'ensemble de la connaissance morale serait *apriorique*. Pour lui, certaines intuitions morales *non fondamentales* pourraient être qualifiées d'*apostérieures* (e.g. l'intuition que le caractère d'Adolf Hitler était vicié). On comprendra néanmoins que la valence morale de ces intuitions *non fondamentales* dépend des contenus moraux *fondamentaux*. En ce sens, la connaissance morale *apostérieure* ne serait pas possible sans une connaissance *apriorique* des faits moraux indépendants (Huemer 2016, 1986).

³¹ Selon Huemer, les adeptes de l'intuitionnisme rationaliste doivent simplement affirmer que *certaines* croyances morales constituent la connaissance de faits moraux. Et donc, il suffirait que *certaines* croyances morales soient issues de mécanismes fiables. Toutefois, un tel réalisme est excessivement modeste et difficilement défaisable (car une seule intuition vraie, justifiée et non défaite suffirait pour valider le modèle) (Huemer 2016, 1984–88).

On croirait que la défaisabilité accrue des intuitions intellectuelles morales, comparativement aux intuitions intellectuelles non morales, lui ferait revoir sa conception ontologique et épistémologique de la moralité. Au lieu de réévaluer sa conception, Huemer s'accroche aux intuitions abstraites formelles, lesquelles s'harmonisent avec son modèle. Or, les intuitions abstraites formelles n'informent aucunement nos intuitions d'un point de vue matériel, ou substantiel. N'est-ce pas là une exigence raisonnable pour tout système moral?

En bref, on comprendra que ma thèse est double : nos croyances morales substantives ne sont pas issues d'une appréhension de vérités nécessaires ou *aprioriques*, comme les énoncés mathématiques. Et, conséquemment, ces croyances morales substantives ne sont pas justifiées, car elles ne sont pas formées par des processus qui sont fiables³². Plus formellement, on affirme que :

1. La *meilleure explication* de notre croyance dans certaines propositions morales substantives ne réside pas dans l'appréhension de vérités *aprioriques* par la fonction rationnelle. Par « propositions morales substantives », nous référons aux intuitions concrètes, intermédiaires et abstraites non formelles.

2. Nos intuitions morales substantives ne sont pas justifiées, au sens où le voudrait Huemer, car elles ne résultent pas d'une capacité qui est épistémologiquement fiable (soit la *rationalité pure*). Par « intuitions morales substantives », nous référons aux intuitions concrètes, intermédiaires et abstraites non formelles³³.

Comme il a été mentionné plus haut, la critique se déroulera deux temps. Dans un premier temps, certaines causes *proximales* de nos croyances morales seront examinées. Dans un second temps, certaines causes *distales* de nos croyances seront mises sous la loupe. La combinaison de ces défaiteurs devrait constituer une preuve suffisante pour montrer qu'une majorité d'intuitions morales ne sont pas fondées.

Chapitre 2 : L'hypothèse du double processus et ses implications

³² Même si l'on préconise une approche internaliste de la justification, on exigera tout de même que l'agent épistémique n'ait aucune raison de croire que les processus impliqués dans la formation de ses croyances morales soient non fiables (*unreliable*).

³³ Cela n'exclut pas la possibilité qu'il existe des propriétés morales objectives. J'affirme, simplement, que nous ne connaissons pas la chose morale de la même manière que nous connaissons la chose mathématique. En ce sens, les mécanismes qui produisent les croyances intellectuelles morales sont très différents des mécanismes qui produisent les croyances intellectuelles non morales. Similairement, on notera que le fait de ranger *toutes* les intuitions morales dans la classe des intuitions intellectuelles contribue à créer une impression de fausse certitude entourant leurs origines.

Au cours de ce second chapitre, de récentes découvertes en neuroscience, psychologie sociale et philosophie expérimentale seront mobilisées afin de clarifier l'étiologie des intuitions morales. En premier lieu, il s'agira d'investiguer le rôle des émotions dans la formation des intuitions morales. Ici, les travaux de Joshua Greene (2004; 2008; 2014; 2016; 2017) et Jonathan Haidt (2001; 2007; 2008; 2010; 2013) pourront être approfondis. En second lieu, il s'agira d'interroger le lien entre règles heuristiques et moralité. Pour ce faire, il faudra aller puiser dans les travaux de Daniel Kahneman (1982; 2011), Cass Sunstein (2004; 2005; 2007; 2008) et Walter Sinnott-Armstrong (2010), notamment. En troisième lieu, nous évaluerons l'impact de ces travaux pour l'intuitionnisme rationaliste de Huemer.

2.1 « Révolution affective » et philosophie morale

Depuis la « révolution cognitive » des années 1960, les théories du développement, comme celles de Jean Piaget et de Lawrence Kohlberg, ont mis l'accent sur la raison calculatrice, l'abstraction et le contrôle cognitif, plutôt que les facteurs affectifs (Greene et al. 2004, 389–90). À leurs yeux, le jugement (moral) était synonyme de délibération, de raisonnement, de traitement informationnel, bref, toutes ces tâches que constitue la cognition de « niveau supérieur ». Or, plusieurs travaux parus depuis les années 1980 et 1990 suggèrent que le rôle des facteurs affectifs dans la formation du jugement aurait été sous-estimé. Contribuant à la « révolution affective », des penseurs tels que Robert Zajonc et Joseph LeDoux ont défendu l'idée d'une primauté de l'affect, selon laquelle les données perceptuelles engageraient le système affectif avant la réflexion consciente (Haidt 2007, 998; Haidt et Bjorklund 2008, 181–88; Railton 2014, 827–28;). Tablant sur l'hypothèse de la primauté affective, les psychologues ont développé la théorie du processus double, un modèle psychologique voulant que le système affectif soit le siège de processus automatiques (Système I) tandis que les aires corticales soient le siège de processus cognitifs supérieurs (Système II). Reconnaisant le pouvoir explicatif de cette théorie, des chercheurs comme Joshua Greene et Jonathan Haidt, vont montrer que l'intuition morale ou, plus largement, les jugements moraux seraient le produit d'une synthèse entre affect et rationalité (Greene et al. 2004, 397). Grâce à l'imagerie par résonance magnétique fonctionnelle (IRMf), cette tension entre émotion et cognition a pu être étudiée en profondeur.

2.1.1 La « trolleyologie » de Joshua Greene

Œuvrant à l'intersection de la neuroscience et de la philosophie morale, Joshua Greene et son équipe ont déterminé que l'exercice du jugement moral mobilise à la fois les processus automatiques (Système I) et les processus délibératifs (Système II). Plus spécifiquement, ils ont établi que les processus automatiques soutiendraient préférentiellement les jugements déontologiques tandis que les processus délibératifs soutiendraient préférentiellement les jugements conséquentialistes. Greene exprime cette découverte sous la forme d'un principe:

The Central Tension : Characteristically deontological judgments are preferentially supported by automatic emotional responses, while characteristically consequentialist judgments are preferentially supported by conscious reasoning and allied processes of cognitive control. (Greene 2014, 699)³⁴

Ici, on constate que la tension entre déontologie et conséquentialisme serait véritablement l'expression d'une tension plus profonde entre les régions cérébrales dédiées aux émotions et à la délibération. Cette hypothèse est substantivée par une série d'expériences effectuées par l'équipe de Greene. Ces expériences portent sur le « problème du trolley » et ses variantes.

Dans la version du « spectateur » (*bystander*) ou du « levier » (*switch*), un trolley hors de contrôle se dirige vers cinq personnes, lesquelles seront imminemment percutées et écrasées. Un spectateur peut déplacer un levier pour sauver les cinq personnes. Cela fera dévier le trolley sur une autre voie, tuant ainsi qu'une seule personne (Foot 1967, 5–15; Thomson 1985, 1397).

Dans la version de la « passerelle » (*footbridge*), un trolley hors de contrôle se dirige vers cinq personnes, lesquelles seront imminemment percutées et écrasées (comme dans le cas précédent). Un témoin situé sur une passerelle peut pousser un homme corpulent sur les rails. L'homme sera tué sur l'impact, mais brisera la course du train, sauvant ainsi cinq personnes (Thomson 1985, 1409).

Lorsque des sujets sont interrogés sur l'alternative à préconiser, ils jugent généralement qu'il est *permissible* de déplacer le levier, mais qu'il n'est pas *permissible* de pousser l'homme sur les rails, *même si* le nombre de vies sauvées est identique dans les deux cas. Et donc, les sujets auraient

³⁴ Par « caractéristiquement déontologique », on entend tous les jugements qui sont définis et justifiés en termes de *droits, de devoirs et d'obligations*. Par « caractéristiquement conséquentialiste », on entend tous les jugements qui sont définis et justifiés en termes de *calcul coût-bénéfice*. Dans aucun cas réfère-t-on spécifiquement à des théories ou écoles de pensées. On comprendra ici que la terminologie de Greene est non standard (Greene 2014, 699–700).

tendance à donner des réponses caractéristiquement *conséquentialistes* au « cas du levier » (*switch case*), mais à donner des réponses caractéristiquement *déontologiques* (*non conséquentialistes*) au « cas de la passerelle » (*footbridge case*) (Greene 2014, 700). Selon Greene, cette différence, culminant dans « l'effet levier-passerelle » (*switch-footbridge effect*), peut être expliquée par la combinaison de deux facteurs : la doctrine du double effet³⁵ et le paramètre de la « force personnelle » (*harm as personal force*) (Greene 2014, 709; Greene 2017, 74–75).

Afin de bien jauger l'impact de la *force personnelle*, l'équipe de Greene a proposé (subséquentement) des variantes à la version standard du « cas de la passerelle » (*footbridge case*), nommément la « passerelle et perche » (*footbridge pole*), la « passerelle et levier » (*footbridge switch*) et la « passerelle distante » (*remote footbridge*). Dans la « passerelle et perche », le sujet pousse la victime avec une perche; dans la « passerelle et levier », le sujet se trouve sur la passerelle et déplace un levier, lequel active une trappe située sous les pieds de la victime; dans la « passerelle distante », le sujet se trouve loin de la passerelle et déplace un levier, lequel active une trappe située sous les pieds de la victime. Dans chaque cas, la proportion de sujets ayant approuvé de l'action suggérée est la suivante : « passerelle » (31%), « passerelle et perche » (33%), « passerelle et levier » (59%) et « passerelle distante » (63%).

À la lecture de ces résultats, on constate que l'effet engendré par le paramètre de la « force personnelle » est saillant. La différence entre les deux premiers cas et les deux derniers cas montre bien que les sujets préfèrent ne pas exercer une force physique sur la victime. Par ailleurs, on constate que le *contact physique* et la *distance spatiale* n'ont qu'un impact mineur (Greene 2014, 709–710; Greene 2016, 183).

À l'aide de l'imagerie par résonance magnétique, Greene et son équipe ont été en mesure d'identifier les régions cérébrales préférentiellement impliquées dans les dilemmes du « levier » et

³⁵ Conformément à la « doctrine du double effet » (DDE), Greene observe que les sujets expérimentaux désapprouvent davantage d'un mal s'il s'agit d'un moyen pour réaliser un plus grand bien, que s'il s'agit d'un effet secondaire (imprévu) pour réaliser un même bien. Dans le « cas de la passerelle », la mort de l'homme obèse est causalement nécessaire pour briser la course du train tandis que, dans le « cas du levier », la mort de l'homme sur les rails constitue uniquement un dommage collatéral. La portée de la « doctrine du double effet » est plus faible que celui de la « force personnelle ». Or, il semblerait que la combinaison de ces deux facteurs ait un effet amplificateur/multiplicatif dans certains scénarios (Greene 2008, 105–17; Greene 2014, 719–23; Greene 2016, 176–77, 179, 186–88; Greene 2017, 67–68; Cushman 2013, 274–75, 283–84).

de la « passerelle ». D'un côté, le « cas du levier » et autres dilemmes *impersonnels*³⁶ stimulaient davantage le cortex préfrontal dorsolatéral (dlPFC) et les régions associées. D'un autre côté, le « cas de la passerelle » et autres dilemmes *personnels* stimulaient davantage l'amygdale, le cortex préfrontal médian (mPFC) et le cortex préfrontal ventromédian (vmPFC) (Greene 2014, 701; Greene 2017, 74).

Dans le premier cas, les parties du cerveau associées à la mémoire de travail (assurant la rétention temporaire et le traitement de l'information) sont activées, mais non les parties liées à l'émotion. Dans le second cas, les parties du cerveau associées à la mémoire de travail sont sous-activées, tandis que celles associées à l'émotion témoignent d'une activité importante (Greene et al. 2001, 2105–8; Sinnott-Armstrong 2006, 203–4; Levy 2006, 573–74).

À la lumière de ces résultats, on constate une association entre les intuitions conséquentialistes et les régions du cerveau liées à la cognition/délibération (Système II) ainsi qu'une association entre les intuitions déontologiques et les régions du cerveau associées aux émotions (Système I).

Hypothèse du double processus : quel soutien empirique?

Une batterie d'études menées indépendamment soutiennent le modèle de Greene. Notamment, l'équipe de Mario Mendez a constaté que les patients souffrant de démence fronto-temporale sont plus à même de donner une réponse conséquentialiste au « dilemme de la passerelle » (Greene 2014, 701). De son côté, Michael Koenigs a observé que les sujets ayant des lésions au cortex préfrontal ventromédian formulaient des réponses anormalement conséquentialistes aux dilemmes moraux personnels, comme le « dilemme de la passerelle » (Koenigs et al. 2007, 908–11). Dans une autre étude, Koenigs a pu établir que les psychopathes, des individus connus pour leurs déficiences sociales et émotionnelles, ont également des affinités pour le conséquentialisme (Greene 2014, 702). Il en va de même pour les patients souffrant d'alexithymie (Greene 2014, 703). Dans les cas énoncés ci-haut, on comprendra que les agents manifestent des préférences utilitaristes, parce que leur fonction socio-émotionnelle est inhibée, et

³⁶ La dichotomie *impersonnel* vs. *personnel* se retrouve surtout dans les premiers articles de Greene. Si ce schème conceptuel n'est plus favorisé par l'auteur, il demeure tout de même possible de l'employer à des fins didactiques ou explicatives (Greene 2014, 700–1).

qu'ils ne sont pas normalement affectés par le paramètre de la « force personnelle ». Cela dit, il existe des situations où cette préférence utilitariste ne relève pas d'une déficience émotionnelle, mais d'un souci accru pour la maximisation de l'utilité globale (ou bien-être général). En effet, une récente étude a montré que les individus ayant des capacités de raisonnement supérieures ont davantage tendance à fournir des réponses utilitaristes aux dilemmes sacrificiels (Patil et al. 2021, 451). Adoptant un style proprement « cognitif », ces individus réfléchissent préférentiellement en termes de coûts-bénéfices, ce pourquoi ils seraient en mesure de surmonter leurs inclinations déontologiques initiales³⁷. Les auteurs de cette même étude ont également observé une corrélation entre l'épaisseur du cortex dorsolatéral préfrontal droit (dlPFC-d) et la fréquence des jugements utilitaristes émis. Un tel résultat est fort heureux, sachant que le cortex dorsolatéral préfrontal droit (dlPFC-d) est associé au raisonnement, et que le raisonnement est associé à l'utilitarisme (Patil et al. 2021, 449–50). De toute évidence, il existe un volume déjà impressionnant, et certes croissant d'études, comme celles présentées ici, qui confirment l'hypothèse du processus double et l'existence d'une « tension » entre les systèmes moraux (Greene 2014, 701–6).

Affect et cognition : une fausse dichotomie?

Si la proposition de Greene est fort bien étayée, des chercheurs comme Railton et Woodward considèrent toutefois que le modèle suggéré est trop simpliste, et aussi trop pessimiste. Notamment, ces auteurs contestent la dichotomie « affect-cognition », laquelle ne rendrait pas justice à la complexité des processus impliqués dans le jugement moral. Railton soutient que le système affectif n'est pas simplement un groupement de fonctions primitives, automatiques et inflexibles. Au contraire, ce système serait capable d'effectuer des tâches statistiques, d'évaluer dynamiquement les contraintes situationnelles et de guider le comportement de manière éclairée (Railton 2014, 833). Émulant la méthode d'inférence bayésienne, le système affectif effectue prédictions et projections, facilitant ainsi la planification et la prise de décision. Conçu comme l'extension des fonctions représentationnelles et cognitives, cet « ordinateur biologique » permettrait à l'agent de s'ajuster à son environnement physique et social (Railton 2014, 835–47). Même s'il procède de manière tacite et pré-délibérative, ce système permettrait au sujet de traiter

³⁷ On notera toutefois que le raisonnement *arithmétique* prédit exclusivement les jugements conséquentialistes, mais que le raisonnement *logique* prédit à la fois les jugements conséquentialistes et déontologiques (Patil et al. 2021, 454–55).

efficacement les données reçues et de répondre aux exigences du quotidien, sans avoir à encourir des coûts cognitifs démesurés. Décidément, on aurait tort de croire que les fonctions affectives constituent un handicap pour le raisonnement moral. À l'inverse, elles seraient une composante essentielle de notre vie morale, et participeraient activement à la prise de décision (Railton 2017, 172–90).

Woodward est d'un avis similaire. S'opposant à la dichotomie « émotion-raison », Woodward stipule, d'une part, que les processus affectifs et rationnels ne sont pas strictement distincts et, d'autre part, que les émotions contribuent positivement au jugement moral (Woodward 2016, 87–88). S'appuyant sur des recherches récentes, Woodward affirme que l'insula, l'amygdale, le cortex orbitofrontal (OFC) et le cortex préfrontal ventromédian (vmPFC), des structures typiquement associées aux émotions, sont impliqués dans des tâches cognitivement sophistiquées (Woodward 2016, 88–89). Plus particulièrement, ces régions s'occupent du traitement des données relatives aux renforçateurs positifs (récompenses) et négatifs (punitions). En ce sens, elles opèrent à partir de représentations binaires de la valeur. Employant un langage probabiliste, elles anticipent et prédisent l'arrivée de renforçateurs, chose particulièrement utile dans des contextes où les variables sont nombreuses et l'incertitude, prégnante (Woodward 2016, 92). Hautement flexibles, ces mécanismes ajustent et réactualisent leurs prédictions en fonction de changements observés dans l'environnement (lesquels engendrent des signaux d'erreur) (Woodward 2016, 93). Capables de traitement informationnel, de calcul probabiliste et d'apprentissage (par renforcement), les quatre structures cérébrales étudiées ici peuvent difficilement être considérées comme non cognitives, affirme Woodward (2016, 97). Évidemment, le chercheur reconnaît que ces régions ne sont pas impliquées dans des tâches conscientes comme la manipulation de propositions logiques ou mathématiques. Cela va de soi. Il n'empêche que la frontière entre affect et rationalité n'est peut-être pas aussi étanche qu'on le croyait auparavant, suppose-t-il (Woodward 2016, 99).

Mobilisant de récentes publications, Woodward suggère que certaines régions cérébrales traditionnellement associées aux émotions, comme l'OFC et le vmPFC, seraient bien plus que des sources de contamination morale. Impliquées dans la représentation de la valeur et des probabilités, ces régions contribueraient à informer, non seulement les jugements déontologiques, mais aussi les jugements conséquentialistes (Woodward 2016, 100–5). En ce sens, la différence entre les jugements déontologiques et conséquentialistes ne résiderait pas dans l'activation de l'OFC et du

vmPFC, mais bien dans l'activation complémentaire du dlPFC pour les jugements conséquentialistes (Woodward 2016, 102). Et donc, les jugements normativement « supérieurs » impliqueraient à la fois les régions dites « émotionnelles » et les régions dites « cognitives », comme le dlPFC. Cette proposition est appuyée par des témoignages empiriques voulant que les traits antisociaux prédisent des jugements conséquentialistes, mais pas des jugements normativement « supérieurs » (Bartels et Pizarro 2011, 154–61).

Réponse de Greene : l'apprentissage « sans modèle » et l'apprentissage « avec modèle »

Soucieux de répondre à ses critiques, Greene proposera certains ajouts et amendements à son modèle, lesquels sont inspirés de la théorie computationnelle (Cushman 2013, 273–92; Greene 2018, 4). Tablant sur la distinction entre apprentissage « sans modèle » et « avec modèle », le chercheur tentera d'expliquer la faillibilité de nos intuitions morales. D'une part, l'apprentissage « sans modèle » (*model-free learning*) désigne un processus par lequel l'agent assigne des valeurs positives ou négatives aux actions, tout dépendamment de la récompense obtenue. Il s'agit ici d'apprentissage par renforcement. D'autre part, l'apprentissage « avec modèle » (*model-based learning*) implique la cueillette active d'informations pour bâtir un modèle causal de l'environnement, lequel est consciemment accessible et modifiable. Ce type d'apprentissage est associé au raisonnement et à la planification. Ceci étant dit, Greene prend le soin de souligner qu'aucune de ces stratégies n'est inhéremment supérieure à l'autre (Greene 2017, 69). L'apprentissage « sans modèle » a l'avantage d'être *peu dispendieux* sur le plan computationnel, mais a le désavantage d'être *inflexible*. L'apprentissage « avec modèle », étant complexe et sophistiqué, est beaucoup *plus dispendieux*, mais il a l'avantage d'être *flexible*. Dans le premier cas, l'agent procède par automatisme, ce pourquoi il est incapable d'ajuster son comportement de manière prospective. Ceci peut entraîner des écarts entre l'action effectuée et l'action exigée par la situation. Dans le second cas, l'agent procède de manière délibérative, ce pourquoi il est capable de réviser activement ses croyances et de mettre en œuvre des stratégies conformes aux exigences contextuelles. Or, les manipulations requises sont laborieuses, et chaque opération cognitive représente un échec potentiel (Greene 2017, 70). Brièvement, on se rappellera que l'apprentissage « sans modèle » est *efficace*, mais il demeure susceptible à une quantité d'erreurs, lesquelles proviendraient de causes internes et externes, comme un entraînement inadéquat ou des changements environnementaux imprévus. À l'inverse, l'apprentissage « avec modèle » est

inefficace, mais il offre toutes les prérogatives d'une délibération explicite et consciente, laquelle assure un réaligement dynamique et intentionné du sujet avec son environnement. Évidemment, un tel exercice nécessite un déploiement d'efforts soutenu, et ouvre la porte à des erreurs procédurales³⁸.

Maintenant, on comprendra que Greene associe l'apprentissage « sans modèle » aux intuitions déontologiques/arétaïques et l'apprentissage « avec modèle » au raisonnement conséquentialiste (Greene 2017, 74–76). Selon le philosophe, l'apprentissage « sans modèle » est adéquat pour résoudre les problèmes du quotidien, mais non pour dénouer des apories proprement philosophiques, ou pour régler les différends moraux entre groupes. Heureusement, l'apprentissage « avec modèle » fournit les ressources et outils nécessaires pour solutionner des enjeux complexes. Ultime arbitre de la moralité, le conséquentialisme permettrait de dépasser les contradictions morales, en proposant une solution objective, ancrée dans une analyse des coûts et bénéfices. En procédant ainsi, il serait possible de pallier les insuffisances du déontologisme, lequel repose sur des dispositions affectives et des règles heuristiques faillibles. Par-delà ces considérations normatives, Greene se prononce également sur la question de la méthodologie en morale. Militant pour l'adoption d'un équilibre réflexif large et double (*double-wide reflective equilibrium*), le philosophe considère que la théorie morale doit non seulement viser une harmonie entre les principes généraux et les intuitions particulières, mais aussi comprendre les limitations des processus cognitifs (psychologiques/biologiques) qui produisent ces jugements (Greene 2014, 726; Greene 2017; 74). Sans surprise, Greene considère qu'une telle méthodologie devrait favoriser le conséquentialisme, puisque cette théorie morale neutralise, ou corrige, de potentiels biais cognitifs et effets heuristiques qui pourraient s'immiscer dans le processus délibératif.

Berker, Sauer et Königs sur l'argument des facteurs moralement non pertinents

Si les arguments pour l'hypothèse du double processus sont désormais légion, certains auteurs suggèrent que le modèle mérite toutefois quelque révision (De Neys et Bialek 2017, 123–36; De Neys 2018; Sauer 2018; Evans 2019, 383–415). Ces critiques ne seront pas adressées ici,

³⁸ Intéressamment, Greene présente l'apprentissage « sans modèle » comme un processus rigide et inflexible, tandis que Railton et Woodward mettent l'accent sur l'adaptabilité des fonctions intuitives, lesquelles seraient conçues, précisément, pour un monde traversé par l'incertitude. Dans ce cas précis, il semble y avoir une coupure conceptuelle entre les auteurs. Or, il ne pourrait s'agir que d'une différence emphatique.

car elles ne réfutent pas explicitement la thèse que le système affectif est impliqué dans les jugements moraux, surtout déontologiques. Par contre, il sera nécessaire de confronter les arguments de Selim Berker (2009), Hanno Sauer (2012) et Peter Königs (2018), lesquels doutent de la valence normative des résultats neuroscientifiques présentés ici.

Pour commencer, Berker souligne que les arguments de Greene contre le déontologisme ne tiennent pas la route (Berker 2009, 314–29). Écartant rapidement trois arguments auxiliaires basés sur les émotions, les heuristiques et l'évolution, Berker se concentre sur l'argument des facteurs moralement non pertinents (*argument from morally irrelevant factors*), voulant que les intuitions déontologiques soient sensibles à des « facteurs personnels » et donc, « non pertinents ». D'une part, Berker souligne que cet argument repose sur une identification arbitraire entre déontologie et « dilemmes personnels », conséquentialisme et « dilemmes impersonnels³⁹ ». D'autre part, l'auteur affirme que cet argument repose sur un postulat suspect, celui que les « facteurs personnels » sont moralement non pertinents. Dans tous les cas, Berker prétend que la neuroscience n'est d'aucun apport épistémologique. Évidemment, Greene s'oppose à un tel discours. Il soutient que la neuroscience permet d'établir l'origine physiologique des jugements et, par le fait même, d'établir si des processus non fiables sont impliqués la formation desdits jugements. À cet égard, Greene réitère que les émotions, heuristiques et structures cérébrales primitives ont une portée normative (Greene 2014, 706, 711)⁴⁰.

Suivant les traces de Berker, le philosophe Hanno Sauer critique l'argument des facteurs moralement non pertinents. Sauer formule l'argument ainsi :

P1 : Les jugements déontologiques proviennent de processus émotionnels. (Prémisse empirique)

³⁹ L'article de Berker a été publié en 2009, époque à laquelle Greene soutenait encore la distinction entre *facteurs personnels* et *impersonnels*. Or, Greene ne soutient plus cette distinction. Afin de mettre l'argument à jour, il faudrait remplacer le prédicat « personnel » par le prédicat « force personnelle ». Plus généralement, on soulignera que la critique de Berker ne tient pas compte des développements les plus récents dans le domaine (Berker 2009, 321–26). Ce point s'applique aussi à la critique de Sauer plus bas.

⁴⁰ En guise de rappel, on soulignera que la critique déployée à l'occasion de ce mémoire s'attaque à l'intuitionnisme rationaliste de Huemer, lequel postule des vérités morales synthétiques *a priori* (analogues aux vérités mathématiques ou logiques). Ceci signifie qu'il est suffisant de montrer que les émotions et heuristiques sous-tendent une majorité d'intuitions pour défaire la thèse rationaliste de Huemer.

P2 : Les processus émotionnels répondent à des facteurs moralement non pertinents.
(Prémisse normative)

C1 : Les intuitions déontologiques et les jugements qui en découlent n'ont aucune force normative authentique. (Sauer 2012, 784)

Sauer tente de montrer que les deux prémisses ne sont pas fondées. Ici, nous ne discuterons pas des résultats empiriques et de l'interprétation des données. (Il s'agit d'un débat complexe, lequel sort du cadre philosophique.) Cela dit, nous aborderons la prémisse normative, et le traitement qu'en fait Sauer (2012, 794–805). L'auteur procède en deux temps : dans un premier temps, il argumente que le critère de la « force personnelle » n'explique pas les réponses différenciées des sujets expérimentaux (Sauer 2012, 794–99) et, dans un second temps, que les jugements moraux proviennent d'heuristiques automatiques, certes, mais corrigibles par le biais de la réflexion (Sauer 2012, 799–805).

Commençons par le critère de la « force personnelle » (Sauer 2012, 797). Sauer suggère que ce critère n'est pas adéquat, car il laisse certains résultats empiriques inexplicés. Invoquant une récente étude de Nichols et Mallon, dans laquelle des tasses de thé sont projetées sur d'autres tasses de thé, Sauer affirme que deux conditions pour l'adoption du critère de la « force personnelle » ne sont pas remplies : d'une part, ces expériences excluent toute interaction physique « agent-agent » et, d'autre part, ces expériences excluent toute application *directe* de force physique. Or, comment expliquer le fait que les sujets expérimentaux répondent d'une manière analogue aux expériences impliquant des humains et aux expériences impliquant des tasses de thé? En l'occurrence, il semble que le critère de la « force personnelle » ne puisse expliquer les résultats observés par Nichols et Mallon (Sauer 2012, 797).

Passons maintenant à l'enjeu des règles heuristiques (Sauer 2012, 800–5). Sauer soutient que nos jugements moraux sont généralement justifiés, car ils sont basés sur des règles heuristiques rapides et frugales (*fast-and-frugal*). Étant fiables dans une majorité de situations, ces règles peuvent malgré tout échouer dans des situations exceptionnelles. Selon Sauer, l'occasionnel échec des heuristiques ne serait pas imputable à la présence d'influences affectives malvenues (*contra* Greene et son *neurosentimentalisme*), mais à la finitude des ressources cognitives disponibles chez l'humain. Nonobstant leur faillibilité, ces règles acquises pourraient être amendées par des processus réflexifs d'ordre supérieur (métacognitifs), conformément à la thèse de la « pénétrabilité

cognitive ». Jouissant d'un statut indubitablement rationnel, les règles heuristiques seraient d'une grande adaptabilité, comme l'affirment Gerrans et Kennett (Sauer 2012, 803).

En écartant le facteur de la « force personnelle » et en adoptant une interprétation laudative des heuristiques, Sauer considère qu'il est capable de réfuter la thèse de Greene, selon laquelle les jugements déontologiques seraient sensibles à des facteurs moralement non pertinents. Pourtant, si l'on consulte la littérature en psychologie sociale, comme les recherches menées par Haidt, il semble difficile de souscrire à l'idée que les intuitions morales résistent à l'influence de facteurs non pertinents, comme la propreté d'un local ou l'odeur environnante. De même, Sauer défend une conception relativement heureuse des heuristiques, affirmant que celles-ci sont des processus hautement flexibles et malléables, lesquels ratent *parfois* la cible. Or, Sauer s'inspire principalement de la tradition suivant Gerd Gigerenzer, laquelle est ostensiblement plus sympathique à l'usage des heuristiques que celle d'un Daniel Kahneman. (Nous aurons l'occasion d'approfondir ces quelques enjeux dans les prochaines sections.)

De son côté, Königs discute de trois arguments mis de l'avant par Greene, nommément, l'argument de la confabulation (1), l'argument fonctionnaliste (2) et l'argument des facteurs moralement non pertinents (3) (Königs 2018, 195–209). Il affirme qu'aucun de ces arguments ne dépend de découvertes neuroscientifiques et qu'aucun d'entre eux n'est logiquement fondé. Le premier argument suppose que la déontologie n'est que confabulation, ou rationalisation *post hoc*. Le second argument comporte que la déontologie fonctionne pour les situations familières, mais non pour les problèmes non familiers, lesquels impliquent souvent l'interaction entre des groupes culturels hétérogènes. (Dans ce cas, Greene suggère que le conséquentialisme est la solution optimale.) Le troisième, déjà exploré plus haut, implique que les réponses déontologiques sont guidées par des intuitions « psychologiques » et que les réponses conséquentialistes sont guidées par des intuitions « philosophiques » (Königs 2018, 198–200). D'une part, on utilise le prédicat « psychologique » pour désigner les intuitions automatiques et émotionnelles. D'autre part, on emploie le prédicat « philosophique » pour désigner les intuitions qui ne seraient pas affectées par des facteurs contingents. Greene affirme que les intuitions « psychologiques » sont particulièrement susceptibles à l'argument des facteurs moralement non pertinents, tandis que ce ne serait pas le cas pour les intuitions « philosophiques ». Selon Königs, cette catégorisation ne fait

rien pour aider la cause de Greene⁴¹. Le critique souligne que, même si le conséquentialisme est soutenu par des intuitions « pures », nous n'avons aucune garantie que ces intuitions sont moralement pertinentes. Königs l'exprime en ces termes :

The fact that responses based on 'psychological' intuitions tend to be sensitive to irrelevant factors because they are based on 'psychological' intuitions simply does not entail that responses that are not based on 'psychological' intuitions are unlikely to be sensitive to irrelevant factors (or insensitive to irrelevant factors). For they may of course have this defect in spite of not being 'psychological' intuitions. (Königs 2018, 200)

Ce commentaire est tout à fait porteur. En effet, comment déterminer quels sont les facteurs moralement pertinents et quels sont les facteurs moralement non pertinents? Quel est notre critère? Ici, Königs résume bien l'enjeu central des trois critiques adressées à Greene : comment savoir que les processus assurant la formation des jugements moraux sont (non) fiables? Cette question sera reprise dans le troisième et dernier chapitre de ce mémoire⁴².

2.1.2 Le « Modèle intuitionniste social » de Jonathan Haidt

Dans leur article intitulé « Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism », Jesse Graham et Jonathan Haidt s'interrogent sur le contenu et les origines des intuitions morales (Graham et al. 2013, 55–130). Afin d'éclairer la nature de ce phénomène, Graham et Haidt ont proposé la « Théorie des fondations morales », laquelle repose sur quatre propositions centrales. Ces quatre hypothèses sont le nativisme (le cerveau est moralement prédisposé), l'apprentissage culturel (la culture détermine les spécificités du cerveau moral), l'intuitionnisme (la primauté des intuitions sur la raison stratégique) et le pluralisme (la pluralité des fondations morales) (Graham et al. 2013, 61–70; Graham et al. 2018, 211–22). À l'occasion, nous pourrions approfondir le troisième élément, lequel concerne le « Modèle intuitionniste social » (*Social Intuitionist Model*).

S'inscrivant à l'intersection de plusieurs courants en psychologie, comme la « révolution affective » et la « révolution de l'automatisme », le modèle intuitionniste social comporte deux

⁴¹ Königs étaye cet argument dans une publication subséquente (Königs 2020, 2609).

⁴² Évidemment, des facteurs comme les effets de cadrage (*framing effects*) ou autres éléments externes de la sorte sont non pertinents sur le plan moral (Sinnott-Armstrong 2006, 352–56; 2006, 204–7). Cela ne fait pas de doute.

clauses principales (Haidt et Bjorklund 2008, 181–88). D’une part, les *croyances morales* proviennent d’un groupe restreint d’*intuitions* innées, lesquelles sont chargées d’une valence affective. D’autre part, le *jugement moral* provient (d’abord) d’*intuitions* rapides et automatiques, lesquelles sont (ensuite) accompagnées du *raisonnement*, un processus lent et délibératif (Haidt et Bjorklund 2008, 181). Bâtissant sur ces deux hypothèses, le modèle est composé de six liens : le lien du jugement intuitif (1), le lien du raisonnement *post hoc* (2), le lien de la persuasion raisonnée (3), le lien de la persuasion sociale (4), le lien du jugement raisonné (5) et le lien de la réflexion privée (6). Les deux premiers liens occuperont la majorité de notre attention.

Pierre d’assise du modèle intuitionniste social, le lien du jugement intuitif (1) veut que les jugements moraux soient des produits de l’intuition. Haidt décrit l’intuition ainsi :

“Moral intuition” is defined as the sudden appearance in consciousness, or at the fringe of consciousness, of an evaluative feeling (like–dislike, good–bad) about the character or actions of a person, without any conscious awareness of having gone through steps of search, weighing evidence, or inferring a conclusion. (Haidt et Bjorklund 2008, 188)

À la lecture de ce passage, on notera que les intuitions sont engendrées par des processus heuristiques, typiquement associés au Système I. Comme les intuitions surgissent avant le raisonnement, on comprendra que le Système I a priorité sur le Système II. Cela dit, le Système II contribue activement à la délibération morale, surtout lorsqu’il s’agit de fournir des raisons explicites pour lesdits jugements. En effet, le lien du raisonnement *post hoc* (2) stipule que la fonction principale du Système II est de rechercher des arguments pour expliquer et justifier un jugement déjà formé (Haidt et Hersh 2001, 193; Haidt et Bjorklund 2008, 189). Également chargé de la persuasion raisonnée (3), le Système II servirait, non pas des desseins scientifiques ou philosophiques, mais bien rhétoriques et sophistiques. En ce sens, la fonction du raisonnement verbal explicite ne serait pas d’atteindre la vérité morale, mais bien de manipuler l’opinion d’autrui, de défendre sa réputation ou de générer l’assentiment (Haidt et Bjorklund 2008, 190–92). Or, toute persuasion n’implique pas nécessairement l’argumentation ou le raisonnement. S’appuyant sur une variété d’études en psychologie et en sociologie, Haidt soutient que la « persuasion » peut aussi s’effectuer de manière inconsciente. À maintes reprises, il a été observé que les humains modulent leur comportement et leurs croyances de manière à se conformer à leurs pairs. Éprouvant un fort besoin d’appartenance, ces animaux ultrasociaux préféreraient la cohésion à la dissension. Évidemment, cette tendance à l’harmonisation traverse aussi la sphère morale. Ce phénomène est

symbolisé par le lien de la persuasion sociale (4). Constituant le cœur du modèle intuitionniste social, ces quatre premiers liens soulignent le rôle éminemment social du raisonnement moral. Cependant, on aurait tort de croire que la moralité n'est qu'un truchement pour la socialité. En effet, il existe deux liens additionnels, lesquels capturent un type de raisonnement proprement philosophique. D'une part, le lien du jugement raisonné (5) suppose qu'il est possible de dominer nos intuitions initiales, par la voie de la logique. D'autre part, le lien de la réflexion privée (6) met l'accent sur le dialogue interne, lequel permet à l'agent de pondérer des intuitions conflictuelles. Ces deux éléments dépendent du Système II. Voilà, en bloc, les six liens qui forment le modèle intuitionniste social (Haidt et Bjorklund 2008, 193–96).

« Perplexité » et confabulation morales

Durant le nombre d'années qui ont permis le développement du modèle, Haidt et son équipe ont eu l'opportunité de mener une batterie d'expériences, afin de vérifier l'hypothèse d'une primauté de l'intuition. Deux phénomènes seront soulignés ici : la « perplexité » (*moral dumbfounding*) et la « confabulation » (*moral confabulation*). L'existence de la « perplexité morale » comme phénomène *sui generis* a été observée dans plusieurs expériences. À des fins illustratives, on rapportera l'exemple de l'inceste, lequel est certainement le plus discuté d'entre tous. Ci-bas, est présentée la vignette distribuée aux sujets expérimentaux :

Julie and Mark. Julie and Mark are brother and sister. They are traveling together in France on summer vacation from college. One night they are staying alone in a cabin near the beach. They decide that it would be interesting and fun if they tried making love. At the very least it would be a new experience for each of them. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy making love, but they decide not to do it again. They keep that night as a special secret, which makes them feel even closer to each other. What do you think about that, was it OK for them to make love? (Haidt 2001, 814)

Confrontés à cette problématique, les participants se positionnent, en forte majorité, contre le comportement des protagonistes. Or, lorsque l'on demande aux sujets de justifier leur prise de position, ceux-ci sont incapables d'articuler des raisons cohérentes. Peut-être évoquent-ils les dangers liés à la consanguinité? Peut-être s'inquiètent-ils de l'intégrité psychologique des personnages? Éventuellement, ils sont contraints d'admettre qu'aucune des raisons fournies ne cadre avec les détails de l'histoire (comme l'usage de contraceptifs, la formulation d'un

consentement explicite et le plaisir encouru). Incapables de défendre leurs convictions, les participants sont « perplexes » (*dumbfounded*) (Brink 2014, 684–688; Railton 2014, 831–32). Ici, deux éléments méritent d’être soulignés : d’une part, les sujets sont incapables d’accéder aux réelles motivations qui sous-tendent leur désapprobation et, d’autre part, les sujets ne révisent pas leurs intuitions spontanées, même s’ils reconnaissent l’innocuité du geste décrit. Comme elles résistent à la « contre-évidence », ces intuitions sont dites robustes, ou « récalcitrantes^{43,44} ».

Multipliant les études empiriques, Haidt observera la « perplexité morale » dans une variété de scénarios impliquant du cannibalisme, des formes de masturbation atypiques ou des pratiques sexuelles controversées (sexe anal homosexuel ou sexe oral lesbien) (Haidt et Hershey 2001, 191–221). Suscitant des réactions affectives fortes, le plus souvent ancrées dans le dégoût, ces expériences démontrent la prépondérance du Système I dans la formation des jugements moraux, surtout lorsque ceux-ci évoquent des comportements tabous (Haidt et Bjorklund 2008, 196–98).

Comme on l’a indiqué plus haut, la « perplexité morale » s’accompagne généralement de confabulation morale, ou rationalisation *post hoc*. Dans une étude largement citée, Wheatley et Haidt soumettent leurs sujets expérimentaux à une suggestion post-hypnotique, laquelle provoque un sentiment soudain de dégoût à la lecture d’un mot particulier (Wheatley et Haidt 2005, 780–84; Haidt 2008, 198–99). Parcourant ensuite six histoires moralement chargées, les sujets hypnotisés formulaient généralement des jugements plus sévères, ou moins favorables, lorsqu’ils rencontraient le mot pertinent durant leur lecture. Une septième histoire contenant le mot concerné, mais aucun méfait, a été incluse dans une version subséquente de l’étude. Curieusement, certains sujets condamnaient les actions du personnage, même si ce dernier n’était coupable d’aucun tort. Interrogés sur leur réponse, les sujets inventaient des motifs, quelque peu farfelus, pour coincer le protagoniste. C’est là un exemple de rationalisation *post hoc*. Dans une autre étude révélatrice, il a été établi que le niveau de propreté de l’environnement immédiat affecte la gravité des jugements.

⁴³ Des auteurs comme Brink et Railton contestent toutefois l’idée que l’inceste, dans le cas de Julie et Mark, est réellement sans dangers psychologiques (Brink 2014, 684–88; Railton 2014, 831–33).

⁴⁴ Dans une récente publication, Royzman, Kim et Leeman (2015, 296–313) présentent les résultats de trois études empiriques, lesquelles jettent le doute sur les conclusions de Haidt. Plus spécifiquement, Royzman et son équipe soutiennent que l’effet de la perplexité morale, observé par Haidt, ne trouverait pas son origine dans l’irrationalité des sujets expérimentaux. Au contraire, l’effet observé serait imputable (principalement) à des vices méthodologiques. Or, on soulignera que l’article de Royzman concerne uniquement le cas de Julie et Mark. On se gardera donc de généraliser, d’autant plus qu’il existe une large quantité de publications attestant de l’existence de la « perplexité morale » comme phénomène *sui generis*.

Plus précisément, les sujets témoignant d'une attention particulière à leurs sensations corporelles (*private body consciousness*) formulaient des jugements plus sévères lorsqu'ils étaient assis dans un local malpropre, sinon dégoûtant (Schnall, Haidt et Clore 2008, 1100). Bien que cet exemple ne comporte aucun raisonnement *post hoc*, il montre tout de même que le jugement moral est affecté par des facteurs strictement non pertinents (Haidt et Bjorklund 2008, 199). Similairement, deux études relativement récentes ont démontré que le simple fait de se laver les mains ou d'être proche d'une bouteille de gel hydroalcoolique peut altérer le jugement moral (Haidt 2013, 288; May et Kumar 2020, 146–47). Autant pour l'objectivité de la morale!

Comme le soulignent Haidt et Bjorklund, des exemples saillants de rationalisation *post hoc* peuvent aussi être retrouvés chez Gazzaniga, dans son répertoire d'études sur le cerveau scindé. (Le syndrome du cerveau scindé comporte que les deux hémisphères sont déconnectés et ne peuvent communiquer.) Dans le cas présenté, le patient effectue un geste, comme se lever ou marcher, lequel relève de l'hémisphère droit. Au lieu de s'interroger sur le comportement adopté, l'hémisphère gauche, chargé du langage et de la verbalisation, tente immédiatement de fournir des justifications. Il en arrive à la conclusion que le patient s'est levé, *car* celui-ci avait soif. Gazzaniga considère qu'un tel commentaire *post hoc* est le produit de régions cérébrales formant le « module d'interprétation ». Ce module aurait justement pour fonction d'expliquer le comportement du sujet, à lui-même et aux autres (Haidt et Bjorklund 2008, 190).

Dans une étude tout aussi révélatrice, Johansson et ses collègues ont montré l'aisance avec laquelle il était possible de manipuler un public et de susciter le réflexe confabulatoire. En l'occurrence, les sujets expérimentaux devaient choisir lequel de deux visages féminins ils trouvaient le plus attirant. Pour certains participants, l'expérimentateur usait d'habileté pour remplacer l'image choisie par l'image rejetée. Ensuite, on demandait aux sujets concernés d'expliquer pourquoi ils avaient choisi l'image présentée (c'est-à-dire l'image rejetée). Bien souvent, les sujets dupés fournissaient des raisons qui n'étaient pas cohérentes avec leur choix original. Par exemple, l'un d'eux a justifié son choix en évoquant spécifiquement les boucles d'oreille portées par la femme. Or, l'image sélectionnée originalement n'affichait aucune boucle d'oreille. Si toutes les confabulations n'étaient pas aussi curieuses, elles témoignent néanmoins de l'hyperactivité et de l'inventivité du « module d'interprétation », lequel a été mentionné plus haut (dans Haidt et Kesebir 2010, 811).

Plus récemment, l'équipe de Fiery Cushman s'est penchée sur la distinction entre les principes moraux endossés (implicitement) et les principes moraux articulés (explicitement). Dans leur étude, les sujets expérimentaux étaient confrontés à plusieurs dilemmes moraux, auxquels ils devaient répondre. Ces dilemmes ont été choisis pour tester trois principes : le contraste entre action et omission (1), la doctrine du double effet (2) et le contact physique comme facteur aggravant (3). Il a été observé que les sujets expérimentaux souscrivaient tacitement à ces trois principes, mais qu'ils étaient seulement capables d'articuler le premier et le second. Bien que les réponses formulées par les participants étaient conformes à la doctrine du double effet, ceux-ci étaient incapables de reconnaître que la distinction entre « moyens » et « effets collatéraux » jouait un rôle prépondérant dans la formation de leur jugement. Au contraire, les sujets invoquaient des explications alternatives (29%), manifestaient de l'incertitude quant à la justification émise (22%), niaient que les cas présentés comportaient des différences morales pertinentes (17%) ou formulaient des justifications strictement inadéquates (16%) (Cushman, Young et Hauser 2006, 1084–88). En l'occurrence, on notera que les « explications alternatives » sont des instances de rationalisation *post hoc*, émergeant d'une incapacité à fournir une justification adéquate. Cette expérimentation confirme l'hypothèse que le jugement moral est un processus mobilisant deux systèmes : l'un conscient et l'autre inconscient (Cushman, Young et Hauser 2006, 1086–88).

Dans une volonté d'approfondissement, l'équipe de Marc Hauser a déployé une étude portant spécifiquement sur la doctrine du double effet (Hauser et al. 2007, 16–18). Ils ont fait deux trouvailles principales : tous les groupes démographiques ciblés effectuaient des jugements moraux conformes à la doctrine du double effet (1) et la majorité des individus interrogés étaient incapables de fournir la justification pertinente pour leurs jugements moraux (2). Encore une fois, les chercheurs ont observé une dissociation entre les jugements effectués et les justifications formulées, entre les *principes opérants* et les *principes exprimés* (Hauser et al. 2007, 17–18). Ces conclusions soutiennent directement le modèle intuitionniste social, voulant que l'intuition soit prioritaire, et le raisonnement stratégique, secondaire (Haidt 2013, 286–89).

Cela dit, il ne faudrait pas croire que le Système I est autosuffisant et le Système II, dispensable. En effet, les deux systèmes travaillent de concert pour aiguillonner le comportement. Cette interdépendance est soulignée par Cushman dans une récente publication sur la rationalisation *post hoc*. S'appuyant sur la théorie de « l'échange représentationnel », Cushman

montre que la rationalisation assure un transfert d'informations entre les différents processus, conscients et inconscients, qui déterminent le comportement. Chargé d'extraire et de traduire des données générées par les systèmes non rationnels, comme des croyances ou des désirs, ce mécanisme fabrique des fictions utiles, lesquelles peuvent ensuite être interprétées et employées par les systèmes rationnels. Ceci permet à l'agent de faire sens de son propre comportement et de celui des autres (Cushman 2020, 1–16). Mais encore, on rajoutera que la rationalisation peut contribuer à la régulation des émotions dans un contexte social (Railton 2020, 31–32). Ainsi décrit, le processus de rationalisation s'aligne avec le modèle intuitionniste social de même qu'une conception fonctionnaliste de la moralité.

Par-delà l'enjeu de la rationalisation comme tel, Haidt et Kesebir mettent l'accent sur le rôle social de la moralité. Proposant une analyse sociale-fonctionnaliste, les auteurs soulignent d'abord l'importance de mécanismes évolutionnaires, comme la sélection de parentèle et l'altruisme réciproque, dans l'émergence d'une proto-moralité (Haidt et Kesebir 2010, 808–9). Reconnaisant que ces deux facteurs n'expliquent pas la totalité du phénomène, les chercheurs ont identifié deux facteurs additionnels qui auraient permis la coopération à grande échelle : l'impératif de défendre sa réputation devant celle des autres et la nécessité de punir les profiteurs (*free-riders*). D'une part, les recherches sur l'hypocrisie et la charité ont montré que l'intuition morale sert, avant tout, à protéger son image. D'autre part, l'intuition permettrait de pister les méfaits d'autrui et de détecter les escrocs plus efficacement. Dans les deux cas, la confabulation serait omniprésente, que ce soit pour défendre ou accuser (Haidt et Kesebir 2010, 808–14). Brièvement, on dira que ces mécanismes de contrôle et de régulation rendent possible la coopération entre compétiteurs.

Maintenant, le modèle intuitionniste social n'est pas sans critiques. La plupart reprochent à Haidt de reléguer la raison (morale) au rang d'épiphénomène ou de peindre cette raison sous un mauvais jour (Pizarro et Bloom 2003, 193–96; Jacobson 2008, 219–32; Narvaez 2008, 233–40; Haidt et Kesebir 2010, 807–8; Gibbs 2019, 37–39). Haidt se défend toutefois du cynisme qu'on lui affuble. Il réitère que les intuitions sont primordiales, mais non toutes-puissantes (*intuitive primacy, but not dictatorship*) (Haidt et Kesebir 2010, 807). Haidt rappelle que la raison participe à quatre des six liens du modèle intuitionniste social, et que les deux derniers liens, celui du jugement raisonné et celui de la réflexion privée, accordent une place de choix à la faculté rationnelle. En l'occurrence, la raison n'est plus simplement instrumentale ou socialement fonctionnelle, elle est

discursive, elle contribue au perfectionnement de soi. On reconnaîtra toutefois que ces deux derniers liens sont plus rarement mobilisés.

Certains critiques encore plus virulents, parmi lesquels figure Joshua May, rejettent carrément la thèse voulant que « nos concepts moraux soient partiellement composés ou nécessairement dépendants de nos émotions » (May 2018, 27). May affirme que les preuves empiriques entourant le processus de moralisation, la stimulation émotionnelle (*emotional priming*), la neuropathologie et la rationalisation *post hoc* sont tout simplement trop faibles pour soutenir le sentimentalisme d'un Greene ou d'un Haidt. S'il reconnaît que les jugements moraux sont corrélés avec les émotions, May argumente que ces émotions ne sont pas des facteurs moralisants, c'est-à-dire des facteurs qui expliquent le passage du conventionnel au moral, ou le saut du moral à l'immoral (May 2018, 28–31). Plus encore, le philosophe soutient que les effets liés aux manipulations émotionnelles sont peu répandus et assez faibles (May 2018, 32–44). Citant des récentes études et méta-analyses, May affirme que les émotions incidentelles, comme le dégoût, la colère, la gaieté (*mirth*) ou le bonheur (*happiness*), ne font qu'amplifier légèrement certains jugements moraux, en changeant très rarement leur orientation (May 2018, 37–38). Cette amplification modeste, prétend May, serait liée au phénomène de « l'attribution erronée », voulant que les sujets expérimentaux attribuent leurs émotions incidentelles au jugement moral. Si cette interprétation était avérée, les émotions incidentelles ne seraient pas les causes de nos jugements moraux, mais de simples conséquences de ces mêmes jugements (May 2018, 41–44). Par ailleurs, May suppose que les travaux en neuropathologie appuient très mal la thèse sentimentaliste. Selon lui, ces études ne prouvent pas que les personnes souffrant de psychopathie ou de sociopathie acquise ont une capacité morale altérée et que cette altération est due à des déficiences émotionnelles (May 2018, 44–51). Si une telle altération devait être identifiée, elle pourrait plus facilement être expliquée par des troubles cognitifs que des troubles affectifs, suppose l'auteur (May 2018, 47–52). Se prononçant sur le thème de la rationalité, May reconnaît que le jugement moral est influencé par des processus inconscients, comme en attestent les travaux sur la « perplexité » et la « confabulation » morales. Or, il soutient que ces influences ne rendent pas causalement inerte le raisonnement conscient et contrôlé (May 2018, 74–77). Au contraire, la délibération et la réflexion actives permettraient de réviser (et non seulement de justifier) certains jugements initiaux. En effet, nombreux sont les agents moraux qui modifient leurs croyances préétablies lorsqu'ils ne sont pas soumis à des contraintes artificielles, d'ordre temporel ou

informationnel. C'est donc dire que la rationalisation *post hoc* ne serait pas la seule forme de rationalité impliquée dans le jugement moral (May 2018, 74–77). Pour toutes ces raisons, et encore, May juge que les nouveaux apôtres du sentimentalisme sont dans l'erreur.

Si l'auteur réussit à montrer que les sentimentalismes *forts* sont intenable, il échoue toutefois à montrer qu'il n'existe *aucune* relation, soit-elle causale ou constitutive, entre l'affect et le jugement moral. Pour cette raison, il ne peut réfuter les sentimentalismes *plus modestes*, comme celui défendu par Haidt, lesquels syncrétisent affect et raison. Cette fois encore, il semblerait que le modèle intuitionniste social résiste à la fronde des nouveaux apôtres du rationalisme.

2.2 « Rationalité limitée » et règles heuristiques

S'inscrivant dans un vaste programme de recherche, celui de la « rationalité limitée » (*bounded rationality*), les processus heuristiques désignent des règles simples pour le jugement ou la prise de décision (Wheeler 2020). Au sein de la tradition, ces règles heuristiques sont interprétées de deux façons : chez Amos Tversky et Daniel Kahneman, il s'agit surtout de biais cognitifs disposant à l'erreur (Kahneman, Slovic et Tversky 1982; Kahneman 2011), tandis que, chez Gerd Gigerenzer, il s'agit de procédures rapides et frugales (*fast-and-frugal*) qui peuvent, sous les bonnes conditions, produire des résultats heureux (Gigerenzer et Gaissmaier 2011, 451–82). En l'occurrence, nous nous concentrerons sur le projet de Tversky et Kahneman, car ceux-ci acceptent la dichotomie entre Système I et Système II, mais non Gigerenzer, qui la rejette vertement (Gigerenzer 2008, 15, 19). Les travaux de Gigerenzer sont d'un intérêt indiscutable, mais ils ne pourront être abordés dans le cadre de cette recherche.

2.2.1 Heuristiques et moralité : un projet au stade embryonnaire

Assis sur les épaules des grands, Cass Sunstein (2004, 1556–97; 2005, 531–42; 2007, 156–70) donnera une nouvelle impulsion à la recherche sur les biais heuristiques. Celui-ci tentera notamment de relier morale et heuristique d'une manière plus explicite. Dans une série d'articles, l'auteur montre que les agents moraux utilisent des règles simples et rapides, lesquelles peuvent toutefois rater la cible (*misfire*) (Sunstein 2005, 533). Bien que des normes déontologiques du genre « il ne faut pas mentir » ou « il ne faut pas faire du mal aux innocents » sont généralement appropriées, il n'en demeure pas moins qu'il existe des situations où celles-ci peuvent produire des

effets néfastes (Sunstein 2008, 27–30). En effet, on pourrait croire qu’il est adéquat de mentir si cela peut sauver des vies, comme le suppose Benjamin Constant dans sa réponse à Kant. Similairement, on pourrait supposer qu’il est justifié de pousser un étranger devant un trolley si cela peut préserver cinq vies. Traversant une vaste gamme d’intuitions, qu’elles soient de niveau intermédiaire ou abstrait, le phénomène de la surgénéralisation peut mener à l’application d’un principe dans une situation où celui-ci ne s’applique pas. Cet effet heuristique affecte particulièrement la déontologie kantienne ou l’utilitarisme de la règle.

À des fins illustratives, Sunstein (2005, 532–33) identifie quelques règles heuristiques, lesquelles seraient susceptibles de s’étendre au domaine moral. En l’occurrence, on retiendra l’heuristique de la disponibilité et celle de la représentation. À l’origine du problème de la surgénéralisation (Huemer 2008, 384), l’heuristique de la disponibilité comporte que les humains surévaluent la valeur épistémique des données présentes en mémoire (Tversky et Kahneman 1982, 163–78). Confrontés à des questions complexes, lesquelles impliquent souvent des probabilités, les agents vont faire appel à des illustrations immédiatement *disponibles*, faciles à interpréter, et souvent stéréotypées. De toute évidence, cette stratégie est fort mal adaptée aux situations non familières, ce pourquoi elle provoque invariablement des erreurs. Tout aussi problématique est l’heuristique de représentativité. Il s’agit là d’une tendance à juger de la probabilité d’un événement incertain (l’appartenance de X à Y) en se basant sur la similarité entre un individu particulier et la population générale (la similarité entre X et Y). Ici, l’agent considère, de manière fautive, que *représentativité* implique nécessairement *probabilité élevée*. Suivant Slovic, on rajoutera qu’il existe aussi un heuristique de l’affect, selon lequel les réponses émotionnelles des agents constituent un critère décisionnel fiable (Slovic et al. 2007, 1333–52). Dans ce cas précis, il y aurait un saut indu du subjectif à l’objectif. Un tel raccourci peut évidemment causer des erreurs sur le plan évaluatif (Railton 2014, 829–30).

Interrogeant le lien entre évolution, moralité et heuristique, Sunstein (2005, 534–35) est d’avis que les règles heuristiques morales peuvent jouer un rôle évolutif. Par exemple, l’heuristique du dégoût (Fischer 2016, 679–93) et de l’outrage comportent une utilité indiscutable lorsqu’il s’agit de condamner l’inceste ou de punir des malfaiteurs. Au sein d’une communauté primitive, ces règles confèrent des avantages patents. Or, si l’on importe ces règles dans le contexte actuel, des effets indésirables peuvent se produire (Cosmides et Tooby 2006). Notamment, l’heuristique de la

« force personnelle », identifiée par Greene, entre en contradiction avec le *desideratum* du plus grand bien. Dans une même veine, Cushman et son équipe ont montré que notre aversion naturelle à causer du tort physique se traduit même dans des situations où aucun tort n'est causé (e.g. en lançant une roche sur une main artificielle). Ici, l'heuristique « ne pas faire du mal » n'est visiblement pas adéquate (Nichols 2020, 127–28). L'idée est que ces deux règles ont été forgées dans un contexte qui n'est pas ou plus opérant, pour des raisons diverses.

Outre les règles mentionnées ci-haut, Sunstein propose quelques heuristiques morales additionnelles, comme « ne pas causer de décès humains intentionnellement », « ne pas permettre à un tiers de commettre un tort moral pour un montant d'argent (rétributif) » ou « punir et ne pas récompenser la trahison » (Sunstein 2005, 535–42; Kelman 2013, 377–80). Parmi les multiples exemples formulés par l'auteur, nul n'est plus évocateur que l'heuristique concernant le rapport de l'homme à la nature. On le résumera ainsi : « ne pas jouer à Dieu ou laisser la nature en paix » (*do not play God nor tamper with nature*) (Sunstein 2005, 539–40). Une telle heuristique possède une portée explicative impressionnante. Elle explique l'aversion ou la méfiance des humains pour les vaccins, les pesticides, la viande synthétisée, le clonage, la recherche sur les cellules souches, les organismes génétiquement modifiés (OGM), etc. Mais surtout, elle motive certaines réactions adverses à l'homosexualité ou à l'inceste. Si cette heuristique explique bien des comportements, elle ne justifie pas ces comportements pour autant. Au contraire, il serait raisonnable de croire que certaines de ses biotechnologies et pratiques sexuelles sont entièrement défendables d'un point de vue moral. Encore une fois, une heuristique *a priori* bénigne peut avoir des effets parfaitement délétères.

Maintenant, Sunstein nous rappelle que son travail reste exploratoire et que la recherche sur les heuristiques morales est encore jeune. Cette affirmation demeure vraie aujourd'hui, près de deux décennies après la publication de son article séminal. Cela dit, des chercheurs comme Walter Sinnott-Armstrong et S. Matthew Liao se sont récemment prononcés sur le sujet. Dans une publication intitulée « Moral Intuitions », Sinnott-Armstrong, Young et Cushman décrivent les intuitions morales comme des processus heuristiques, lesquels impliquent « une substitution inconsciente d'attributs » (*unconscious attribute substitution*) (Sinnott-Armstrong, Young et Cushman 2010). Tiré d'un article de Kahneman et Frederick, le concept de « substitution inconsciente d'attributs » comporte qu'il y a substitution entre deux propriétés, la propriété

recherchée et la propriété heuristique (Sinnott-Armstrong, Young et Cushman 2010, 250–51). Lorsqu'un objet possède un attribut qui est difficile à détecter, en raison d'un manque de ressources, l'agent aura tendance à substituer cet attribut pour un attribut heuristique, déjà présent dans l'esprit. Ceci revient à dire que l'usage d'une règle heuristique consiste dans un processus de substitution. Cette substitution s'effectue entre une propriété objective (*target attribute*) et une propriété subjective (*heuristic attribute*). Pour cette raison, les heuristiques seraient des défaiteurs sur le plan moral. En d'autres mots, les heuristiques ne seraient pas des processus épistémiquement fiables, car elles ne prennent pas en compte les propriétés de l'objet visé, mais bien les propriétés du sujet.

L'hypothèse de la « substitution inconsciente d'attributs » sera toutefois rejetée par Liao, lequel stipule que les intuitions ne sont simplement pas des heuristiques (Liao 2016, 314). L'argument formulé par Liao se décline en trois points :

P1 : Les intuitions morales ne constituent pas une forme de raisonnement.

P2 : Les heuristiques constituent une forme de raisonnement.

C1 : Donc, les intuitions ne peuvent être des heuristiques. (Liao 2016, 328)

Clarifions d'abord la prémisse #1. Liao prétend que les intuitions ne sont pas une forme de raisonnement. En effet, si les intuitions sont définies comme un processus non inférentiel et que le raisonnement est défini comme un processus inférentiel, nous pouvons déduire que les intuitions ne constituent pas une forme de raisonnement (Liao 2016, 319–20). Pour Liao, ce jugement n'est pas controversé.

Passons maintenant à la prémisse #2. Liao soutient que les règles heuristiques peuvent être comprises de trois manières différentes : des substitutions conscientes d'attributs (1), des substitutions inconscientes d'attributs (2) ou des procédés rapides et frugaux facilitant la résolution de problèmes (3). Liao tente de montrer que chacune de ces définitions correspond à une forme de raisonnement, et donc, que toutes les conceptions envisageables des heuristiques sont des formes de raisonnement (Liao 2016, 319–27). En ce qui concerne le *premier* élément, Liao argumente que les heuristiques de la disponibilité, de la représentativité ou de l'affect sont toutes des processus inférentiels (Liao 2016, 319–21). Pour ce qui est du *second* élément, Liao défend l'idée

(controversée) que le raisonnement peut être inconscient et l'idée (tout aussi controversée) que l'intuition n'est pas nécessairement inconsciente (Liao 2016, 321–25). Quant au *troisième* élément, Liao procède par arguments successifs : d'abord, il montre qu'un processus cognitif peut être rapide sans être une intuition; ensuite, il montre qu'un jugement peut être le résultat d'une intuition sans avoir été formé rapidement; et finalement, il montre que les intuitions ne peuvent être des heuristiques qui simplifient un problème complexe, car les simplifications seraient aussi une forme de raisonnement (Liao 2016, 325–27).

En conjuguant P1 & P2, nous avons que les intuitions ne peuvent être des heuristiques. Ceci implique que les intuitions morales, plus spécifiquement, ne constituent pas des heuristiques. Plus encore, si les intuitions morales ne sont pas basées sur des règles heuristiques, cela signifie que ces mêmes intuitions ne peuvent être défaits par des règles heuristiques. Et donc, tout argument sceptique qui présuppose un lien entre intuition et heuristique s'en retrouve neutralisé (Liao 2016, 328).

Toutefois, il serait prématuré de rejeter les thèses de Greene, Sunstein ou Sinnott-Armstrong, car la seconde prémisse de l'argument avancé par Liao demeure plutôt controversée. Comportant nombre de ramifications et de points de rupture, cette prémisse dépend elle-même de sous-prémises discutables. Notamment, il n'est pas évident que l'heuristique de l'affect est un processus inférentiel ou que le raisonnement puisse être inconscient. Les arguments succincts que propose Liao ne suffisent guère à prouver de telles affirmations. Cela dit, l'auteur aura réussi à jeter un doute raisonnable sur l'idée que les intuitions morales constituent des règles heuristiques, sans plus.

Mis à part la proposition de Liao, les travaux de Sunstein et de ses contemporains suggèrent que les intuitions morales sont des produits de règles heuristiques, logées au sein du Système I, lesquelles peuvent échouer dans des situations non familières⁴⁵. Non seulement ces règles peuvent-elles échouer, mais nous n'avons aucune raison de croire que celles-ci ont émergé pour pister la vérité morale. Au contraire, il semblerait que celles-ci découlent de facteurs écologiques contingents (davantage là-dessus plus loin).

⁴⁵ Si les effets de cadrage (*framing effects*) constituent aussi des « illusions morales » au sens de Sinnott-Armstrong, ils n'ont pas été abordés ici pour deux raisons : ils ne relèvent pas nécessairement du Système I et, pour cette raison, sortent du cadre étiologique proposé (Sinnott-Armstrong 2006, 352–56; 2006, 204–7).

2.3 Intuitionnisme huémérien et modèle du double processus : une épreuve

Au cours de ce chapitre, nous avons établi que les intuitions morales sont instables, ou aisément manipulables, et que ces intuitions trouvent leur origine dans le Système I, lequel regroupe des processus automatiques et inconscients, compris comme des mécanismes d'apprentissage sans modèle (*model-free learning*). Les processus impliqués sont porteurs d'une valence affective et prennent la forme des règles d'action heuristiques. Ces règles, comportant des bienfaits évolutifs, occupent une fonction essentiellement sociale. Appliquées de manière pré-délibérative, ces règles invitent le déploiement du Système II, lequel est responsable de la rationalisation ou du raisonnement contrôlé. Est-ce que cette conception de l'étiologie des jugements moraux est compatible avec l'intuitionnisme huémérien?

Huemer nous dit que les intuitions morales sont des états phénoménaux comportant du contenu propositionnel, du genre « sujet-copule-prédicat ». Il nous dit que ces intuitions intellectuelles correspondent à des vérités morales synthétiques *a priori*. Huemer souligne toutefois que nos intuitions sont faillibles, et révisables. Si cette perspective est déflationniste à certains égards, il n'en demeure pas moins qu'elle est trop ambitieuse, surtout lorsqu'il s'agit d'intuitions concrètes, intermédiaires et abstraites non formelles. En ce sens, les résultats empiriques présentés ci-haut suggèrent fortement que nos croyances morales substantives ne proviennent pas d'une faculté rationnelle et qu'elles ne correspondent pas à des faits indépendants. Il y aurait donc incompatibilité entre le modèle du double système et l'épistémologie rationaliste que propose Huemer. Formalisons cet argument :

P1. Prémisse doxastique / explanandum : S a des croyances morales substantives (P).

P2. Prémisse explicative / explanans : La croyance de S que P est expliquée par le Système I, un type de cognition automatique et inconscient, lequel repose sur l'affect. Le rôle du Système II est secondaire : il s'agit d'un arbitre qui délibère et gère les conflits, au besoin.

P3. Prémisse épistémique / normative : Le Système I est un processus qui n'est pas épistémiquement fiable.

C1 : Conclusion : La croyance de S que P n'est pas justifiée.

Cet argument ne remet aucunement en question la phénoménologie morale de Huemer. L'expérience morale demeure réaliste et la sémantique morale demeure cognitiviste. De même, la justification demeure de nature internaliste. Or, cet argument remet en question l'idée que nos intuitions morales substantives proviennent d'une faculté rationnelle (1), qu'elles correspondent à des vérités synthétiques *a priori* (2) et, surtout, qu'elles sont justifiées au sens où le voudrait Huemer (3).

Si l'on se fie aux arguments étiologiques présentés jusqu'ici, nous avons bonne raison de croire que nos intuitions morales trouvent leurs origines dans l'affect et, pour cette raison, qu'elles ne correspondent pas à des vérités rationnelles synthétiques *a priori*. En effet, les données empiriques récoltées par les chercheurs en neurosciences et en psychologie suggèrent fortement que les structures dites « émotionnelles » sont causalement impliquées dans une vaste proportion de jugements moraux (Woodward et Allman 2007, 182–83, 186, 193, 200; Allman et Woodward 2008, 169, 181; Haidt et Kesebir 2010, 808–14; Sinnott-Armstrong, Young et Cushman 2010, 268–70; Greene 2014, 718; Woodward 2016, 106–10). Sans cet apport affectif, la moralité serait un concept vain :

Reason, narrowly conceived, is too weak—too lacking in substantive content—to supply sound derivations of substantive moral requirements of the sort required. This is why no one has been able to produce them. The additional processing that is at work in generating the moral judgments we make is supplied by the “emotional” structures discussed in this chapter. When we leave these out of the picture, not enough remains (just in “reason” itself) to generate, either causally or as a matter of logic, the content of our moral judgments. (Woodward 2016, 108)

Comme le souligne Woodward, il est peu probable que la raison seule soit capable d'informer la moralité, surtout lorsqu'il s'agit d'intuitions concrètes, intermédiaires et abstraites non formelles. En effet, les recherches de Greene et Haidt ont montré que la prise de décision mobilise (le plus souvent) des structures émotionnelles, et que ces structures n'obéissent pas à une raison désincarnée. De même, il semblerait que les règles heuristiques employées en situation capturent des propriétés internes du sujet et non des propriétés externes au sujet. Il suffit de songer à l'heuristique de l'affect pour s'en convaincre.

Rétrospectivement, on constate que l'intuition morale diffère, à plusieurs égards, de l'intuition métaphysique, logique ou mathématique. Bien que la structure propositionnelle de ces

intuitions soit similaire, leur étiologie est tout à fait différente, que ce soit sur le plan psychoaffectif ou neurophysiologique⁴⁶. Cette différence explique, notamment, la faillibilité accrue des intuitions évaluatives, devant leurs contreparties non évaluatives. Or, il importe de souligner que toutes les intuitions évaluatives ne sont pas équivalentes. Effectivement, il semblerait que les intuitions évaluatives formelles ressemblent davantage aux intuitions intellectuelles non évaluatives, ce pourquoi elles ne seraient pas susceptibles à l'argument étiologique développé ici. Par exemple, il semblerait que les principes de transitivité, d'additivité et de survenance résistent à l'argument étiologique, sachant qu'ils ne mobilisent aucun contenu psychologique contingent. À l'inverse, il semblerait que les principes du conséquentialisme, du kantisme, du contractualisme, de l'éthique arétaïque ou chrétienne, ne puissent résister à l'offensive sceptique. Participant de l'éthique normative, ces jugements substantifs s'accompagnent vraisemblablement d'une valence affective, ce qui expliquerait leur défaisabilité accrue⁴⁷. Au même titre que les intuitions intermédiaires, les intuitions abstraites non formelles pourraient être considérées comme des rationalisations *post hoc* à partir de réponses affectives prépotentes. C'est donc dire que les grands canons de la rationalité morale ne seraient que préjugés arborant des habits d'objectivité!

Ceci étant dit, on soulignera que la dichotomie « émotion-rationalité » est de plus en plus contestée dans la sphère de la psychologie morale et de la neuroscience. Comme le soulignent Cushman (2013, 273–92), Huebner (2015, 427–40), Woodward (2016, 108) et bien d'autres (Demaree-Cotton et Kahane 2020), l'émotion désigne un type de cognition plus complexe, plus riche que l'on pensait auparavant. Associées à l'apprentissage sans modèle (*model-free learning*), les processus émotionnels faciliteraient l'acquisition de connaissances ainsi que la prise de décision. Véritable calculatrice biologique, le système affectif effectue des opérations rapides et

⁴⁶ Bien que cela puisse sembler trivial, les structures cérébrales impliquées dans la formation des intuitions morales (concrètes/intermédiaires/abstraites non formelles) diffèrent de celles impliquées dans le raisonnement mathématique ou logique. Notamment, on sait que les régions dites « émotionnelles » contribuent à la formation des intuitions morales, mais non au raisonnement mathématique. Toutefois, cela ne signifie pas que les intuitions morales sont nécessairement fausses. Il s'agirait là d'un *sophisme génétique*. Au contraire, il faut simplement reconnaître que les intuitions morales sont d'une étiologie particulière, et que cette étiologie ne cadre pas avec l'épistémologie rationaliste prônée par Huemer (Allman et Woodward 2008, 169; Woodward 2016, 106–7; Moll, de Oliveira-Souza et Zahn 2008, 161–80; Zahn, de Oliveira-Souza et Moll 2015; de Oliveira-Souza, Moll et Zahn 2021, 247–53).

⁴⁷ La thèse que les intuitions abstraites non formelles s'accompagnent nécessairement d'une charge affective mériterait une discussion plus élaborée et une défense plus robuste. Pour ce mémoire, il suffira de mentionner qu'il s'agit d'une idée tout à fait plausible, laquelle s'aligne avec les résultats empiriques provenant de la psychologie sociale et de la neuroscience. Ces résultats suggèrent, pour la majeure partie, que l'assignation de prédicats moraux à un état de fait mobiliserait les fonctions affectives.

implicites, lesquelles reflètent le mode d'inférence bayésien. Les projections qui en résultent aiguillonnent le comportement et régissent les rapports interpersonnels. Adopté par un nombre croissant de chercheurs, le modèle neuro-computationnel possède un pouvoir explicatif largement supérieur au modèle philosophique traditionnel, ancré dans la dichotomie entre émotion et raison.

Or, si l'on s'en tient à ces dualismes⁴⁸, on soulignera simplement que l'argument étiologique présenté ici jette un doute sérieux sur l'intuitionnisme rationaliste de Huemer et, plus précisément, sur la nature *apriorique* des vérités qui en découleraient. Sachant désormais que les émotions déterminent causalement une vaste gamme d'intuitions morales, il n'est plus possible de tenir que nos intuitions morales ont un statut ontologique et épistémique aussi solide que les vérités logiques et mathématiques, à l'exception des intuitions évaluatives formelles⁴⁹. Pour ces quelques raisons, on dira qu'une majorité de nos jugements moraux sont compromis, d'un point de vue justificatif, par l'argument étiologique dressé ci-haut.

Pourtant, certains pourraient soutenir que le Système I est capable de pister la vérité morale, même s'il n'opère pas selon les prémisses de la raison logico-déductive. Une telle suggestion sera rapidement écartée par les sceptiques, affirmant que nos facultés morales servent des fins évolutives et non épistémiques. Cette critique généalogique évolutionnaire fera l'objet du prochain chapitre.

Chapitre 3 : La critique généalogique évolutionnaire de Street

Au cours de ce chapitre, la question étiologique sera abordée dans une perspective évolutionnaire. Plusieurs auteurs, comme Michael Ruse (1986), Richard Joyce (2001; 2007; 2016) et Sharon Street (2006; 2016), ont formulé des critiques généalogiques évolutionnaires, lesquelles jettent un doute sur la fiabilité de nos intuitions morales. Or, c'est la critique de Sharon Street qui nous intéressera ici, comme celle-ci est la mieux formulée et la plus discutée dans la littérature. Si certains (Klenk 2017; Schechter 2018) ont tracé des parallèles entre l'argument de Street et le défi de Field-Benacerraf (1989) ou encore le défi explicatif de Harman (1977), il n'en demeure pas moins qu'il s'agit d'un argument original, lequel mérite un traitement tout aussi original. Une fois

⁴⁸ À travers ses écrits, Huemer entretient la distinction classique entre « émotion » et « raison ». Or, comme on l'affirme plus haut, ce présupposé est conceptuellement et empiriquement suspect.

⁴⁹ On discutera plus directement du statut épistémique des intuitions métadiscursives dans le prochain chapitre.

l'argument présenté, il s'agira d'introduire certaines objections, surtout centrées autour d'un potentiel « troisième facteur ». Une fois ces objections présentées et analysées, il sera possible d'évaluer les implications de la critique évolutionnaire de Street pour l'intuitionnisme huémérien.

3.1 Le dilemme darwinien de Street : contre les théories réalistes de la valeur

L'argument formulé par Street, dans son article de 2006, a pour cible les théories réalistes de la valeur. Par « théories réalistes », on entend toutes ces théories métaéthiques, lesquelles supposent qu'il existe des faits évaluatifs indépendants de *toutes* attitudes évaluatives adoptées par les agents moraux (Street 2006, 110–12). Ces théories réalistes se divisent en deux camps : les théories non naturalistes et les théories naturalistes. Les théories non naturalistes, dont celles de Huemer, prétendent que les faits évaluatifs ne sont pas réductibles à des faits naturels et, par le fait même, que ces faits sont causalement inertes. Au contraire, les théories naturalistes supposent que les faits évaluatifs sont réductibles à des faits naturels et, par le fait même, que ces faits sont causalement efficaces. Telles sont les théories visées par l'argument de Street.

Ayant d'abord identifié les cibles de son défi sceptique, Street présente ensuite la prémisse centrale de son argument : « la sélection naturelle a eu une influence énorme sur le contenu de nos jugements évaluatifs » (Street 2006, 113). Employant la notion darwinienne de « succès reproductif », Street conjecture que les pressions évolutives auraient favorisé les individus dont les jugements évaluatifs étaient bénéfiques sur le plan reproductif. Corollairement, ces mêmes pressions évolutives auraient défavorisé les individus dont les jugements évaluatifs étaient coûteux sur le plan reproductif. Pour se convaincre de telles spéculations, il suffit de comparer le potentiel reproductif d'un individu qui fuit le danger contre celui d'un individu qui recherche le danger, ou encore, d'un individu qui prend soin de sa progéniture contre celui d'un individu qui malmène sa progéniture (Street 2006, 114–15).

Afin d'appuyer la thèse que la sélection naturelle a influencé le contenu de nos jugements moraux, Street souligne deux constats empiriques. D'une part, les humains, malgré les époques changeantes et la diversité des cultures, formulent des jugements évaluatifs drôlement uniformes, lesquels promeuvent la survie et la reproduction. D'autre part, ces jugements largement répandus s'alignent avec les tendances évaluatives fondamentales des animaux et, plus particulièrement, des animaux apparentés aux humains.

En ce qui concerne le premier constat, Street identifie six jugements évaluatifs approuvés d'une manière quasi universelle par les humains :

1. Le fait que quelque chose promouvrait notre survie est une raison de favoriser cette chose.
2. Le fait que quelque chose promouvrait les intérêts d'un proche est une raison de le faire.
3. Nous avons une obligation plus grande d'aider nos propres enfants que d'aider les enfants d'autrui.
4. Le fait que quelqu'un nous ait bien traité est une raison de bien traiter cette personne en retour.
5. Le fait qu'une personne soit altruiste est une raison d'admirer, d'applaudir et de récompenser cette personne.
6. Le fait qu'une personne ait commis un mal délibéré est une raison d'éviter cette personne ou de la punir. (Street 2006, 115; Street 2015, 688)

Street explique la quasi-universalité de ces jugements par la biologie évolutionnaire ou, plus précisément, par la sélection de la parentèle et par l'altruisme réciproque. Ces jugements sont exactement les types de jugements que la sélection naturelle aurait favorisés, car ils maximisent la survie et la reproduction, affirme Street.

En ce qui concerne le second constat, il s'agit de reconnaître que les jugements identifiés ci-haut sont tout aussi répandus chez des formes de vie animales moins évoluées, comme les primates. Justement, ces primates semblent avoir une expérience morale similaire à celles des humains. Sans surprise, les tendances évaluatives exprimées par ces animaux ressemblent également à celles des humains. Selon Street, cette continuité dans l'expérience morale des primates et des humains appuie fortement la thèse que la sélection naturelle a informé le contenu de nos jugements évaluatifs (Street 2006, 117).

Cela dit, Street reconnaît que son modèle évolutionnaire comporte deux lacunes importantes, lesquelles peuvent toutefois être corrigées sans trop de difficultés. L'une concerne le développement de nos capacités évaluatives tandis que l'autre concerne l'héritabilité génétique (supposée) des jugements évaluatifs. Pour chaque aspect, Street procède de la même manière : elle identifie l'hypothèse problématique, réfute cette hypothèse et propose une solution.

D'une part, Street constate que son modèle évolutionnaire a des implications contre-intuitives pour la généalogie de nos capacités morales. Si l'on s'en fie à ce modèle, on serait tenté d'admettre

l'hypothèse suivante : *d'abord*, nos ancêtres ont formulé des jugements évaluatifs et, *ensuite*, le contenu de ces jugements a été soumis à un processus sélectif. Or, comme Street le souligne, il est improbable que nos ancêtres aient commencé par former des jugements évaluatifs réfléchis et verbalisés (*full-fledged evaluative judgment*). Au contraire, il est bien plus probable que des tendances comportementales et motivationnelles se soient d'abord développées chez nos ancêtres et qu'une capacité aux jugements évaluatifs réfléchis et verbalisés se soit ensuite développée. Une telle logique est beaucoup plus intuitive, estime Street (2006, 118).

D'autre part, Street remarque que sa formule a des conséquences très lourdes sur le plan de l'héritabilité génétique de nos dispositions évaluatives. Si l'on s'appuie sur la théorie de la sélection naturelle, on pourrait supposer que la capacité à faire des jugements évaluatifs réfléchis et verbalisés soit un trait génétiquement héritable. Or, il s'agit là d'une hypothèse totalement improbable, beaucoup trop exigeante. À l'inverse, il serait sensé de croire que nos tendances évaluatives fondamentales préréflexives et préverbales constituaient des traits génétiquement hérissables. C'est là une hypothèse beaucoup plus parcimonieuse, prétend Street (2006, 119).

Ayant corrigé ces deux lacunes conceptuelles, Street en vient à la conclusion que les pressions sélectives n'ont eu qu'une influence *indirecte* sur le contenu de nos jugements évaluatifs. En ce sens, la sélection naturelle aurait façonné *directement* nos tendances évaluatives fondamentales, mais *indirectement* nos jugements évaluatifs réfléchis et verbalisés. Cependant, cela ne revient pas à dire que la sélection naturelle aurait eu un impact modeste sur le contenu de nos jugements évaluatifs. Au contraire, l'impact subit aurait été énorme, mais cet impact n'est ressenti qu'*indirectement*. Cela signifie que les pressions évolutives agissent de manière *médiée* sur les jugements évaluatifs réfléchis et verbalisés (Street 2006, 119–21).

L'influence de la sélection naturelle sur le contenu de nos jugements évaluatifs ayant été fermement établie, Street peut désormais procéder à la formulation du dilemme darwinien, lequel constitue l'élément central de son argument étiologique. Le dilemme posé par Street est le suivant : l'adepte du réalisme moral peut soit *nier* ou *affirmer* l'existence d'une relation entre les *forces sélectives* ayant influencé les contenus de nos jugements évaluatifs et les *vérités morales* postulées. Dans tous les cas, le partisan du réalisme moral se retrouve coincé.

3.1.1 Branche I : négation d'une relation

Si le réaliste *nie* l'existence d'une relation entre les forces sélectives et les vérités morales postulées, le réaliste doit aussi admettre que les forces sélectives ont façonné nos tendances évaluatives d'une manière qui n'a rien à voir avec la vérité morale. Afin d'illustrer la chose, Street en va d'une analogie nautique :

On this view, allowing our evolutionary judgements to be shaped by evolutionary influences is analogous to setting out for Bermuda and letting the course of your boat be determined by the wind and tides: just as the push of the wind and tides on your boat has nothing to do with where you want to go, so the historical push of natural selection on the content of our evaluative judgements has nothing to do with evaluative truth. (Street 2006, 121)

Tout comme le navigateur qui se laisse porter au gré des vagues, l'agent moral est éprouvé par les forces évolutives, lesquelles guident le jugement d'une manière parfaitement contingente et certainement indépendante des vérités morales. Et donc, si le réaliste devait choisir la première branche du dilemme, il serait contraint d'accepter que les jugements évaluatifs actuels ne pistent pas la vérité, et seraient fort probablement faux. Évidemment, il existe une probabilité non nulle qu'une portion de nos jugements évaluatifs intersecte une portion des vérités morales, mais cela constituerait une coïncidence, un coup de chance étrangement convenable pour le réaliste, soutient Street (2006, 121–22; 2016, 304–6).

Malgré ces arguments quasi irrésistibles, le partisan du réalisme pourrait toujours s'en remettre au pouvoir émancipateur de la réflexion rationnelle, laquelle permettrait de transcender l'influence indésirable des facteurs évolutifs. Pour reprendre l'analogie nautique, le compas de la rationalité ne pourrait-il pas sauver l'agent moral de son inexorable dérive? Street répond par la négative. Selon l'autrice, l'exercice de la raison ne peut affranchir l'agent moral des influences de l'évolution, car la raison est elle-même prisonnière de ces contingences naturelles. En effet, que peut la raison, que peut la méthode de l'équilibre réflexif, sachant que tous les jugements évaluatifs que nous formulons sont déjà contaminés? À quoi sert de classer, de cribler et de départager, si toutes les alternatives sont viciées? Réalisant l'impotence de la raison, Street dresse un constat sceptique :

[...] all our reflection over the ages has really just been a process of assessing evaluative judgments that are mostly off the mark in terms of others that are mostly off the mark. And reflection of this kind isn't going to get us any closer to evaluative truth,

any more than sorting through contaminated materials with contaminated tools is going to get on closer to purity. (Street 2006, 124)

Selon l'autrice, l'impureté préexistante de nos jugements évaluatifs rend toute intervention de la raison futile. C'est pourquoi elle juge que le réaliste est sans recours, pour ce qui est de la première branche du dilemme darwinien.

3.1.2 Branche II : affirmation d'une relation

Maintenant, si le réaliste *affirme* l'existence d'une relation entre les forces sélectives et les vérités morales postulées, le réaliste doit aussi fournir une explication plausible quant à la nature de la relation affirmée. Selon *l'hypothèse du pistage*, les jugements évaluatifs sélectionnés par les forces naturelles constituent précisément ces jugements qui sont vrais. Cette hypothèse comporte deux clauses : d'une part, elle suppose que la meilleure explication de nos jugements évaluatifs est la vérité de ces jugements et, d'autre part, que la capacité de détecter de telles vérités est avantageuse sur le plan de la survie et de la reproduction (Street 2006, 125–26). Or, Street considère que cette explication scientifique fait bien piètre figure devant ses rivales.

Pour l'autrice, *l'hypothèse du pistage* serait aisément supplantée par *l'hypothèse du lien adaptif*. Street définit la proposition ainsi :

According to what I call the *adaptive link account*, tendencies to make certain kinds of evaluative judgments rather than others contributed to our ancestor's reproductive success not because they constituted perceptions of independent evaluative truths, but rather because they forged adaptive links between our ancestor's circumstances and their responses to those circumstances, getting them to act, feel, and believe in ways that turned out to be reproductively advantageous. (Street 2006, 127)

En l'occurrence, nous avons que certaines tendances évaluatives ont été favorisées par l'évolution, car les agents qui possédaient ces tendances ont développé des mécanismes de réponse avantageux sur le plan reproductif. Les six jugements évaluatifs énumérés plus haut constituent de tels mécanismes. Dans chacun de ces cas mentionnés, il y aurait un jumelage entre les circonstances matérielles et la réponse fournie par l'agent moral. Pour illustrer ce phénomène, il suffit d'analyser le quatrième jugement évaluatif introduit dans la liste : « Le fait que quelqu'un nous ait bien traité est une raison de bien traiter cette personne en retour ». Dans ce cas précis, un lien adaptif aurait été formé entre « le fait que quelqu'un nous ait bien traité » et « le fait de bien traiter cette personne

en retour ». Conformément au principe d'altruisme réciproque, une circonstance favorable invite une réponse favorable.

Ayant présenté deux explications divergentes de nos convictions morales, soit *l'hypothèse du pistage* et *l'hypothèse du lien adaptif*, Street donne trois raisons pour lesquelles elle préfère la seconde option. Ces trois raisons sont: la parcimonie, la clarté et le pouvoir explicatif.

Premièrement, Street soutient que *l'hypothèse du lien adaptif* est plus parcimonieuse que *l'hypothèse du pistage*, car elle explique le phénomène moral sans postuler d'entités métaphysiques censées correspondre à des vérités morales indépendantes (Street 2006, 129). Deuxièmement, Street affirme que *l'hypothèse du lien adaptif* est plus claire que *l'hypothèse du pistage*, car elle ne suppose aucune relation obscure entre vérité morale et avantage reproductif. En l'occurrence, Street inculpe les réalistes, non naturalistes et naturalistes, d'un manque de clarté, car ceux-ci sont incapables d'expliquer pourquoi la perception de vérités évaluatives indépendantes serait un gage de succès reproductif. S'il est généralement avantageux d'entretenir des croyances vraies, que ce soit pour éviter des prédateurs ou localiser des réserves alimentaires, il n'est pas évident qu'il soit tout aussi avantageux de pister des vérités morales, lesquelles semblent plutôt superflues d'un point de vue évolutionnaire⁵⁰. Troisièmement, Street prétend que *l'hypothèse du lien adaptif* possède un pouvoir explicatif largement supérieur à celui de *l'hypothèse du pistage*. En ce sens, *l'hypothèse du lien adaptif* serait beaucoup mieux à même d'expliquer les tendances évaluatives partagées par une majorité d'humains (Street 2006, 132). Selon Street, lesdites tendances doivent être conçues comme des réponses adaptives, et non pas comme des dispositions intellectuelles à saisir une réalité immuable. Street donne trois arguments pour défendre sa prise de position. D'abord, les réalistes sont incapables d'expliquer le fait que l'ensemble des vérités évaluatives postulées correspond précisément (hormis quelques exceptions) à l'ensemble des jugements évaluatifs sélectionnés par l'évolution. Pourquoi donc se réclamer de *l'hypothèse du pistage*, une théorie qui postule des entités superflues sur le plan explicatif? Ensuite, il semblerait que les réalistes peuvent difficilement expliquer notre prédisposition à effectuer certains jugements évaluatifs, lesquels nous abandonnons après révision. Street prend pour exemple notre prédisposition à favoriser les

⁵⁰ Ceci n'exclut pas qu'il est avantageux de se plier à certaines *normes* morales. Il s'agit simplement de reconnaître que les normes opérantes pourraient bien ne pas être sous-tendues par des *vérités* morales indépendantes. Plus encore, il semble que ces vérités morales ne sont aucunement nécessaires à l'explication causale de nos croyances.

membres du groupe auquel nous appartenons et à défavoriser les non-membres. Ici, les réalistes sont incapables d'expliquer notre disposition au favoritisme endogroupe, compte tenu du fait que cette tendance à l'exclusion n'est plus considérée comme une vérité évaluative⁵¹. De toute évidence, *l'hypothèse du pistage* ne peut rendre compte des jugements évolutivement avantageux que nous répudions désormais. Finalement, les réalistes ne peuvent expliquer l'univers des jugements évaluatifs que nous pouvons faire, mais ne faisons pas. Cette fois-ci, Street donne l'exemple de quelques jugements farfelus comme « les plantes ont plus de valeur que les humains » ou « le fait que quelque chose soit de couleur pourpre est une raison de la soumettre à nos cris » (Street 2006, 133). En l'occurrence, *l'hypothèse du pistage* peut seulement affirmer la fausseté de ces jugements, sans toutefois donner de raisons substantielles pour justifier une telle proposition. Au contraire, *l'hypothèse du lien adaptif* peut expliquer pourquoi de tels jugements ne sont pas entérinés, en soulignant leur inutilité. Plus généralement, *l'hypothèse du lien adaptif* explique nos jugements évaluatifs avec clarté et parcimonie. En ce sens, nous avons les jugements que nous avons, car ceux-ci auraient permis un jumelage vertueux entre réponse et circonstance (Street 2006, 134).

Au final, il semblerait que le réaliste soit dans une bien mauvaise posture. N'ayant d'autre recours que *l'hypothèse du pistage*, le réaliste peut difficilement rivaliser avec l'antiréaliste, lequel est armé de *l'hypothèse du lien adaptif*, une explication causale décidément plus convaincante.

3.1.3 Objections envisagées par Street

Une fois son argument présenté, Street envisage trois réponses potentielles à son dilemme. La première objection, concernant le réalisme naturaliste, ne sera pas abordée ici. Toutefois, il sera pertinent d'aborder la seconde et la troisième objection, comme celles-ci sont généralisables à toutes les formes de réalisme.

Selon « l'hypothèse du sous-produit », notre aptitude à percevoir les vérités morales serait dérivée d'une *autre capacité*, laquelle possède une explication évolutionnaire. Au même titre que notre capacité à effectuer du calcul intégral, notre aptitude à la morale n'aurait pas été directement

⁵¹ Notons que Street n'a pas accès à la notion de vérité au sens robuste, considérant son positionnement antiréaliste. Selon toute vraisemblance, l'autrice devra mobiliser une conception plus faible de la vérité, laquelle s'inscrit en continuité avec son constructivisme humien (Street 2006, 152–54).

sélectionnée, mais elle proviendrait d'une *autre capacité* façonnée par les forces évolutives. Or, cette hypothèse est stérile, affirme Street, car elle ne ferait que déplacer le problème. Ici encore, le dilemme darwinien est source d'embarras, car le réaliste doit expliquer en quoi cette *autre faculté* serait reliée aux vérités morales indépendantes. Deux choix se présentent à lui : nier ou affirmer une relation. La première branche suppose une coïncidence improbable tandis que la seconde suppose une explication scientifique insatisfaisante (Street 2006, 142–44).

Si l'argument précédent est contestable, le réaliste n'est toutefois pas sans recours. Afin de se tirer d'affaire, il peut faire appel à des intuitions largement partagées. Parmi celles-ci, figure l'intuition voulant que la douleur soit intrinsèquement mal. En l'occurrence, la relation entre évolution et vérité morale est plus évidente : ressentir de la douleur est avantageux sur le plan évolutif et ressentir de la douleur est mal, indépendamment de toute contingence. Sans surprise, Street ne mord pas à l'hameçon. Elle soutient que le caractère évaluatif de la douleur est fonction de nos attitudes particulières. Pour défendre ce point, elle déploie un argument intitulé « le dilemme de la douleur ». Ici, le choix présenté au réaliste est de nier ou d'accepter une définition de la douleur. Cette définition se lit ainsi: « Pain is a sensation such that the creature having the sensation unreflectively [and necessarily] takes that sensation to count in favor of doing whatever would avoid, lessen, or stop it. » (Street 2006, 146) Si le réaliste devait nier cette définition, car il considère que la douleur est une sensation telle que l'agent n'éprouve *pas* préréflexivement et nécessairement cette sensation comme une raison de l'éviter, de l'amenuiser ou de la cesser, il sera contraint d'accepter la possibilité que la douleur soit à la fois une sensation que l'on recherche et un mal intrinsèque. En effet, si un médicament miraculeux devait voir le jour, et si ce médicament amenait ses consommateurs à apprécier la douleur, donc, nous avons que la douleur est à la fois plaisante et intrinsèquement mal (Street 2006, 148–49). Voilà qui est curieux. Sachant que cet argument ne séduira pas tous les réalistes, Street affirme qu'il est toutefois possible d'adapter le présent exemple pour le rendre compatible avec le dilemme darwinien. Encore une fois, si le réaliste considère que la douleur est une sensation telle que l'agent n'éprouve *pas* préréflexivement et nécessairement cette sensation comme une raison de l'éviter, de l'amenuiser ou de la cesser, il devra expliquer pourquoi les humains ont choisi, choisissent et choisiront de fuir la douleur plutôt que de la rechercher. Ici aussi, le réaliste est confronté à une coïncidence improbable ou à une explication scientifiquement intenable. Cela dit, le réaliste peut toujours opter pour la seconde branche du dilemme de la douleur, acceptant ainsi la définition originale donnée par Street. Or, s'il

en fait ainsi, le réaliste sera forcé de reconnaître que le statut moral de la douleur, en tant que sensation, dépend précisément de nos attitudes préreflexives. Et donc, si la douleur est mal, c'est en vertu de nos attitudes évaluatives particulières, et non de vérités *sui generis*. Embarrassé par le « dilemme de la douleur », le partisan du réalisme est contraint de retirer son argument, sachant désormais que la douleur ou le plaisir ne peuvent constituer un point fixe de la morale (Street 2006, 150–51).

3.1.4 Schéma de l'argument

Ayant délinéé les grands moments du dilemme darwinien de Street, nous pouvons désormais rendre l'argument de manière plus formelle :

P1. Hypothèse évaluative : Le réalisme évaluatif est vrai.

P2. Hypothèse de la sélection naturelle : La sélection naturelle a *directement* façonné nos tendances évaluatives fondamentales. Pour cette même raison, la sélection naturelle influence *indirectement* les contenus de nos jugements évaluatifs réfléchis et verbalisés.

P3. Le dilemme :

Branche I / Nier une relation: Le réaliste peut nier qu'il existe une relation entre les vérités évaluatives et les pressions évolutives. Ici, le réaliste suppose que les pressions évolutives contaminent nos jugements évaluatifs. Faire appel à une *coïncidence fortuite* ou au *pouvoir émancipateur de la raison* est une stratégie stérile. Donc, « nier une relation » est une option inacceptable pour le réaliste, celui-ci devant accepter le *scepticisme normatif*.

Branche II / Affirmer une relation : Le réaliste peut affirmer qu'il existe une relation entre les vérités évaluatives et les pressions évolutives. En l'occurrence, il doit se rabattre sur *l'hypothèse du pistage*. Or, cette explication est scientifiquement et philosophiquement inférieure à *l'hypothèse du lien adaptif*. Donc, « affirmer une relation » est une option inacceptable pour le réaliste.

C1. Conclusion : Ne pouvant accepter ni l'une, ni l'autre des branches, le réaliste est contraint à l'abdication. En d'autres mots, le réalisme évaluatif est faux.

Regroupant les éléments centraux de l'argument, le schéma ci-haut montre bien l'ampleur des difficultés qui attendent le réaliste. Or, depuis la publication du manuscrit en 2006, une variété de réponses ont été suggérées par d'excellentissimes philosophes. À l'occasion de ce mémoire, nous

aborderons un type de réponse relativement répandu, connu sous le nom d'explications par un « troisième facteur ».

3.2 L'explication par un « troisième facteur » : une solution tangible?

Plusieurs auteurs, dont David Enoch (2010, 413–38; 2011, 168–75), Erik J. Wielenberg (2010, 441–64; 2014, 134–77) et Knut Olav Skarsaune (2011, 229–43), ont suggéré une explication par un « troisième facteur⁵² ». Afin d'introduire le sujet, je tâcherai d'explicitier la proposition d'Enoch.

3.2.1 Enoch : la survie comme « troisième facteur »

Dans un article intitulé « The epistemological challenge to metanormative realism [...] », Enoch propose une version unifiée du défi épistémique concernant les interprétations robustes du réalisme normatif. Souvent conçu en termes d'accès épistémique, de justification, de fiabilité, de connaissance, ou d'arguments sceptiques, ce défi épistémique serait mieux compris en termes d'explication, prétend Enoch (2010, 413). En l'occurrence, il s'agirait « d'expliquer la corrélation entre les croyances normatives ainsi que les vérités morales indépendantes » (Enoch 2010, 415–20). Postulant des vérités morales causalement inertes, l'adepte du réalisme robuste est contraint de fournir une explication non causale de la correspondance entre croyances et vérités morales⁵³. Cette relation sera nécessairement indirecte (Lutz 2018, 1107).

Afin de répondre au défi épistémique énoncé, Enoch avance une explication par un « troisième facteur ». Un tel argument comporte que la corrélation entre les croyances normatives et les vérités morales indépendantes serait expliquée par un facteur commun, le « troisième

⁵² Ces auteurs ne sont toutefois pas les seuls à avoir proposé une explication par un « troisième facteur », ou toute autre explication similaire dans la forme. En plus des auteurs présentés, on pensera à Michael Huemer (2008, 218–19), David Copp (2008, 186–206), Karl Schafer (2010, 471–88), Kevin Brosnan (2011, 51–64), William J. Talbott (2015, 691–708), Marc Artiga (2015, 3357–74), Andrew Moon (2017, 208–26) et Dan Baras (2017), notamment. La proposition de Huemer n'a pas été étudiée ici, car elle est assez peu développée. Cela dit, Huemer a présenté un argument plus élaboré contre les anti-réalistes dans un article intitulé « A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism ». En l'occurrence, il prétend montrer que les anti-réalistes sont incapables d'expliquer la libéralisation des valeurs à travers l'histoire (Huemer 2016).

⁵³ Si les deux facteurs corrélés sont A et B, il existe deux types d'explications pouvant rendre compte de la corrélation, et ce, sans faire intervenir un tiers facteur. D'une part, nous avons que le facteur A est causalement, ou constitutivement, responsable du facteur B. D'autre part, nous avons que le facteur B est causalement, ou constitutivement, responsable du facteur A. (C'est la relation converse.) (Enoch 2010, 429)

facteur⁵⁴ ». On comprendra alors que la relation entre croyances et vérités est médiée par ledit « troisième facteur ». En ce sens, il n'existerait aucun lien direct entre les deux entités.

Conscient qu'il existe plusieurs types d'explications par un « troisième facteur », Enoch opte pour une version proprement leibnizienne de l'argument, laquelle renvoie au concept « d'harmonie préétablie ». En l'occurrence, Enoch suppose une « harmonie préétablie » entre les croyances normatives et les faits moraux. Cette relation non causale repose sur une prémisse substantive, laquelle fait office de « troisième facteur ».

Pour Enoch, les contenus « visés⁵⁵ » par l'évolution, comme la survie et le succès reproductif, sont minimalement bons, et certainement meilleurs que les contenus qui ne sont pas visés, comme la mort et l'extinction (Enoch 2010, 430–31). Puisque les pressions darwiniennes visent des contenus moralement bons, il s'ensuit que les croyances normatives façonnées par les forces évolutives sont corrélées avec les vérités morales indépendantes. Autrement dit, le fait que ces contenus évolutifs soient moralement bons garantit une harmonie préétablie entre croyances et vérités morales. Cette explication par « l'harmonie préétablie » implique que les visées évolutives causent nos croyances morales (1) et que ces visées sont constitutivement reliées aux vérités morales (2) (Enoch 2010, 431–32).

Maintenant, Enoch reconnaît que la relation constitutive unissant les visées évolutives aux vérités morales est drôlement heureuse, sinon miraculeuse. Or, ce « miracle » a le bénéfice épistémique d'expliquer une corrélation entre un grand nombre de croyances et de vérités morales, affirme l'auteur (Enoch 2010, 433). Et puis, si la relation postulée n'est pas d'une robustesse contrefactuelle totale, elle n'est pas sans robustesse contrefactuelle. En ce sens, il existerait une pluralité de mondes possibles et rapprochés pour lesquels les forces évolutives, bien que légèrement différentes, auraient engendré des jugements normatifs plutôt bien arrimés aux vérités morales (Enoch 2010, 434). Centrés sur la survie, la reproduction ou tout autre bien évolutif, ces jugements normatifs agiraient comme de véritables points d'ancrage pour notre système de croyances morales. Étant étroitement reliés à des biens afférents, comme « la douleur est mal » et « la

⁵⁴ Si les deux facteurs corrélés sont A et B, nous avons qu'un « troisième facteur », le facteur C, explique la corrélation entre les facteurs A et B. Ici, le facteur C est responsable, simultanément, des facteurs A et B (Enoch 2010, 430).

⁵⁵ Enoch se défend toutefois de prêter un *télôs* à l'évolution (Enoch 2010, 430).

coopération est bien », ces points d’ancrage suffiraient pour expliquer une corrélation assez étendue entre croyances et faits normatifs (Enoch 2010, 431, 435).

Pour toutes ces raisons, Enoch juge que l’explication par un « troisième facteur » – édition « harmonie préétablie » – constitue une réponse plausible au défi épistémologique qui attend le réaliste robuste. Plus spécifiquement, cette relation permettrait de renverser l’argument de Street, en présentant l’évolution comme un facteur garant d’une corrélation entre croyances et vérités normatives.

3.2.2 Wielenberg : certaines facultés cognitives comme « troisième facteur »

S’inscrivant dans le sillon d’Enoch, le philosophe Erik J. Wielenberg propose une explication par un « troisième facteur », laquelle concerne certaines « facultés cognitives » (Wielenberg 2010, 450). Ces facultés cognitives, occupant le rôle de « troisième facteur », seraient causalement responsables de notre croyance en certains droits moraux inaliénables, et la présence de ces mêmes facultés impliquerait logiquement l’existence de ces droits moraux. Ainsi, les facultés cognitives pertinentes garantiraient une corrélation entre nos croyances morales et les faits moraux (Wielenberg 2010, 459).

Façonnées par les forces évolutives, lesdites « facultés cognitives » seraient des mécanismes fiables pour la formation des croyances morales puisque l’évolution favorise des créatures entretenant des croyances vraies (1) et puisque les mécanismes sélectionnés par l’évolution, comme la perception et la mémoire, sont historiquement fiables (2) (Wielenberg 2010, 450–51). Pour Wielenberg, ces observations empiriques sont suffisantes pour soutenir l’hypothèse que certaines croyances morales sont produites par des facultés fiables et que ces croyances possèdent une qualité épistémique nommée « garantie » (*warrant*). Par « garantie », Wielenberg entend toute propriété qui permet à une croyance vraie d’accéder au statut de connaissance. L’auteur considère que la fiabilité est suffisante pour obtenir la « garantie épistémique » (Wielenberg 2010, 447–48). Notons que Wielenberg ne conçoit pas les croyances morales comme des perceptions de faits moraux (externes), ni comme des projections attitudeles (internes). Sans spécifier la nature exacte des mécanismes impliqués, Wielenberg maintient que nos croyances morales sont les produits de mécanismes fiables, ce pourquoi elles pourraient constituer de la connaissance (Wielenberg 2010, 453).

Ayant tracé les contours de son modèle, Wielenberg peut désormais confronter l'argument de Street. L'auteur comprend le dilemme en termes de connexions causales : la première branche suppose que les croyances morales ne sont pas causalement reliées aux faits moraux objectifs, celles-ci étant plutôt déterminées par des contingences évolutives. La deuxième branche suppose que les croyances morales sont causalement reliées aux faits moraux objectifs, en vertu de l'hypothèse du pistage. Wielenberg se prononce sur la première branche du dilemme, car son explication ne suppose aucun lien causal entre croyances et vérités morales. Il n'a donc pas besoin de se soucier de la seconde.

Selon Wielenberg, il suffit d'invoquer une explication par un « troisième facteur » pour éviter le scepticisme normatif. En l'occurrence, l'explication favorisée comporte que:

[...] the presence of the very cognitive faculties that cause (or at least causally contribute to) my belief that I have certain rights also entails that I have those very rights. Furthermore, the process that produces this belief is a reliable one, and hence the belief constitutes knowledge. Something similar is true of my belief that other people have the same rights that I have. Those others have certain nonmoral features that make them relevantly like me. Those features both causally contribute to my belief that they have rights and entail that they have those same rights. And, once again, the process that produces this belief is a reliable one. (Wielenberg 2010, 458)

Si l'on accepte une telle explication, il en découle que la corrélation entre croyances et faits moraux est assurée, sans même qu'il existe une relation causale directe entre les deux classes d'entités. La charge du sceptique serait *de facto* neutralisée. Or, on pourrait être tenté de rejeter une telle explication, puisque celle-ci postule une relation suspecte entre des facultés cognitives contingentes et des faits moraux indépendants (Wielenberg 2010, 459). N'est-ce pas là une autre coïncidence fortuite? Wielenberg reconnaît l'existence d'une telle coïncidence, mais soutient que cette coïncidence n'est qu'apparente. Afin de répondre aux critiques, l'auteur propose d'abord une analyse originale du problème et entrevoit ensuite trois solutions potentielles, bien que spéculatives. Au terme de son exposé, il n'adoptera aucune solution spécifique.

Pour Wielenberg, le défi est d'expliquer la conjonction entre deux classes de faits : les uns causaux et les autres moraux. Sous-ensemble des lois naturelles, les faits causaux concernent des relations de la forme « ceci cause cela » tandis que les faits moraux concernent des relations de la forme « cette propriété morale survient à cette propriété non morale » (Wielenberg 2010, 459).

Dans le cas présent, les faits causaux désignent la relation de dépendance entre nos facultés cognitives et nos croyances morales tandis que les faits moraux désignent la relation de survenance unissant les faits moraux aux facultés cognitives (Wielenberg 2010, 463). Comment donc expliquer la corrélation entre ces deux catégories de faits (Wielenberg 2010, 459–60)? Wielenberg envisage trois solutions :

La première est celle d'un Dieu personnifié, lequel aurait intentionnellement aligné les causes naturelles avec les faits moraux non naturels, de manière à assurer la fiabilité de nos processus cognitifs. La seconde est celle d'une loi universelle (non personnifiée) garantissant la corrélation entre les causes et les faits moraux. La troisième comporte que les causes et les faits moraux postulés sont des vérités nécessaires de l'univers. Si tel est le cas, la correspondance souhaitée serait aussi une vérité nécessaire de l'univers. Wielenberg admet que ces trois explications sont « spéculatives » et, pour cette raison, qu'elles peuvent sembler insatisfaisantes, mais l'auteur maintient que celles-ci fournissent une réponse plausible au problème de la coïncidence fortuite. Ainsi, le réaliste peut adopter la première branche du dilemme darwinien sans être contraint au scepticisme normatif. Selon toute vraisemblance, l'existence de facultés cognitives, en tant que « troisième facteur », nous permettrait d'expliquer la corrélation entre nos croyances morales et les faits moraux correspondants. C'est, du moins, ce que Wielenberg prétend (2010, 461).

3.2.3 Skarsaune : le plaisir et la douleur comme « troisième facteur »

Knut Olav Skarsaune déploie, à son tour, une explication par un « troisième facteur », laquelle repose sur l'idée que « le plaisir est habituellement bien et la douleur est habituellement mal » (Skarsaune 2011, 232–33). Adoptant une formulation du genre « si-donc », l'argument central de Skarsaune se lit comme suit : « S'il est vrai que le plaisir est habituellement bien et la douleur est habituellement mal, on peut inférer qu'il existe effectivement une relation entre les forces évolutives et les faits évaluatifs (1) ET que cette relation conduit à la vérité (*is truth conducive*), dans la mesure où elle biaiserait nos croyances vers la vérité (2). » (Skarsaune 2011, 233) Cet argument central est suppléé par un argument périphérique, censé montrer que la relation (1) et l'effet épistémique (2) sont suffisamment robustes pour permettre au réaliste d'éviter le scepticisme normatif (Skarsaune 2011, 233).

Penchons-nous d'abord sur l'argument central. Cet argument suppose que la relation entre pressions évolutives et faits évaluatifs est médiée par un « troisième facteur », le plaisir et la douleur comme bien et mal. Dans sa présente forme, cet argument n'est toutefois pas complet, puisque le lien entre évolution et états affectifs (plaisir/douleur) demeure inexpliqué. Afin de pallier cette lacune, Skarsaune propose une explication provisoire: « Evolution has caused us to value reproductively beneficial things by making us such that we take pleasure in these things, and caused us to disvalue reproductively harmful things by making us such that these things cause us pain. » (Skarsaune 2011, 233) Plus spécifiquement, Skarsaune prétend que notre capacité à *ressentir* et à être *motivé* par le plaisir ou la douleur constitue un produit de l'évolution. Ce mécanisme, assure-t-on, aurait permis aux humains de sauvegarder des données cruciales sur les vertus et vices évolutifs de certains faits ou états dans le code génétique. À cette prédisposition émotionnelle et motivationnelle s'ajoute une tendance à *valoriser* et *juger* comme bien/mal certains états de choses. Opérant une distinction entre dispositions motivationnelles et croyances morales, ce modèle à deux niveaux s'aligne tout à fait avec la proposition de Street, voulant que l'évolution ait *directement* influencé nos attitudes primitives et *indirectement* influencé nos jugements moraux articulés. Ayant expliqué le lien causal entre évolution et croyances morales, Skarsaune est désormais capable de compléter son argument central, concrétisant ainsi la triade « évolution – plaisir – bien » (Skarsaune 2011, 234–235). Cela dit, on notera que l'argument de Skarsaune est seulement conditionnel. En ce sens, le « troisième facteur » proposé, soit « le plaisir comme bien », est simplement une clause hypothétique et non un fait démontré. Cela ne signifie pas pour autant que l'argument de Skarsaune est vain. Au contraire, cet argument prouve que le réaliste n'est pas limité à l'*hypothèse du pistage* s'il opte pour la deuxième branche du dilemme (Skarsaune 2011, 235). Satisfait de sa démonstration, Skarsaune ne déclare toutefois pas victoire sur Street. Reconnaissant que le mécanisme du plaisir n'explique pas toutes nos croyances morales, Skarsaune envisage la possibilité que d'autres mécanismes non fiables viennent contaminer nos croyances. Ceci condamnerait le réaliste au scepticisme normatif.

Passons maintenant à l'argument périphérique, lequel permettrait au réaliste d'éviter la conclusion pessimiste impliquée par la première branche du dilemme. Cet argument est composé de deux sous-arguments. Le premier affirme que le mécanisme du plaisir et de la douleur supplante, en importance, tout autre mécanisme envisageable. Parmi ces mécanismes non fiables, on recense la réciprocité, l'altruisme et le favoritisme endogroupe (Skarsaune 2011, 235–36). Skarsaune

explique que ces tendances évaluatives innées font l'objet de profonds désaccords, et que leur portée effective demeure modeste. Si tel est le cas, il s'ensuit que le réaliste n'a pas besoin de s'inquiéter de l'influence délétère de ces processus secondaires. La seule influence significative proviendrait du mécanisme du plaisir et de la douleur, mécanisme dont les vertus épistémiques sont assurées⁵⁶. Étant plus général, le second argument affirme que le réalisme est préférable à l'antiréalisme, car le réalisme prédit un taux d'erreur plus élevé que l'alternative. Comme elles sont plus exigeantes, les épistémologies réalistes seraient également plus plausibles que leurs contreparties antiréalistes, lesquelles inviteraient des résultats souvent contre-intuitifs, et parfois même désastreux. Centrée sur le mécanisme du plaisir et de la douleur, l'épistémologie réaliste préférée par Skarsaune réussit à accommoder, systématiser et expliquer nos intuitions morales, évitant ainsi les conséquences indésirables d'un constructivisme ou d'un scepticisme, suppose-t-on. Pour ces deux raisons, Skarsaune estime qu'il est capable de réfuter Street sur son propre terrain (Skarsaune 2011, 236).

Au final, si le réalisme moral est vrai, il est vrai en vertu de la condition énoncée par Skarsaune, soit « le plaisir comme bien et la douleur comme mal ». Expliquant le lien entre évolution et vérité morale, ce « troisième facteur » permettrait au réaliste d'adopter la seconde branche du dilemme darwinien, sans avoir à recourir à l'*hypothèse du pistage*, une avenue décidément problématique.

3.2.4 Discussion sur les explications par un « troisième facteur »

Bien qu'elles constituent des instruments dialectiques ingénieux, les explications par un « troisième facteur » sont susceptibles à un barrage d'objections, ma foi, redoutables. Plusieurs auteurs, dont Street et Dyke, clament que ces explications sont d'une circularité maligne. En ce sens, ils prétendent que les arguments par un tiers facteur sont parfaitement vains, car ils reposent sur des processus épistémiques irrecevables. D'autres, comme Berker et Vavova, affirment que cette circularité est en fait bénigne, sinon vertueuse. Ceux-ci soutiennent qu'il est nécessaire

⁵⁶ À l'inverse, si le plaisir n'était pas bien et la douleur n'était pas mal, à la fois subjectivement et objectivement, une vaste majorité de nos croyances morales auraient été fausses. Ce résultat sceptique est impliqué par la première branche du dilemme darwinien. Or, Skarsaune affirme que ceci n'est pas un problème pour sa théorie, car celle-ci n'a pas besoin de garantir la vérité du réalisme moral pour les deux branches du dilemme. Une seule suffit (Skarsaune 2011, 236–38).

d'assumer certains faits normatifs pour évaluer la fiabilité de la faculté morale. Au cours de cette section, il sera possible de présenter ce débat, avec un certain détail. En l'occurrence, nous constaterons qu'il y a impasse sur le statut épistémique des explications par un « troisième facteur ». Au même moment, nous verrons comment l'intuitionnisme éthique de Huemer permet d'éclairer le débat.

Dans l'œuvre de Street, on retrouve au moins deux réponses potentielles aux arguments avancés par les réalistes non naturalistes. La première, laquelle a été abordée plus haut, soutient qu'il est possible de réappliquer le dilemme darwinien à un niveau supérieur et la seconde consiste à dire que les explications par un « troisième facteur » ont la forme logique d'une pétition de principe, d'un argument circulaire (dans Berker 2014, 227–32)⁵⁷.

D'une part, Street soutient que les explications par un « troisième facteur » ne résolvent pas le dilemme darwinien; elles ne feraient que déplacer le problème à un niveau supérieur (dans Berker 2014, 230–31). Pourquoi en serait-il ainsi? Lorsque le réaliste invoque une telle explication, il sélectionne un « troisième facteur », constituant un fait moral de premier ordre, pour expliquer la corrélation entre des faits contingents et des faits moraux indépendants. Aussi plausible soit-il, ce « troisième facteur » n'est pas couvert d'une « immunité spéciale », lui permettant de défléchir automatiquement l'offensive sceptique. Sans exception, le réaliste est sommé de prouver que le facteur sélectionné résiste au dilemme darwinien. Encore une fois, deux choix s'offrent à lui : nier ou affirmer une relation entre les forces évolutives contingentes et les faits moraux indépendants. La première option le contraint au scepticisme tandis que la seconde le condamne à une explication scientifique intangible. Visiblement, l'explication par un « troisième facteur » ne permet pas au réaliste de répondre au dilemme darwinien, mais seulement de déplacer ce dilemme à un niveau supérieur. On comprendra, par ailleurs, que ce jeu argumentatif entre le réaliste et le sceptique peut se répercuter à l'infini.

D'autre part, Street soutient que les explications par un « troisième facteur » sont irrecevables, car elles sont basées sur une forme de raisonnement invalide, la pétition de principe (dans Berker 2014, 231–32). Une telle faute logique sera imputée à tous ceux et celles qui posent

⁵⁷ Une telle critique s'applique aussi aux réponses mooréennes ou aux réponses qui font appel au sens commun, plus généralement (Koon 2021, 12170–71; del Rio 2021, 7439–40). On retrouve un tel genre de réponse dans Cuneo et Shafer-Landau (2014).

comme prémisses ce qu'ils prétendent prouver. Figurant parmi ces délinquants, les réalistes qui se rabattent sur les explications par un « troisième facteur » fournissent une réponse strictement inadéquate au dilemme darwinien, réitère Street :

It is no answer to [the Darwinian] challenge simply to assume a large swath of substantive views on how we have reason to live [...] and then note that these are the very views evolutionary forces pushed us toward. Such an account merely trivially reasserts the coincidence between the independent normative truth and what the evolutionary causes pushed us to think; it does nothing to explain the coincidence. (Street 2008, 214)

Relevant de la sophistique, les explications par un « troisième facteur » sont dialectiquement stériles, car elles ne montrent pas pourquoi les forces évolutionnaires coïncident avec les vérités morales. Elles font seulement réaffirmer, à l'aide de prémisses normatives reçues et acceptées, une telle coïncidence. Plus encore, une telle stratégie argumentative est condamnable, parce qu'elle ne peut pas échouer! Street l'affirme en ces mots :

[S]uppose you hold certain views about how you have reason to live (and you think there are independent truths about such matters). Then one day you learn that you acquired these normative views in large part by having had them shaped by evolutionary forces. You're concerned that this might not have been a reliable method for arriving at them. In answer to that concern, it is no help to repeat your views about how you have reason to live and then point out that those are the very same ones that evolutionary forces shaped you to have. *No method of arriving at your views about how you have reason to live – no matter how bizarre and unreliable – could fail this test, since if having led you to your actual views is the test of a method's reliability, then whatever method you actually used will come out as reliable* [mes italiques]. (Street 2008, 215)

En d'autres mots, les explications par un « troisième facteur » sont illégitimes, car elles supposent que *toute méthode ayant produit nos croyances est une méthode fiable*. Si tel était le cas, il suffirait d'avoir une croyance pour que celle-ci soit justifiée. Or, on sait très bien que le descriptif n'implique pas le normatif, et qu'un tel critère épistémique est simplement trop généreux. En effet, un critère qui ne discrimine point mérite-t-il véritablement une telle désignation? De toute évidence, le réaliste peut faire mieux, et il doit faire mieux, s'il prétend expliquer la corrélation entre nos croyances morales et les faits moraux indépendants.

Néanmoins, le réaliste pourrait rétorquer que les explications par un « troisième facteur » ne font pas que réaffirmer nos convictions morales, elles fournissent une explication plausible et

plutôt robuste du lien entre croyances et vérités morales. Or, il appert que chacune des explications fournies par Enoch, Wielenberg et Skarsaune, respectivement, laisse une place au doute et à la spéculation. Bien entendu, la philosophie n'est pas une science exacte, ni un moulin à certitudes, mais elle exige tout de même un certain niveau de rigueur épistémique qui ne semble pas être atteint par ces explications, en quelque sorte, circulaires.

Rajoutant à la critique de Street, la philosophe Michelle M. Dyke prétend que les explications par un « troisième facteur » constituent une stratégie argumentative douteuse, voulant que la fiabilité d'une source puisse être attestée par cette même source (*bad bootstrapping*) (Weisberg 2012, 597; Dyke 2020, 2120–22). Malgré ses contours suspects, cette forme de raisonnement est nécessitée par toute théorie qui admet la possibilité de connaissance fondamentale (*basic knowledge*), une connaissance obtenue avant que la fiabilité de la source de cette connaissance ait été établie (Weisberg 2012, 598). S'appuyant sur de récentes contributions académiques, Dyke recense quatre manières distinctes d'identifier les théories basées sur une telle stratégie : la présence de règles circulaires (1), un usage inadéquat des lemmes (2), une absence de risque (3) et une impossibilité de défaite épistémique (4) (Dyke 2020, 2122–27). Dyke présente ces quatre méthodes séquentiellement, et souligne que, pour chacune d'entre elles, les explications par un « troisième facteur » constituent une violation de la bonne pratique épistémique. Étant d'une étendue et d'une complexité considérables, l'argument de Dyke ne pourra, au cours des prochaines lignes, être reproduit dans son intégralité. Toutefois, il sera possible de reprendre le premier élément de son argument, lequel rassemble plusieurs points d'importance.

Comme on l'a mentionné, l'usage indu d'une « règle circulaire » témoigne d'une pratique épistémique déviante. Afin d'éviter de tels écarts de conduite, la contrainte suivante peut être énoncée :

A belief that an epistemic rule *R* is reliable cannot be justified by the application of *R*. That is, neither the conclusion itself nor any belief which supports the conclusion may be justified in virtue of the application of *R*. (Dyke 2020, 2122)

En l'occurrence, on suggère que la croyance dans la fiabilité d'une règle épistémique donnée ne peut être justifiée par l'application de cette même règle. Selon Dyke, c'est précisément ce que font Enoch, Wielenberg et Skarsaune, lorsqu'ils avancent une explication par un « troisième facteur ». Pourquoi? Car ils présupposent la fiabilité de notre faculté morale, eu égard aux vérités

correspondantes. Or, comme le remarquent Weisberg (2012, 599) et Dyke (2020, 2122–23), cette contrainte épistémique est probablement trop forte, car elle interdirait des formes de circularité désirables. En effet, comment vérifier la fiabilité de notre mémoire sans faire intervenir cette même faculté? Plus encore, comment vérifier la fiabilité de notre perception visuelle sans faire intervenir celle-ci? N'étant pas seulement incontournable, il semblerait que la circularité puisse aussi être vertueuse! Qui l'eût cru? Maintenant, il s'agit de déterminer si les explications par un « troisième facteur » peuvent, à leur tour, être vertueuses dans leur circularité.

Convaincus par l'argumentaire d'un Selim Berker (2014, 248–50) et d'une Katia Vavova (2014, 82–84), plusieurs philosophes⁵⁸ soutiennent qu'il est impossible d'établir la fiabilité d'une faculté sans faire intervenir des croyances engendrées par cette même faculté. Assumant que cette règle s'applique à la faculté morale, chose dont nous n'avons aucune raison de douter (*a priori*), il nous est permis d'inférer deux propositions :

1. D'une part, si l'on affirme qu'il est *légitime* de présupposer la vérité de propositions dans quelque domaine de connaissance que ce soit, alors il est *légitime* de présupposer la vérité de propositions normatives du domaine moral.

Par contraposition

2. D'autre part, si l'on affirme qu'il est *illégitime* de présupposer la vérité de propositions normatives du domaine moral, alors il est *illégitime* de présupposer la vérité de propositions dans quelque domaine de connaissance que ce soit.

Constituant une implication triviale de la forme « si A alors B », la première proposition comporte que le réaliste peut se réclamer d'une explication par un « troisième facteur », et ce, en toute légitimité. Constituant une implication de la forme « si non-B alors non-A », la seconde proposition comporte que le sceptique ne peut nier la légitimité des explications par un « troisième facteur » sans sombrer dans un scepticisme généralisé. En ce sens, si l'on suppose que la faculté morale ne diffère en aucun point épistémiquement significatif des autres facultés (comme la perception sensible, l'introspection et la mémoire), nous pouvons inférer qu'il est impossible d'être

⁵⁸ Parmi ceux-ci figurent Matt Lutz (2018, 1107–08, 1120–22), Michael Klenk (2017, 228; 2018, 430–31), Folke Tersman (2017, 766; 2022, 2, 11–12) et Daniel Z. Korman (2019, 345–48; 2020, 314–16).

sceptique pour la moralité, sans être sceptique pour tous les autres domaines de la connaissance. C'est le problème de la *surgénéralisation*.

Or, les ennuis du sceptique ne s'arrêtent pas là! Si le sceptique rejette toute prémisse morale substantive, il ne sera pas en mesure de montrer que nous avons tort à propos des contenus de la moralité. Comme Vavova l'affirme :

If we cannot make any substantive assumptions about particular moral norms, then morality could (conceptually) be about anything. But if morality could be about anything, then we have no idea what morality is about. So we have no idea if evolutionary forces would have pushed us toward or away from the truth. *So we have no reason to think we are mistaken* [mes italiques]. (Vavova 2015, 112)

En effet, si nous n'avons aucune idée sur les contenus de la morale, comment savoir que nous avons tort à propos de la morale? Comment savoir que l'évolution nous a guidés dans la mauvaise direction? Visiblement, le sceptique doit faire certaines concessions (Vavova 2015, 107–13; Vavova 2021, 723). Mais quelles sont la nature et l'étendue de ces concessions? On pourrait penser que cet argument contraint le sceptique à reconnaître *la vérité* de certaines prémisses normatives. Or, rien n'est plus faux. Cet argument ne contraint pas le sceptique à des concessions ontologiques, mais seulement à des concessions épistémologiques modestes. En ce sens, le sceptique doit reconnaître que nos croyances morales jouissent d'une justification *prima facie* (*default entitlement*). Néanmoins, une question centrale demeure : est-ce que cette justification *prima facie* est défaite par l'étiologie profonde de nos croyances morales? Il est difficile de répondre à cette question, car on ne sait pas ce qui constitue véritablement un processus fiable pour la formation des croyances morales. Mais encore, on ne sait même pas comment vérifier, de manière indépendante, la corrélation de nos croyances morales avec les faits moraux postulés! Incapables de déterminer ce qui est une source de *confirmation* ou de *défaite* pour nos croyances morales, nous sommes condamnés à faire dialoguer des apparences *prima facie* justifiées, sans réellement savoir celles qui sont défaites et celles qui ne le sont pas⁵⁹. C'est là une morale irrémédiablement provisoire!

⁵⁹ Toutefois, cela n'exclut pas que des critères formels comme la cohérence ou l'accord intersubjectif demeurent des facteurs épistémiquement désirables.

Cette discussion nous ramène à l'intuitionnisme éthique de Huemer, théorie épistémologique voulant que nos intuitions morales soient *prima facie* justifiées. Cette théorie constitue une avenue prometteuse en philosophie morale, car elle accorde suffisamment au réaliste, sans contrarier le sceptique. Plus exactement, l'intuitionnisme huémérien prévoit que les intuitions morales sont présumées innocentes (*default entitlement*), tout en demeurant défaisables. On voit aisément comment une telle approche permet d'éviter un scepticisme généralisé, sans toutefois empêcher un scepticisme restreint. Cela dit, l'orientation rationaliste de la théorie demeure un point de discorde, lequel sera traité plus bas.

3.3 Intuitionnisme huémérien et critique généalogique évolutionnaire : quelles implications?

Jusqu'ici, nous avons abordé la critique généalogique évolutionnaire de Street, ainsi qu'une réponse typiquement envisagée par les réalistes, soit l'explication par un « troisième facteur ». Différentes versions de cet argument ont été explorées, celles d'Enoch, de Wielenberg et de Skarsaune. Si ces explications comportent des faiblesses patentes, elles demeurent philosophiquement pertinentes, dans la mesure où elles contribuent à faire progresser la dialectique entre le réaliste et le sceptique. Effectivement, ces arguments suggèrent qu'il est difficile, voire impossible, d'attester de la fiabilité d'une faculté sans présupposer certaines choses à propos de cette même faculté. En ce sens, il semblerait qu'une défense (*vindication*) et une critique (*debunking*) de nos facultés morales soient, toutes deux, impossibles sans tenir pour justifiés certains contenus normatifs. Une certaine circularité serait, donc, inévitable.

Pourtant, il appert que la critique de Street nous donne ample raison de croire que nos intuitions morales substantives ne sont pas justifiées. *Contra* les réalistes, Street montre qu'il est fort *improbable* que nos intuitions morales – tributaires de processus évolutionnaires contingents – soient alignées avec les vérités morales (Vavova 2015, 112; Street 2016, 305–6). Décidément, si les forces évolutives ont informé le contenu de nos intuitions morales, alors, nous pouvons difficilement croire que ces intuitions sont épistémiquement justifiées. En l'occurrence, on pourrait dire que l'évolution agit comme un défacteur discréditant pour une gamme d'intuitions morales substantives, ou non formelles :

P1. Prémisse doxastique / *explanandum* : S a des croyances morales substantives (P).

P2. Prémisse explicative / explanans : La croyance de S que P est expliquée par l'évolution, un processus arbitraire qui rend *improbable* toute correspondance entre nos croyances et les vérités morales substantives, quelles qu'elles soient.

P3. Prémisse épistémique / normative : L'évolution est un processus qui n'est pas épistémiquement fiable.

C2 : Conclusion : La croyance de S que P n'est pas justifiée.

Encore une fois, il apparaît que les origines distales de nos croyances morales substantives constituent un défaiteur discréditant pour ces mêmes croyances. Par « croyances morales substantives », j'entends toutes ces croyances évaluatives et normatives de la forme : « je crois que la douleur est mal », « je crois que la survie est bien », « je crois que nous avons des droits inaliénables », etc. Allant d'un niveau de généralité minimal à élevé, ces croyances morales seraient défaites par leurs origines évolutives, car leur contenu est informé par l'évolution (prémisse explicative), et l'évolution n'est pas un processus fiable pour l'acquisition de croyances (prémisse épistémique). De même, les jugements moraux concernant la réciprocité, l'altruisme et le favoritisme filial peuvent tous être considérés comme des produits évolutifs. Élevant le particulier au général, ces jugements moraux de niveau intermédiaire sont susceptibles à l'argument généalogique présenté ici, cela ne fait pas de doute.

Néanmoins, on pourrait s'interroger sur la vulnérabilité des jugements substantifs abstraits à ce même argument. Effectivement, peut-on considérer que les éthiques conséquentialiste, déontologique, arétaïque, contractualiste ou chrétienne sont traversées par des influences évolutives épistémiquement perverses? Il est raison de croire que oui. Si le bonheur est valorisé, si l'obéissance aux normes est valorisée, si la vertu est valorisée, c'est bien parce que ces choses sont avantageuses sur le plan reproductif. Il serait inutile d'effectuer une analyse fine pour s'en convaincre. Maintenant, les réalistes pourraient invoquer des explications par un « troisième facteur » afin de montrer que nos croyances substantives abstraites correspondent à la vérité morale. Or, comme nous avons pu le voir précédemment, ces arguments demeurent plutôt spéculatifs, et sont certainement inconcluants. Si la discussion méta-épistémologique sur la validité des explications par un « troisième facteur » est encore ouverte, nous pouvons tout de même retenir l'hypothèse que nos intuitions substantives sont les produits de l'évolution ET que l'évolution est un défaiteur pour ces mêmes intuitions.

Toutefois, on se gardera d'étendre un tel argumentaire aux croyances abstraites formelles. Comme Behrens (2013, 494–95) et Leibowitz (2021, 64) le suggèrent, l'évolution ne nous donne aucune raison d'abandonner des règles évaluatives formelles, ni des intuitions morales métadiscursives, que nous tenons pour vraies. En effet, il est difficile de concevoir le phénomène moral sans les lois quantitatives de la transitivité⁶⁰ et de l'additivité axiologiques⁶¹, ou sans certains faits métadiscursifs comme la survenance du moral sur le non-moral⁶². La relation de nécessitation entre le concept de moralité et ces prédicats est à ce point étroite, qu'il nous serait permis d'invoquer la notion de connaissance analytique *a priori*. Cela revient à dire que la définition de la morale impliquerait logiquement la transitivité, l'additivité, la survenance, etc. Cette relation logique serait d'une telle robustesse contrefactuelle que, dans tous les mondes possibles où le concept de moralité existe, ces propriétés seraient instanciées. Et donc, évolution ou pas, le concept de moralité contiendrait nécessairement certains sous-éléments, en vertu de sa propre définition et par nécessité conceptuelle.

Tout bien considéré, il semblerait que les intuitions morales forment un corps épistémiquement hétérogène. D'une part, nous avons les intuitions morales substantives (non formelles), lesquelles ne résulteraient pas de l'appréhension de vérités rationnelles *aprioriques*. Étant donné leurs origines douteuses, ces intuitions évaluatives et normatives ne seraient pas justifiées. D'autre part, nous avons les intuitions morales formelles et métadiscursives, lesquelles résulteraient de l'appréhension de vérités *aprioriques* nécessaires, tout comme les vérités logiques ou mathématiques. Formées à partir de processus fiables, ces intuitions seraient justifiées, et donc

⁶⁰ Si nous considérons que la morale est un système de quantités axiologiques, nous avons que l'expression « si A est mieux que B, et B est mieux que C » est vraie dans tous les mondes possibles. Stuart Rachels (1998) argumente qu'il existe des exemples d'*intransitivité* en morale, mais Huemer suppose que ceux-ci ne reflèteraient que l'irrationalité des agents moraux (Huemer 2008, 386–87).

⁶¹ Si nous considérons que la morale est un système de quantités axiologiques, nous avons aussi que l'expression « la valeur morale d'un état de fait est égale à la somme de la valeur de ses parties (e.g. le nombre de vies sauvées, la quantité de CO₂ émise dans l'atmosphère, la quantité de plaisir, etc.) » est vraie dans tous les mondes possibles. Pour être clair, il est possible de concevoir des systèmes de quantités non additifs, mais il semble que la moralité ne puisse être un tel système. Huemer aborde la question dans son article « Revisionary Intuitionism » (2008, 384).

⁶² Le principe de survenance du moral sur le non-moral veut que « deux états de fait identiques en termes non moraux soient aussi identiques en termes moraux ». Cette intuition métadiscursive, ou métaéthique, est largement admise. D'autres intuitions conceptuelles comme « la morale porte sur le bien et le mal » ou « la qualité morale d'un état de fait dépend de notre critère moral » sont aussi largement admises, et ce, pour des raisons évidentes. Or, ces intuitions métaphysiques ne disent rien sur le contenu des énoncés normatifs eux-mêmes. Elles décrivent simplement le fonctionnement de la morale.

non défaites, au sens préconisé par Huemer. Or, le réaliste non naturaliste peut-il véritablement se contenter d'une morale de signes? Rien n'est moins certain.

Conclusion

L'intuitionnisme rationaliste de Huemer affirme que nous avons des intuitions morales et qu'une *partie* de ces intuitions morales, étant le produit de mécanismes épistémiquement faibles, constituent de la connaissance synthétique *a priori*, tout comme les vérités métaphysiques et mathématiques (Huemer 2016, 1986–88, 2003, 2007). Reconnaisant la possibilité qu'une majorité d'intuitions morales soient effectivement défaites, Huemer prétend toutefois que les intuitions abstraites formelles sont d'excellents candidats à titre de vérités morales aprioriques (Huemer 2008, 383–87). Est-ce que de telles thèses sont corroborées par l'analyse proposée ici?

Interrogeant les causes *proximales* et *distales* de nos intuitions morales, la critique déployée ci-haut montre que les intuitions morales substantives (ou non formelles) sont largement défaites par les arguments étiologiques. Cette enquête montre également que les intuitions morales formelles ne sont pas défaites par ces mêmes arguments. Une telle asymétrie suggère qu'il existe des différences épistémiquement significatives entre les intuitions morales substantives et formelles, même si leur phénoménologie demeure similaire. En effet, les différences étiologiques notées indiquent que les intuitions substantives (concernant l'assignation des prédicats moraux à certains états de fait) sont influencées par des contenus non cognitifs, et que les intuitions formelles (concernant les relations entre les quantités morales) ne sont pas influencées par de telles contingences. À cet égard, il nous est permis de rapprocher les intuitions morales formelles aux intuitions métaphysiques ou mathématiques. Or, rien, mais absolument rien, n'indique que ce rapprochement est justifié dans le cas des intuitions substantives. En ce sens, les intuitions morales substantives seraient sujettes à des problèmes spécifiques, lesquels n'affectent pas les intuitions métaphysiques ou mathématiques.

Comme nous avons pu le constater, les processus non cognitifs influencent largement le contenu de nos intuitions morales. Ces processus affectifs et conatifs constituent des défaiteurs pour les croyances morales substantives, étant présupposé que la cognition soit la seule source de connaissance reconnue par Huemer. Si l'auteur mentionne au passage les travaux de Michael Tye et Graham Oddie dans un récent article, il demeure peu bavard sur le potentiel épistémique des

fonctions affectives et conatives, lesquelles peuvent constituer des mécanismes fiables pour la formation des croyances morales (Huemer 2016, 1985). Similairement, l'auteur n'explique pas si et comment son intuitionnisme rationaliste peut rendre compte des vertus épistémiques des émotions, en tant que représentations évaluatives. À ce titre, on pourrait se demander : quelle place pour les émotions dans l'intuitionnisme huémérien? Plus encore, pourquoi s'en tenir à la pure raison, si les émotions peuvent aussi constituer des sources de connaissance morale fiable? Maintenant, si Huemer devait accorder une place de choix aux émotions dans son épistémologie morale, il devrait concurremment reconnaître que les intuitions morales diffèrent d'une manière épistémiquement significative des intuitions métaphysiques et mathématiques, lesquelles sont entièrement étrangères aux considérations affectives et conatives. Encore une fois, on constate que l'intuitionnisme huémérien repose sur une psychologie morale hautement schématique, laquelle ne peut rendre justice à la complexité et à la variété des procédés impliqués dans la formation du jugement moral. Une analyse plus profonde nous permettrait d'expliquer pourquoi certains types d'apparences seraient plus fiables que d'autres.

Cela dit, on aura compris que les origines proximales de nos croyances morales ne sont pas les seuls facteurs qui pèsent dans la balance épistémique. Encore faut-il considérer les origines distales de ces mécanismes proximaux, soient-ils cognitifs ou non cognitifs. La thèse défendue ici implique que les jugements moraux substantifs sont causés par des mécanismes proximaux, lesquels résultent à leur tour de forces évolutives contingentes. Ne pouvant pister les faits moraux, ces pressions évolutives ne garantiraient d'aucune manière la fiabilité de nos jugements moraux. De même, il semblerait que toute corrélation entre ces jugements et les faits correspondants serait purement accidentelle. Afin de faire sens de cette corrélation (possible), plusieurs ont avancé des explications par un « troisième facteur ». Si ces explications apparaissent foncièrement circulaires, elles montrent néanmoins qu'il est nécessaire d'accorder une justification *prima facie* à nos jugements moraux substantifs. Maintenant, la question est de savoir s'il existe des raisons *quelconques* pour considérer que l'ensemble des jugements moraux favorisés par l'évolution intersecte l'ensemble des faits moraux postulés. J'ai prétendu que de telles raisons ne sont pas en notre possession, et qu'il était plus probable que l'évolution soit un défaiteur discréditant pour nos croyances qu'une source de confirmation pour ces mêmes croyances. Mais comment sais-je qu'il est plus probable que l'évolution soit une source de défaite que de confirmation épistémique? En toute franchise, je ne le sais pas. Par contre, la nécessité de recourir à des arguments aussi suspects

que les explications par un « troisième facteur » pour dissoudre les arguments évolutionnaires suggère que notre situation épistémique est bien piètre. Ultimement, nous sommes forcés de reconnaître que nous n'avons aucune idée des contenus substantifs de l'objet moral. Ceci nous ramène au défi posé par Enoch : comment expliquer la corrélation entre nos croyances morales et les faits moraux correspondants? Même si le réaliste produisait une explication plausible, il serait sommé d'expliquer pourquoi son explication serait plus plausible que celle d'autrui, et ainsi à l'infini. Décidément, le problème de fond est que nous ne savons pas sous quelles conditions épistémiques appliquer ou retirer le prédicat de justification à nos croyances morales. Cela dit, s'il devait exister une moralité indépendante de nos attitudes évaluatives, nous avons raison de croire que cette moralité serait régie par certaines contraintes métadiscursives, comme la survenance du moral sur le non-moral. Similairement, nous avons raison de croire que cette morale contiendrait certaines règles formelles, nous permettant de manipuler des quantités évaluatives. Il serait difficile de faire sens de la morale sans ces quelques éléments.

Après cette tempête d'idées, que reste-t-il de l'intuitionnisme huémérien? S'agit-il d'une solution tangible au scepticisme épistémologique moral? Si la théorie huémérienne ne protège pas nos intuitions morales contre une éventuelle défaite, elle permet néanmoins d'éviter un scepticisme moral généralisé. Comme Berker et Vavova l'ont également suggéré, il est nécessaire de faire certaines concessions, sans quoi il serait impossible de faire de l'éthique normative. À cet égard, le conservatisme phénoménal atteint sa cible. La théorie est simple, élégante, et ses vertus épistémiques sont claires. Là où le bât blesse, c'est que Huemer calque son épistémologie morale sur l'épistémologie des langages formels. Non seulement cette thèse rationaliste cadre mal avec les sciences empiriques, elle ne rend pas compte des différences épistémiques saillantes entre les groupes d'intuitions. En ce sens, elle n'explique pas pourquoi, d'un point de vue rationaliste, certains types d'intuitions morales (substantives) seraient largement défaits tandis que d'autres intuitions (formelles) ne le seraient pas. À mon humble avis, cette asymétrie épistémique exprime une différence profonde entre les types d'intuitions morales. En l'occurrence, il s'agit de reconnaître que les intuitions substantives constituent de la connaissance synthétique *a posteriori* tandis que les intuitions formelles/métadiscursives constituent de la connaissance synthétique *a priori* (possiblement analytique). *Contra* Huemer, je considère qu'aucune intuition substantive ne serait purement rationnelle, ce pourquoi on observe une défaisabilité et une défaite épistémiques accrues dans ce secteur de la moralité. Cependant, si l'on sortait du carcan rationaliste imposé par

Huemer, il semble que nous pourrions accommoder une gamme d'intuitions élargie. En effet, si l'on permettait que la justification ne découle pas seulement de la raison pure, mais aussi de l'expérience, peut-être serait-on en mesure d'accéder à certains faits évaluatifs. Comme l'affirme Tappolet, il est raisonnable de croire que les émotions, en tant qu'expériences (quasi-) perceptuelles, constituent une voie d'accès privilégiée aux faits axiologiques (Tappolet 2000; Rossi et Tappolet 2017, 177–83). Si tel est le cas, il serait sensé d'adopter une épistémologie hybride, laquelle nous permettrait de réintégrer certaines intuitions substantives au sein de notre discours, sans toutefois rejeter les intuitions formelles et métadiscursives auxquelles nous tenons. Un tel agencement serait heureux, sachant que ces deux théories perceptualistes reposent sur un principe de justification *prima facie*.

Dans tous les cas, il semblerait que les partisans du réalisme moral ne soient pas entièrement sans recours. En optant pour une théorie comme l'intuitionnisme huémérien, ceux-ci seraient capables d'articuler et de justifier un bon nombre d'intuitions morales, surtout formelles et métadiscursives. Par contre, ils auraient une certaine difficulté à composer avec les intuitions substantives, lesquelles ne proviendraient pas de l'appréhension de vérités *aprioriques*. Pour cette raison, et encore, je considère que l'intuitionnisme huémérien est une solution imparfaite au scepticisme épistémologique moral.

Bibliographie

- Allman, John, et Jim Woodward. 2008. « What Are Moral Intuitions and Why Should We Care About Them? A Neurobiological Perspective ». *Philosophical Issues* 18 (1): 164–85.
- Appiah, Anthony. 2010. *Experiments in Ethics*. Cambridge, MA: Harvard University Press.
- Artiga, Marc. 2015. « Rescuing Tracking Theories of Morality ». *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition* 172 (12): 3357–74.
- Audi, Robert. 2004. *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton, NJ: Princeton University Press.
- Baras, Dan. 2017. « Our Reliability Is in Principle Explainable ». *Episteme* 14 (2): 197–211.
- Baron, Jonathan. 2014. « Heuristics and Biases ». Dans *The Oxford Handbook of Behavioral Economics and the Law*. Sous la direction d'Eyal Zamir et Doron Teichman, 3–27. Oxford: Oxford University Press.
- Baron, Jonathan. 2014. « Moral Judgment ». Dans *The Oxford Handbook of Behavioral Economics and the Law*. Sous la direction d'Eyal Zamir et Doron Teichman, 61–89. Oxford: Oxford University Press.
- Bartels, Daniel M., et David A. Pizarro. 2011. « The Mismeasure of Morals: Antisocial Personality Traits Predict Utilitarian Responses to Moral Dilemmas ». *Cognition* 121 (1): 154–61.
- Bealer, George. 2000. « A Theory of the a Priori ». *Pacific Philosophical Quarterly* 81 (1): 1–30.
- Bedke, Matthew S. 2008. « Ethical Intuitions: What They Are, What They Are Not, and How They Justify ». *American Philosophical Quarterly* 45 (3): 253–69.
- Bedke, Matthew S. 2010. « Intuitional Epistemology in Ethics ». *Philosophy Compass* 5 (12): 1069–83.

- Bedke, Matthew S. 2020. « Moral Intuition ». Dans *The Routledge Handbook of Moral Epistemology*. Sous la direction d'Aaron Zimmerman, Karen Jones et Mark Timmons, 360–74. New York, NY: Routledge.
- Behrends, Jeff. 2013. « Meta-Normative Realism, Evolution, and Our Reasons to Survive ». *Pacific Philosophical Quarterly* 94 (4): 486–502.
- Bengson, John. 2015. « Grasping the Third Realm ». Dans *Oxford Studies in Epistemology, Volume 5*. Sous la direction de Tamar Szabó Gendler et John Hawthorne, 1–38. Oxford University Press: Oxford.
- Berker, Selim. 2009. « The Normative Insignificance of Neuroscience ». *Philosophy & Public Affairs* 37 (4): 293–329.
- Berker, Selim. 2014. « Does Evolutionary Psychology Show That Normativity Is Mind-Dependent? ». Dans *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*. Sous la direction de Justin D'Arms et Daniel Jacobson, 215–52. Oxford: Oxford University Press.
- Brink, David O. 2014. « Principles and Intuitions in Ethics: Historical and Contemporary Perspectives ». *Ethics* 124 (4): 665–94.
- Brosnan, Kevin. 2011. « Do the Evolutionary Origins of Our Moral Beliefs Undermine Moral Knowledge? ». *Biology & Philosophy* 26 (1): 51–64.
- Chudnoff, Elijah. 2011. « The Nature of Intuitive Justification ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 153 (2): 313–33.
- Chudnoff, Elijah. 2011. « What Intuitions Are Like ». *Philosophy and Phenomenological Research* 82 (3): 625–54.
- Clarke-Doane, Justin. 2015. « Justification and Explanation in Mathematics and Morality ». Dans *Oxford Studies in Metaethics, Volume 10*. Sous la direction de Russ Shafer-Landau, 80–103. Oxford: Oxford University Press.

- Clarke-Doane, Justin. 2017. « What is the Benacerraf Problem? » Dans *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity*. Sous la direction de Fabrice Pataut, 17–43. Cham: Springer.
- Copp, David. 2008. « Darwinian Skepticism About Moral Realism ». *Philosophical Issues* 18 (1): 186–206.
- Cosmides, Leda, et John Tooby. 2006. « Evolutionary Psychology, Moral Heuristics, and the Law ». Dans *Heuristics and the Law*. Sous la direction de Gerd Gigerenzer et Christoph Engel, 175–205. Cambridge, MA: MIT Press en coopération avec Dahlem University Press.
- Cuneo, Terence, et Russ Shafer-Landau. 2014. « The Moral Fixed Points: New Directions for Moral Nonnaturalism ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 171 (3): 399–443.
- Cushman, Fiery. 2013. « Action, Outcome, and Value: A Dual-System Framework for Morality ». *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc* 17 (3): 273–92.
- Cushman, Fiery. 2020. « Rationalization Is Rational ». *Behavioral and Brain Sciences* 43.
- Cushman, Fiery, Liane Young, et Marc Hauser. 2006. « The Role of Conscious Reasoning and Intuition in Moral Judgment: Testing Three Principles of Harm ». *Psychological Science* 17 (12): 1082–89.
- De Neys, Wim, dir. 2018. *Dual Process Theory 2.0*. New York, NY: Routledge, Taylor & Francis Group.
- De Neys, Wim, et Michal Bialek. 2017. « Dual Processes and Conflict During Moral and Logical Reasoning: A Case for Utilitarian Intuitions ». Dans *Moral Inferences*. Sous la direction de Jean-François Bonnefon et Bastien Trémolière, 123–36. London: Psychology Press.

- de Oliveira-Souza, Ricardo, Jorge Moll, et Roland Zahn. 2021. « Moral Cognition in Neurology ». Dans *Encyclopedia of Behavioral Neuroscience*. Sous la direction de Sergio Della Sala, 247–53. Amsterdam: Elsevier.
- Demaree-Cotton, Joanna, et Guy Kahane. 2020. « The Neuroscience of Moral Judgment ». Dans *The Routledge Handbook of Moral Epistemology*. Sous la direction d'Aaron Zimmerman, Karen Jones et Mark Timmons, 84–104. New York, NY: Routledge.
- DePaul, Michael R. 2009. « Phenomenal Conservatism and Self-Defeat ». *Philosophy and Phenomenological Research* 78 (1): 205–12.
- Dyke, Michelle M. 2020. « Bad Bootstrapping: The Problem with Third-Factor Replies to the Darwinian Dilemma for Moral Realism », *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition* 177 (8): 2115–28.
- Enoch, David. 2010. « The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 148 (3): 413–38.
- Enoch, David. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.
- Evans, Jonathan St. B. T. 2019. « Reflections on Reflection: The Nature and Function of Type 2 Processes in Dual-Process Theories of Reasoning ». *Thinking & Reasoning* 25 (4): 383–415.
- Evans, Jonathan St. B. T, et Keith E. Stanovich. 2013. « Dual-Process Theories of Higher Cognition : Advancing the Debate ». *Perspectives on Psychological Science* 8 (3): 223–41.
- Field, Hartry H. 1989. *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.
- Fischer, Robert William. 2016. « Disgust As Heuristic ». *Ethical Theory and Moral Practice* 19 (3): 679–93.

- FitzPatrick, William J. 2015. « Debunking Evolutionary Debunking of Ethical Realism ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 172 (4): 883–904.
- Foot, Philippa. 1967. « The Problem of Abortion and the Doctrine of the Double Effect ». *Oxford Review* (5): 5–15.
- Gibbs, John C. 2019. « Beyond Haidt’s New Synthesis ». Dans *Moral Development and Reality: Beyond the Theories of Kohlberg, Hoffman, and Haidt*, 19–44. New York: Oxford University Press.
- Gigerenzer, Gerd. 2008. « Moral Intuition = Fast and Frugal Heuristics? » Dans *Moral Psychology, Vol. 2*. Sous la direction de Walter Sinnott-Armstrong et Christian B. Miller, 1–26. Cambridge, MA: MIT Press.
- Gigerenzer, Gerd, et Wolfgang Gaissmaier. 2011. « Heuristic Decision Making ». *Annual Review of Psychology* 62 (1): 451–82.
- Graham, Jesse, Jonathan Haidt, Matt Motyl, Peter Meindl, Carol Iskiwitch, et Marlon Mooijman. 2018. « Moral Foundations Theory: On the Advantages of Moral Pluralism over Moral Monism ». Dans *Atlas of Moral Psychology*. Sous la direction de Kurt Gray et Jesse Graham, 211–22. New York, NY: Guilford Press.
- Graham, Jesse, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, et Peter H. Ditto. 2013. « Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism ». Dans *Advances in Experimental Psychology, Vol. 47*. Sous la direction de Patricia Devine et Ashby Plant, 55–130. Burlington: Academic Press.
- Greene, Joshua D. 2008. « Reply to Mikhail and Timmons ». Dans *Moral psychology, Vol. 3*. Sous la direction de Walter Sinnott-Armstrong, 105–17. Cambridge, MA: MIT Press.
- Greene, Joshua D. 2008. « The Secret Joke of Kant’s Soul ». Dans *Moral psychology, Vol. 3*. Sous la direction de W. Sinnott-Armstrong, 35–80. Cambridge, MA: MIT Press.

- Greene, Joshua D. 2013. *Moral Tribes : Emotion, Reason, and the Gap between Us and Them*. New York, NY: Penguin Press.
- Greene, Joshua D. 2014. « Beyond Point-And-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics ». *Ethics* 124 (4): 695–726.
- Greene, Joshua D. 2016. « Solving the Trolley Problem ». Dans *A Companion to Experimental Philosophy*. Sous la direction de Justin Sytsma et Wesley Buckwalter, 173–89. Hoboken, NJ: Wiley.
- Greene, Joshua D. 2017. « The Rat-A-Gorical Imperative: Moral Intuition and the Limits of Affective Learning ». *Cognition* 167: 66–77.
- Greene, Joshua D. 2018. « Can We Understand Moral Thinking without Understanding Thinking? » Dans *Atlas of Moral Psychology*. Sous la direction de Kurt Gray et Jesse Graham, 3–8. New York, NY: Guilford Press.
- Greene, Joshua D., Leigh E. Nystrom, Andrew D. Engell, John M. Darley, et Jonathan D. Cohen. 2004. « The Neural Bases of Cognitive Conflict and Control in Moral Judgment ». *Neuron* 44 (2): 389–400.
- Greene, Joshua D., R. Brian Sommerville, Leigh E. Nystrom, John M. Darley, et Jonathan D. Cohen. 2001. « An Fmri Investigation of Emotional Engagement in Moral Judgment ». *Science* 293 (5537): 2105–8.
- Haidt, Jonathan. 2001. « The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment ». *Psychological Review* 108 (4): 814–34.
- Haidt, Jonathan. 2007. « The New Synthesis in Moral Psychology ». *Science* 316 (5827): 998–1002.
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York, NY: Pantheon Books.

- Haidt, Jonathan. 2013. « Moral Psychology for the Twenty-First Century ». *Journal of Moral Education* 42 (3): 281–97.
- Haidt, Jonathan, et Fredrik Bjorklund. 2008. « Social Intuitionists Answer Six Questions about Moral Psychology ». Dans *Moral Psychology, Vol. 2*. Sous la direction de Walter Sinnott-Armstrong, 181–217. Cambridge, MA: MIT Press.
- Haidt, Jonathan, et Matthew A. Hersh. 2001. « Sexual Morality: The Cultures and Emotions of Conservatives and Liberals ». *Journal of Applied Social Psychology* 31: 191–221.
- Haidt, Jonathan, et Selin Kesebir. 2010. « Morality ». Dans *Handbook of social psychology 5th ed., Vol. 2*. Sous la direction de Susan T. Fiske, Daniel T. Gilbert et Gardner Lindzey, 797–832. New York, NY: Wiley.
- Hannon, Michael. 2018. « Intuitions, Reflective Judgments, and Experimental Philosophy ». *Synthese: An International Journal for Epistemology, Methodology and Philosophy of Science* 195 (9): 4147–68.
- Harman, Gilbert. 1977. *The Nature of Morality: An Introduction to Ethics*. New York, NY: Oxford University Press.
- Hasan, Ali et Richard Fumerton. 2018. « Foundationalist Theories of Epistemic Justification ». Sous la direction d'Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2018/entries/justep-foundational/>.
- Hauser, Marc, Fiery Cushman, Liane Young, R. Kang-Xing Jin, et John Mikhail. 2007. « A Dissociation Between Moral Judgments and Justifications ». *Mind and Language* 22 (1): 1–21.
- Hooker, Brad. 2002. « Intuitions and Moral Theorizing ». Dans *Ethical Intuitionism: Re-Evaluations*. Sous la direction de Philip Stratton-Lake, 161–83. Oxford: Clarendon Press.
- Huebner, Bryce. 2015. « Do Emotions Play a Constitutive Role in Moral Cognition? » *Topoi: An International Review of Philosophy* 34 (2): 427–40.

- Huemer, Michael. 2001. *Skepticism and the Veil of Perception*. Lanham, MD: Rowman & Littlefield Publishers.
- Huemer, Michael. 2006. « Phenomenal Conservatism and the Internalist Intuition ». *American Philosophical Quarterly* 43 (2): 147–58.
- Huemer, Michael. 2007. « Compassionate Phenomenal Conservatism ». *Philosophy and Phenomenological Research* 74 (1): 30–55.
- Huemer, Michael. 2008. *Ethical Intuitionism*. Basingstoke: Palgrave Macmillan.
- Huemer, Michael. 2008. « Revisionary Intuitionism ». *Social Philosophy and Policy* 25 (1): 368–92.
- Huemer, Michael. 2009. « Apology of a Modest Intuitionist ». *Philosophy and Phenomenological Research* 78 (1): 222–36.
- Huemer, Michael. 2013. « Phenomenal Conservatism Über Alles ». Dans *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*. Sous la direction de Chris Tucker, 328–50. New York, NY: Oxford University Press.
- Huemer, Michael. 2016. « A Liberal Realist Answer to Debunking Sceptics: The Empirical Case for Realism ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 173 (7): 1983–2010.
- Jacobson, Daniel. 2008. « Does Social Intuitionism Flatter Morality or Challenge It? » Dans *Moral Psychology, Vol. 2*. Sous la direction de Walter Sinnott-Armstrong, 219–32. Cambridge, MA: MIT Press.
- Joyce, Richard. 2001. *The Myth of Morality*. Cambridge, MA: Cambridge University Press.
- Joyce, Richard. 2007. *The Evolution of Morality*. Cambridge, MA: MIT Press.
- Joyce, Richard. 2016. *Essays in Moral Skepticism*. Oxford: Oxford University Press.

- Kahane, Guy. 2010. « Evolutionary Debunking Arguments ». *Noûs* 45 (1): 103–25.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.
- Kahneman, Daniel, Paul Slovic, et Amos Tversky. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, MA: Cambridge University Press.
- Kelman, Mark. 2013. « Moral Realism and the Heuristics Debate ». *Journal of Legal Analysis* 5 (2): 339–97.
- Kirchin, Simon. 2005. « What Is Intuitionism and Why Be an Intuitionist? ». *Social Theory and Practice* 31 (4): 581–606.
- Kitcher, Philip. 2005. « Biology and Ethics ». Dans *The Oxford Handbook of Ethical Theory*. Sous la direction de David Copp, 163–85. New York, NY: Oxford University Press.
- Klenk, Michael. 2017. « Can Moral Realists Deflect Defeat Due to Evolutionary Explanations of Morality? ». *Pacific Philosophical Quarterly* 98: 227–48.
- Klenk, Michael. 2017. « Old Wine in New Bottles : Evolutionary Debunking Arguments and the Benacerraf-Field Challenge ». *Ethical Theory and Moral Practice : An International Forum* 20 (4): 781–95.
- Klenk, Michael. 2018. « Third Factor Explanations and Disagreement in Metaethics ». *Synthese : An International Journal for Epistemology, Methodology and Philosophy of Science* 197 (1): 427–46.
- Koenigs, Michael, Liane Young, Ralph Adolphs, Daniel Tranel, Fiery Cushman, Marc Hauser, et Antonio Damasio. 2007. « Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgements ». *Nature* 446 (7138): 908–11.
- Königs, Peter. 2018. « On the Normative Insignificance of Neuroscience and Dual-Process Theory ». *Neuroethics* 11 (2): 195–209.

- Königs, Peter. 2020. « Experimental Ethics, Intuitions, and Morally Irrelevant Factors ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 177 (9): 2605–23.
- Koon, Justis. 2021. « The Epistemology of Evolutionary Debunking ». *Synthese : An International Journal for Epistemology, Methodology and Philosophy of Science* 199 (5-6): 12155–76.
- Korman, Daniel Z. 2019. « Debunking Arguments ». *Philosophy Compass* 14 (12): 1–17.
- Korman, Daniel Z. 2019. « Debunking Arguments in Metaethics and Metaphysics ». Dans *Metaphysics and Cognitive Science*. Sous la direction d’Alvin I. Goldman et Brian P. McLaughlin, 337–63. New York, NY: Oxford University Press.
- Korman, Daniel Z., et Dustin Locke. 2020. « Against Minimalist Responses to Moral Debunking Arguments ». Dans *Oxford Studies in Metaethics, Volume 15*. Sous la direction de Russ Shafer-Landau, 309–32. Oxford: Oxford University Press.
- Kumar, Victor, et Richmond Campbell. 2012. « On the Normative Significance of Experimental Moral Psychology ». *Philosophical Psychology* 25 (3): 311–30.
- Leibowitz, Uri D. 2021. « A Dilemma for Evolutionary Debunking Arguments ». *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition* 178 (1): 45–69.
- Levy, Neil. 2006. « Cognitive Scientific Challenges to Morality ». *Philosophical Psychology* 19 (5): 567–87.
- Liao, S. Matthew. 2016. « Are Intuitions Heuristics? » Dans *Moral Brains: The Neuroscience of Morality*. Sous la direction de S. Matthew Liao, 312–330. New York, NY: Oxford University Press.
- Lillehammer, Hallvard. 2011. « The Epistemology of Ethical Intuitions ». *Philosophy* 86 (2): 175–200.

- Lutz, Matt. 2018. « What Makes Evolution a Defeater? » *Erkenntnis : An International Journal of Scientific Philosophy* 83 (6): 1105–26.
- Markie, Peter. 2005. « The Mystery of Direct Perceptual Justification ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 126 (3): 347–73.
- Mason, Kelby. 2010. « Debunking Arguments and the Genealogy of Religion and Morality ». *Philosophy Compass* 5 (9): 770–78.
- Mason, Kelby. 2011. « Moral Psychology and Moral Intuition: A Pox on All Your Houses ». *Australasian Journal of Philosophy* 89 (3): 441–58.
- May, Joshua. 2018. *Regard for Reason in the Moral Mind*. Oxford, United Kingdom: Oxford University Press.
- May, Joshua, et Victor Kumar. 2020. « Moral Reasoning and Emotion ». Dans *The Routledge Handbook of Moral Epistemology*. Sous la direction d’Aaron Zimmerman, Karen Jones et Mark Timmons, 139–56. New York, NY: Routledge.
- Moll, Jorge, Ricardo de Oliveira-Souza, et Roland Zahn. 2008. « The Neural Basis of Moral Cognition : Sentiments, Concepts, and Values ». *Annals of the New York Academy of Sciences* 1124 (1): 161–80.
- Moon, Andrew. 2017. « Debunking Morality: Lessons from the Eaan Literature ». *Pacific Philosophical Quarterly* 98: 208–26.
- Moretti, Luca. 2020. *Seemings and Epistemic Justification : How Appearances Justify Beliefs*. Cham: Springer.
- Narvaez, Darcia. 2008. « The Social Intuitionist Model: Some Counter-Intuitions ». Dans *Moral Psychology, Vol. 2*. Sous la direction de Walter Sinnott-Armstrong, 233–40. Cambridge, MA: MIT Press.

- Nichols, Shaun. 2020. « Moral Learning ». Dans *The Routledge Handbook of Moral Epistemology*. Sous la direction d'Aaron Zimmerman, Karen Jones et Mark Timmons, 124–38. New York, NY: Routledge.
- Patil, Indrajeet, Micaela Maria Zucchelli, Wouter Kool, Stephanie Campbell, Federico Fornasier, Marta Calò, Giorgia Silani, Mina Cikara, et Fiery Cushman. 2021. « Reasoning Supports Utilitarian Resolutions to Moral Dilemmas Across Diverse Measures ». *Journal of Personality and Social Psychology* 120 (2): 443–60.
- Pizarro, David A., et Paul Bloom. 2003. « The Intelligence of the Moral Intuitions: Comment on Haidt (2001) ». *Psychological Review* 110 (1): 193–96.
- Pözlzer, Thomas. 2018. *Moral Reality and the Empirical Sciences*. New York, NY: Routledge.
- Pust, Joel. 2019. « Intuition ». Sous la direction d'Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2019/entries/intuition/>.
- Rachels, Stuart. 1998. « Counterexamples to the Transitivity Of better Than ». *Australasian Journal of Philosophy* 76 (1): 71–83.
- Railton, Peter. 2014. « The Affective Dog and Its Rational Tale: Intuition and Attunement ». *Ethics* 124 (4): 813–59.
- Railton, Peter. 2017. « Moral Learning: Conceptual Foundations and Normative Relevance ». *Cognition* 167: 172–90.
- Railton, Peter. 2020. « Rationalization of Emotion Is Also Rational ». *The Behavioral and Brain Sciences* 43: 43.
- Ross, William David. 1930. *The Right and the Good*. Oxford: Oxford University Press.
- Rossi, Mauro, et Christine Tappolet. 2017. « The Evolutionary Debunker Meets Sentimental Realism ». Dans *Facts and Values: The Ethics and Metaphysics of Normativity*. Sous la

direction de Giancarlo Marchetti and Sarin Marchetti, 176–95. New York, NY: Routledge, Taylor & Francis Group.

Royzman, Edward B., Kwanwoo Kim, et Robert F. Leeman. 2015. « The Curious Tale of Julie and Mark: Unraveling the Moral Dumfounding Effect ». *Judgment and Decision Making* 10 (4): 296–313.

Ruse, Michael. 1986. *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*. Oxford: Blackwell.

Ruse, Michael, et Robert J. Richards, dir. 2017. *The Cambridge Handbook of Evolutionary Ethics*. Cambridge, United Kingdom: Cambridge University Press.

Sauer, Hanno. 2012. « Morally Irrelevant Factors: What's Left of the Dual Process-Model of Moral Cognition? ». *Philosophical Psychology* 25 (6): 783–811.

Sauer, Hanno. 2017. *Moral Judgments as Educated Intuitions*. Cambridge, MA: MIT Press.

Sauer, Hanno. 2018. *Debunking Arguments in Ethics*. New York, NY: Cambridge University Press.

Sauer, Hanno. 2018. *Moral Thinking, Fast and Slow*. London: Routledge.

Schafer, Karl. 2010. « Evolution and Normative Scepticism ». *Australasian Journal of Philosophy* 88 (3): 471–88.

Schechter, Joshua. 2018. « Explanatory Challenges in Metaethics ». Dans *Routledge Handbook of Metaethics*. Sous la direction de Tristram McPherson et David Plunkett, 443–59. New York, NY: Routledge.

Schnall, Simone, Jonathan Haidt, Gerald L. Clore, et Alexander H. Jordan. 2008. « Disgust As Embodied Moral Judgment ». *Personality & Social Psychology Bulletin* 34 (8): 1096–1109.

Shafer-Landau, Russ. 2003. *Moral Realism: A Defence*. Oxford: Oxford University Press.

- Shafer-Landau, Russ. 2017. « Evolutionary Debunking, Moral Realism and Moral Knowledge ». *Journal of Ethics and Social Philosophy* 7 (1): 1–38.
- Singer, Peter. 1972. « Famine, Affluence, and Morality ». *Philosophy & Public Affairs* 1 (3): 229–43.
- Singer, Peter. 2005. « Ethics and Intuitions ». *The Journal of Ethics* 9 (3-4): 331–52.
- Sinnott-Armstrong, Walter. 2006. « Moral Intuitionism Meets Empirical Psychology ». Dans *Metaethics after Moore*. Sous la direction de Terry Horgan et Mark Timmons, 339–66. Oxford: Oxford University Press.
- Sinnott-Armstrong, Walter. 2006. *Moral Skepticisms*. Oxford: Oxford University Press.
- Sinnott-Armstrong, Walter. 2011. « An Empirical Challenge to Moral Intuitionism ». Dans *The New Intuitionism*. Sous la direction de Jill Graper Hernandez, 11–28. Londres: Continuum.
- Sinnott-Armstrong, Walter, Liane Young, et Fiery Cushman. 2010. « Moral Intuitions ». Dans *The Moral Psychology Handbook*. Sous la direction de John M. Doris et The Moral Psychology Research Group, 246–72. Oxford: Oxford University Press.
- Skarsaune, Knut Olav. 2011. « Darwin and Moral Realism: Survival of the Fittest ». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 152 (2): 229–43.
- Slovic, Paul, Melissa L. Finucane, Ellen Peters, et Donald G. MacGregor. 2007. « The Affect Heuristic ». *European Journal of Operational Research* 177 (3): 1333–52.
- Smart, J. J. C, et Bernard Williams. 1973. *Utilitarianism : For and Against*. Cambridge: Cambridge University Press.
- Sosa, Ernest. 2006. « Experimental Philosophy and Philosophical Intuition ». *Philosophical Studies* 132 (1): 99–107.

- Stratton-Lake, Philip. 2020. « Intuitionism in Ethics ». Sous la direction d'Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/sum2020/entries/intuitionism-ethics/>.
- Street, Sharon. 2006. « A Darwinian Dilemma for Realist Theories of Value ». *Philosophical Studies* 127 (1): 109–66.
- Street, Sharon. 2008. « Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About ». *Philosophical Issues* 18 (1): 207–28.
- Street, Sharon. 2015. « Does Anything Really Matter or Did We Just Evolve to Think So ». Dans *The Norton Introduction to Philosophy*. Sous la direction de Gideon A. Rosen, Alex Byrne, Joshua Cohen et Seana Valentine Shiffrin, 685–695. New York, NY: W.W. Norton & Company.
- Street, Sharon. 2016. « Objectivity and Truth: You'd Better Rethink It ». Dans *Oxford Studies in Metaethics, Volume 11*. Sous la direction de Russ Shafer-Landau, 293–334. Oxford: Oxford University Press.
- Sunstein, Cass R. 2004. « Moral Heuristics and Moral Framing ». *Minnesota Law Review* 88 (6): 1556–97.
- Sunstein, Cass R. 2005. « Moral Heuristics ». *The Behavioral and Brain Sciences* 28 (4): 531–42.
- Sunstein, Cass R. 2007. « Moral Heuristics and Risk ». Dans *Risk: Philosophical Perspectives*. Sous la direction de Tim Lewens, 156–70. Londres: Routledge.
- Sunstein, Cass R. 2008. « Fast, Frugal, and (Sometimes) Wrong ». Dans *Moral Psychology, Vol. 2*. Sous la direction de Walter Sinnott-Armstrong et Christian B. Miller, 27–30. Cambridge, MA: MIT Press.
- Talbott, William J. 2015. « How Could a 'Blind' Evolutionary Process Have Made Human Moral Beliefs Sensitive to Strongly Universal, Objective Moral Standards? » *Biology & Philosophy* 30 (5): 691–708.

- Tappolet, Christine. 2000. *Émotions et valeurs*. Paris: Presses Universitaires de France.
- Tersman, Folke. 2017. « Debunking and Disagreement: Debunking and Disagreement ». *Noûs* 51 (4): 754–74.
- Tersman, Folke, et Olle Risberg. 2022. « Hope for the Evolutionary Debunker: How Evolutionary Debunking Arguments and Arguments from Moral Disagreement Can Join Forces ». *Ethical Theory and Moral Practice*: 1–17.
- Thomson, Judith Jarvis. 1985. « The Trolley Problem ». *The Yale Law Journal* 94 (6): 1395–1415.
- Tversky, Amos, et Daniel Kahneman. 1982. « Availability: A Heuristic for Judging Frequency and Probability ». Dans *Judgment under Uncertainty: Heuristics and Biases*. Sous la direction de Daniel Kahneman, Paul Slovic et Amos Tversky, 163–78. Cambridge, MA: Cambridge University Press.
- Unger, Peter. *Living High and Letting Die: Our Illusion of Innocence*. New York: Oxford University Press, 1996.
- Vavova, Katia. 2014. « Debunking Evolutionary Debunking ». Dans *Oxford Studies in Metaethics, Volume 9*. Sous la direction de Russ Shafer-Landau, 76–101. Oxford: Oxford University Press.
- Vavova, Katia. 2015. « Evolutionary Debunking of Moral Realism ». *Philosophy Compass* 10 (2): 104–16.
- Vavova, Katia. 2021. « The Limits of Rational Belief Revision: A Dilemma for the Darwinian Debunker ». *Noûs* 55 (3): 717–34.
- Weisberg, Jonathan. 2012. « The Bootstrapping Problem ». *Philosophy Compass* 7 (9): 597–610.
- Wheatley, Thalia, et Jonathan Haidt. 2005. « Research Report: Hypnotic Disgust Makes Moral Judgments More Severe ». *Psychological Science* 16 (10): 780–84.

- Wheeler, Gregory. 2020. « Bounded Rationality ». Sous la direction d' Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/fall2020/entries/bounded-rationality/>.
- Wielenberg, Erik J. 2010. « On the Evolutionary Debunking of Morality ». *Ethics* 120 (3): 441–64.
- Wielenberg, Erik J. 2014. *Robust Ethics : The Metaphysics and Epistemology of Godless Normative Realism*. Oxford: Oxford University Press.
- Woodward, James. 2016. « Emotion versus Cognition in Moral Decision-Making: A Dubious Dichotomy ». Dans *Moral Brains: The Neuroscience of Morality*. Sous la direction de S. Matthew Liao, 87–116. New York, NY: Oxford University Press.
- Woodward, James, et John Allman. 2007. « Moral Intuition: Its Neural Substrates and Normative Significance ». *Journal of Physiology - Paris* 101 (4-6): 179–202.
- Zahn, Roland, Ricardo de Oliveira-Souza et Jorge Moll. 2015. « Neural Foundation of Morality ». Dans *International Encyclopedia of the Social & Behavioral Sciences*. Sous la direction de James D. Wright, 606–18. Amsterdam: Elsevier.
- Zimmerman, Aaron. 2010. *Moral Epistemology*. Milton Park, Abingdon, Oxon: Routledge.
- Zimmerman, Aaron. 2013. « Skepticism, Moral ». Dans *The International Encyclopedia of Ethics*. Sous la direction de Hugh LaFollette, 1–10. Hoboken, NJ: John Wiley & Sons.