

Université de Montréal

Faits, valeurs et non-cognitivism

Une analyse critique

Par

Sandra Froidevaux

Département de philosophie

Faculté des arts et sciences

Mémoire présenté à la Faculté des études supérieures en vue de l'obtention
du grade de *maître es art* en philosophie

Décembre 2006

© Sandra Froidevaux, 2006



B
29
US4
2007
V.011

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

Faits, valeurs et non-cognitivism

Une analyse critique

Présenté par:

Sandra Froidevaux

A été évalué par un jury composé des personnes suivantes:

[REDACTED]

Président-rapporteur

[REDACTED]

Directeur de recherche

[REDACTED]

Membre du jury

SOMMAIRE

Faits, valeurs et non-cognitivismes

Une analyse critique

La dichotomie fait-valeur, telle qu'elle est défendue par le non-cognitivismes, établit que les énoncés moraux diffèrent des énoncés factuels en ce sens que les seconds, contrairement aux premiers, sont des croyances. Le non-cognitivismes soutient que les jugements moraux ne possèdent pas de valeur de vérité parce qu'ils ne représentent pas des états de choses. Cette dichotomie implique qu'il ne peut être question d'objectivité en éthique et qu'un débat moral ne saurait déboucher sur la découverte d'un jugement moral vrai. Dans ce travail, je me propose d'examiner les différents arguments apportés par le non-cognitivismes et de défendre le réalisme moral. Je commencerai par un examen de l'émotivisme qui fonde la dichotomie fait-valeur sur des critères du positivisme logique. Je serai amenée à discuter du vérificationnisme et de la difficulté qu'a le non-cognitivismes à rendre compte du fonctionnement des énoncés moraux dans des contextes non assertés. Je discuterai, ensuite, des théories de la motivation morale puisque le non-cognitivismes défend sa thèse que les jugements moraux ne sont pas des croyances en invoquant la thèse internaliste de la motivation ainsi que la théorie humienne de la motivation. Selon cette dernière, des croyances seules ne sont pas capables de motiver et selon l'internalisme les jugements moraux sont nécessairement motivants. Le non-cognitivismes en conclut alors que les jugements moraux ne peuvent pas être des croyances. Je défendrai ici une position externaliste. Finalement, je me pencherai sur les théories du réalisme moral. En effet, le non-cognitivismes refuse d'accorder une valeur de vérité aux énoncés moraux en notant que puisqu'il n'existe pas de faits moraux, ceux-ci ne réfèrent à rien. S'il existe des faits moraux, il n'y a plus de raison de penser que les énoncés moraux ne sont pas des croyances. Je soutiendrai que les faits moraux surviennent sur des faits naturels et qu'ils ne posent ainsi pas de problème à une ontologie naturaliste.

MOT CLÉS

Philosophie, méta-éthique, dichotomie fait-valeur, externalisme, réalisme moral

ABSTRACT

Facts, values and non-cognitivism

A critical analysis

The fact-value dichotomy, as it has been defended by non-cognitivism, establishes that moral claims differ from factual claims in the sense that the second, but not the first, are beliefs. Non-cognitivism claims that moral judgments are not truth-apt because they do not represent facts of matter. This dichotomy implies that there is no question of objectivity in ethics and that a moral debate cannot bring about the discovery of a true moral judgment. In this paper, I propose myself to look at the different arguments given by non-cognitivism and to defend moral realism. I will start with an examination of emotivism which founds the fact-value dichotomy on criterias defended by logical positivism. I will discuss verificationism and the difficulty shown by non-cognitivism to describe the working of moral claims in unasserted contexts. I will next discuss moral motivation theories. Non-cognitivism defends its thesis that moral judgments are not beliefs using the internalist thesis of motivation as well as the Humean theory of motivation. The Humean theory establishes that beliefs alone cannot motivate and internalism claims that necessarily, moral judgments motivate. Non-cognitivism concludes that moral judgments cannot be beliefs. I will defend externalism. Finally, I will consider theories of moral realism. Non-cognitivism denies that moral claims have a truth value because moral facts do not exist and therefore moral claims do not refer to anything. If moral facts exist, there is no reason to deny that they are beliefs anymore. I will defend that moral facts supervene on natural facts and that they are not a problem for a naturalistic ontology.

KEY WORDS

Philosophy, metaethics, fact-value dichotomy, externalism, moral realism

Table des matières

| | |
|---|----|
| Sommaire | i |
| Abstract | ii |
| Introduction | 1 |
| 1. Le non-cognitivism | 5 |
| 1.1 L'émotivisme d'Ayer | 7 |
| 1.2 La notion de fait du positivisme logique | 10 |
| 1.2.1 <i>La notion de fait est-elle neutre au niveau axiologique?</i> | 16 |
| 1.3 La signification émotive | 20 |
| 1.4 Les actes de langages | 21 |
| 1.5 Les attitudes morales | 23 |
| 1.6 Le problème Frege-Geach | 25 |
| 1.6.1 <i>La solution de Blackburn</i> | 26 |
| 1.6.2 <i>La solution de Gibbard</i> | 32 |
| 1.7 Conclusion | 35 |
| 2. Non-cognitivism et motivation | 38 |
| 2.1 L'externalisme | 39 |
| 2.2 La théorie humienne de la motivation | 54 |
| 2.3 Les théories rationalistes de la motivation | 61 |
| 2.4 Conclusion | 64 |
| 3. Le réalisme moral | 66 |
| 3.1 L'argument de la question ouverte | 68 |
| 3.2 La survenance | 72 |
| 3.2.1 <i>L'objection de Blackburn</i> | 75 |
| 3.2.2 <i>L'explication de la relation de survenance</i> | 83 |
| 3.3 Les explications morales | 93 |

| | |
|----------------|-----|
| 3.4 Conclusion | 106 |
|----------------|-----|

| | |
|-------------------|-----|
| Conclusion | 108 |
|-------------------|-----|

| | |
|----------------------|-----|
| Bibliographie | 115 |
|----------------------|-----|

Introduction

Dans ce travail, je me propose d'examiner la dichotomie fait-valeur. Cette distinction est importante, car c'est au nom de celle-ci que les non-cognitivistes ont relégué les énoncés axiologiques au rang de pseudo-propositions, dans le meilleur des cas, ou d'énoncés dépourvus de signification littérale. Le non-cognitivism implique que nos discussions concernant des questions éthiques ne sont pas de même nature que les débats scientifiques et, qu'en fin de compte, chacun ne fait que défendre ses propres opinions ou sentiments et qu'il n'y a pas réellement de matière à débattre. En effet, pour le non-cognitivism, les énoncés évaluatifs sont des projections de nos émotions ou attitudes. Les discussions éthiques seraient alors similaires à une défense de nos préférences puisque le non-cognitivism implique qu'il ne saurait être question d'objectivité dans le champ moral. Si aucun énoncé moral n'est vrai, s'engager dans des débats sur des questions morales ne fait pas vraiment sens. En effet, le débat ne peut pas amener à la découverte d'un jugement moral qui soit vrai. Ce que je tenterai de montrer, c'est qu'il y a de bonnes raisons pour ne pas soutenir une dichotomie fait-valeur au sens fort. Je ne cherche pas à nier qu'il s'agisse d'une distinction utile, mais je ne pense pas qu'il s'agisse d'une dichotomie exhaustive. Je défendrai ici le réalisme moral. Je pense que les énoncés moraux sont des croyances et qu'ils réfèrent à des états de choses.

C'est pourquoi, j'examinerai, dans ce travail, les arguments en faveur du non-cognitivism pour voir si cette position méta-éthique est convaincante. Je commencerai avec la théorie émotiviste d'Alfred Ayer, qui est un des pères du non-cognitivism puis considérerai des arguments plus contemporains, en particulier ceux de Simon Blackburn qui est un défenseur du quasi-réalisme et d'Allan Gibbard qui soutient l'expressivisme des normes.

On peut distinguer trois thèses au sein de la théorie non-cognitiviste. Une thèse sémantique affirmant que les énoncés axiologiques ne sont ni vrais ni faux. Une thèse ontologique postulant que les propriétés morales n'existent pas et que nos jugements moraux sont des expressions de nos émotions ou attitudes. Et finalement, une thèse psychologique selon laquelle les jugements moraux sont nécessairement motivants. La thèse sémantique a pour corollaire une thèse épistémologique. Si les jugements moraux ne sont ni vrais, ni faux, il n'y a pas de possibilité d'erreur ou de connaissance morale.

L'avantage du non-cognitivism, c'est qu'il est compatible avec le naturalisme, il peut donc limiter son ontologie à celle employée par les sciences. Le naturalisme affirme que seuls les propriétés et objets naturels existent. Cependant, le non-cognitivism ne cherche pas à réduire

les propriétés morales à des propriétés naturelles. C'est parce qu'il nie qu'il existe des propriétés évaluatives qu'il est naturaliste. Il ne postule donc en rien l'existence d'entités mystérieuses. Il évite ainsi aussi des problèmes épistémologiques. En effet, puisqu'il n'y a pas de propriétés morales et donc pas de vérités morales à connaître, il n'a pas besoin d'expliquer comment nous arriverions à connaître ces dernières.

J'envisagerai trois manières de contrer la dichotomie fait-valeur. Premièrement, je regarderai les problèmes rencontrés par l'émotivisme. L'émotivisme adopte les thèses du positivisme logique concernant la signification et la vérification. Il opère une distinction entre énoncés analytiques, c'est-à-dire vrais en vertu de leur signification, énoncés factuels empiriquement vérifiables et énoncés dépourvus de signification. Selon les critères positivistes, les jugements moraux ne sont ni des énoncés analytiques, ni des énoncés empiriquement vérifiables. Ils ne sont donc pas de véritables propositions et sont donc dépourvus de valeur de vérité. On peut alors se demander si cette distinction est pertinente et j'examinerai les arguments d'Hilary Putnam pour montrer que les faits ne sont pas neutres au niveau axiologique. Putnam met en évidence l'importance des valeurs épistémiques, comme la cohérence ou la simplicité, par exemple, dans la pratique scientifique. Ce sont, en effet, ces valeurs qui nous permettraient d'effectuer un choix entre les différentes théories rivales. Il n'y aurait pas de connaissance factuelle possible sans présupposer au préalable une connaissance axiologique. Quine a aussi argumenté en faveur de l'abandon de la distinction entre énoncés analytiques et synthétiques. Si cette distinction tombe, la distinction tripartite employée par les émotivistes perd également de sa pertinence.

Par ailleurs, le non-cognitivism fait face à une série de problèmes, dont le problème Frege-Geach qui souligne la difficulté qu'a le non-cognitivism à rendre compte du fonctionnement des énoncés moraux dans des contextes non assertés. Des auteurs expressivistes contemporains comme Simon Blackburn et Allan Gibbard ont cherché à donner une solution expressiviste à ce problème et je verrai si ces solutions sont convaincantes. Les autres problèmes que j'envisagerai concernent la difficulté du non-cognitivism à offrir un critère de la signification émotive, à tenir compte de la variété de nos actes de langages et de la richesse de notre vocabulaire moral.

Deuxièmement, je me pencherai sur le problème de la motivation exercée par les énoncés moraux. En effet, les non-cognitivistes observent qu'une des différences entre les jugements

moraux et les énoncés factuels, c'est que les premiers incitent à agir, alors que pas les seconds. Si je juge que la générosité est une bonne chose, je dispose d'une raison motivante à me montrer généreuse. Si j'affirme qu'il faut respecter l'environnement, mais que je ne trie pas mes déchets, on en conclura que je ne suis pas vraiment convaincue des thèses environnementalistes. La thèse non-cognitivistique soutient que les jugements moraux ne peuvent pas être des croyances, mais seulement des expressions d'émotions ou d'attitudes. Elle repose sur deux autres thèses, à savoir la thèse internaliste de la motivation et la théorie humienne de la motivation. La première soutient que les jugements moraux sont nécessairement motivants. La théorie humienne de la motivation propose une analyse du comportement humain selon laquelle des croyances seules ne suffisent pas à motiver. Il faut toujours, pour cela, la présence d'un désir. Or, relève le non-cognitivistique, si des croyances seules ne suffisent pas à motiver et que les jugements moraux sont nécessairement motivants, ils ne sauraient être des croyances. Puisqu'ils ne sont pas des croyances, mais sont alors l'expression d'états affectifs, ils ne possèdent pas de conditions de vérité. Je me pencherai donc sur ces deux thèses pour voir si elles sont convaincantes. Notons qu'il suffit qu'une des deux soit fautive pour que l'argument non-cognitivistique échoue.

Troisièmement, on peut chercher à montrer qu'il existe des faits moraux et qu'il n'y aurait donc pas de raison de postuler un fossé métaphysique insurmontable entre faits et valeurs. Je me tournerai donc vers les théories du réalisme moral. Un philosophe comme Ayer commence par affirmer qu'il n'existe pas de faits moraux. Comme il adhère au principe selon lequel toutes les propositions sont vérifiables, il en conclut que les énoncés moraux ne sont pas de véritables propositions et qu'ils ne possèdent pas de valeur de vérité. En effet, comment pourrions-nous vérifier quelque chose qui n'existe pas? S'il existe des faits moraux, par conséquent, il n'y a plus de raison de refuser des conditions de vérité aux énoncés éthiques.

Une possibilité pour expliquer l'existence de faits moraux consiste à dire qu'ils sont réductibles à des faits naturels. Cependant, G.E. Moore objecte au réalisme naturaliste réductionniste l'argument de la question ouverte. Il cherche à montrer qu'il n'est pas possible d'établir des relations d'identité analytique entre propriétés morales et propriétés naturelles. Je me pencherai donc sur cet argument et verrai comment un réaliste naturaliste peut y répondre. Moore défendait un réalisme non naturaliste. Il pensait que les propriétés morales sont des propriétés simples, non analysables, et non naturelles que l'on détecte grâce à l'intuition.

Pour expliquer comment les propriétés évaluatives peuvent être réelles sans avoir à

postuler l'existence d'entités mystérieuses ou non naturelles, on peut montrer que les propriétés évaluatives surviennent sur des propriétés physiques. De cette manière, le fait de postuler des propriétés évaluatives n'est pas incompatible avec le naturalisme. Or Blackburn, ainsi que Horgan et Timmons ont formulé des arguments pour montrer que la survenance pose problème aux réalistes. Je regarderai donc si les réalistes ont des moyens de répondre à ces objections. Pour terminer, je verrai s'il est possible que des faits moraux aient un rôle explicatif. En effet, pour qu'une entité soit admise dans notre ontologie, il est généralement requis qu'elle joue un rôle explicatif. La question est donc de savoir s'il est possible de donner des explications morales de nos observations qui ne puissent pas être remplacées par des explications en termes naturels.

1. Le non-cognitivism

Pour débiter mon examen de la dichotomie fait-valeur, je commencerai par regarder la position non-cognitivist et en particulier l'émotivisme tel qu'il a été présenté par Alfred Ayer. Le non-cognitivism considère que les énoncés évaluatifs ne sont pas aptes à recevoir une valeur de vérité parce qu'ils ne réfèrent pas à des propriétés réelles. Par conséquent, les jugements moraux n'exprimeraient pas des croyances, mais serviraient plutôt à exprimer des attitudes non-cognitives. C'est cela qui les distingue des énoncés factuels qui, eux, expriment des croyances et possèdent une valeur de vérité. Il existe, outre l'émotivisme, d'autres sortes de non-cognitivism. Le prescriptivism a été défendu en particulier par R.M. Hare et affirme que les jugements moraux servent à exprimer des prescriptions. Selon Hare, dire qu'une action est juste, revient à prescrire cette action ainsi que toute action similaire. Le quasi-réalisme et l'expressivism des normes sont des formes contemporaines d'expressivism et sont défendus respectivement par Simon Blackburn et par Allan Gibbard. Ces versions du non-cognitivism cherchent à accommoder l'aspect réaliste du discours moral tout en niant qu'il y ait des faits moraux. Le quasi-réalisme pense que nous sommes justifiés d'agir comme si les jugements moraux avaient une valeur de vérité, alors qu'en fin de compte ils ne sont ni vrais, ni faux. Il faut donc distinguer deux niveaux du langage, un niveau superficiel et un niveau profond. Les jugements moraux expriment, selon lui, notre disposition à former des sentiments d'approbation et de désapprobation. L'expressivism des normes suggère que les jugements moraux expriment l'acceptation d'un système de normes. Pour Gibbard, une action est blâmable s'il est rationnel que l'auteur s'en sente coupable ou que d'autres le jugent tel. Puisque la culpabilité et le ressentiment sont des états non-cognitifs, les jugements moraux ne possèdent pas de valeur de vérité.

L'émotivisme fournit des arguments en faveur d'une distinction tranchée entre faits et valeurs en faisant appel à des thèses des positivistes logiques, parmi lesquels on peut citer Rudolph Carnap et Moritz Schlick. La dichotomie fait-valeur a été défendue par les positivistes logiques sur la base d'une distinction entre énoncés pourvus et dépourvus de signification. Il y a, selon eux, deux manières pour un énoncé d'être signifiant: soit il est analytique, soit il est empiriquement vérifiable. Or, comme les jugements évaluatifs ne peuvent pas être confirmés ou infirmés par l'expérience et qu'ils ne sont pas non plus des énoncés analytiques, cela implique qu'ils ne possèdent pas de signification. Ceci a donc pour conséquence que toute connaissance axiologique est impossible et les positivistes purs et durs comme Carnap pensaient que des

domaines tels que l'éthique ou l'esthétique auraient dû être expulsés du champ de la philosophie. Le positivisme logique est, en ce sens, un héritier de la première philosophie de Wittgenstein. En effet, dans le *Tractatus logico-philosophicus*, il distingue le dicible de l'indicible. Seules les propositions ayant une valeur de vérité possèdent un sens et sont dicibles. Ceci implique qu'une proposition est sensée si nous comprenons quel état de choses lui correspondrait si elle était vraie. Les propositions ne peuvent donc énoncer que des faits. Comme Wittgenstein pense que, dans le monde, nous ne trouvons que des faits bruts dont aucun n'a plus de valeur que les autres, l'éthique se trouve, par conséquent, du côté de l'indicible, de même que l'esthétique et la religion. Wittgenstein opère ainsi une distinction entre le monde des faits et le domaine des valeurs.

Pourtant, un auteur comme Alfred Ayer ne réserve pas le même sort aux énoncés évaluatifs qu'aux énoncés métaphysiques qu'il considère comme du verbiage et du non-sens. Les jugements de valeur ont, selon lui, une fonction même s'ils ne sont pas littéralement signifiants. Ils "sont simplement des expressions de l'émotion qui ne peuvent être ni vraies ni fausses"¹. Les jugements de valeur sont clairement distingués des énoncés factuels qui, eux, sont empiriquement vérifiables. L'erreur, c'est de croire que la fonction du langage moral est de décrire des faits. La présence d'un prédicat moral dans une phrase n'amène aucune nouvelle information puisqu'il ne décrit rien mais ne sert qu'à exprimer une attitude. La théorie d'Ayer ne permet donc pas non plus la possibilité d'une connaissance morale.

Dans ce chapitre, je débiterai par un examen de l'émotivisme d'Ayer ainsi que les arguments apportés par Putnam et Quine remettant en cause la conception positiviste des faits et de l'analyticité. Je considérerai ensuite une série d'objections au non-cognitivism. Une des difficultés majeures rencontrées par le non-cognitivism a été exposée par Peter Geach et découle de l'emploi d'énoncés évaluatifs dans des contextes non assertés. En effet, le non-cognitivism a du mal à expliquer la validité des *modus ponens* moraux. J'examinerai alors les solutions expressivistes offertes par Blackburn et Gibbard à ce problème. Les autres objections portent sur l'échec du non-cognitivism à offrir un critère de la signification émotive, à rendre compte de la diversité des actes de langage et sur sa difficulté à tenir compte de la richesse de notre vocabulaire moral.

¹Ayer, 1956, p.144

1.1 L'émotivisme d'Ayer

Ayer est un représentant du positivisme logique et le premier à proposer une théorie émotiviste de l'éthique. D'autres auteurs, comme C.L. Stevenson auteur de *Facts and Value*, sont également des représentants de l'émotivisme, mais je n'examinerai ici que la théorie d'Ayer qui a le mérite d'être très clairement exposée et d'être le point de départ des théories expressivistes contemporaines.

Dans *Langage, Vérité et Logique*, Ayer se penche sur les énoncés éthiques, mais ce qu'il dit à leur sujet est également valable pour les énoncés évaluatifs d'autres types. Ayer adopte le critère de la signification littérale soutenu par le positivisme logique. Un énoncé peut posséder une signification seulement de deux manières. Soit il exprime une croyance *a posteriori* qui peut être vérifiée de manière empirique, soit il exprime un énoncé analytique, c'est-à-dire un énoncé vrai ou faux en vertu de la signification de ses termes. Le principe de vérification établit les conditions qu'un énoncé doit satisfaire pour avoir une signification en vertu d'être empiriquement vérifiable. Ce principe peut être résumé par le slogan de Schlick disant que la signification d'un énoncé consiste en ses méthodes de vérification.² Les énoncés pouvant être vérifiés empiriquement ont une signification factuelle. Les seuls énoncés dépourvus d'une signification factuelle, mais néanmoins signifiants, sont les énoncés analytiques. Pour savoir si une proposition est vérifiable, il faut savoir quelles observations rendraient la proposition vraie ou fausse, ou pour le formuler autrement, il faut connaître les conditions de satisfaction de la proposition. Puisque, dans le cas des énoncés moraux, on ne sait pas quelles observations rendraient l'énoncé vrai ou faux, ceux-ci ne possèdent pas de signification factuelle. Comme ils ne sont pas non plus des énoncés analytiques, ils ne possèdent pas de signification littérale.

En affirmant que les jugements de valeur ne sont pas vérifiables empiriquement, Ayer nie qu'il soit possible de réduire les termes éthiques à des termes non évaluatifs. Pour ce faire, il s'oppose à l'utilitarisme et au subjectivisme. Il affirme qu'il n'est "pas contradictoire de dire qu'il est quelquefois immoral de réaliser l'action qui produirait effectivement ou probablement le plus grand bonheur ou la plus grande supériorité de plaisir sur la peine ou du désir satisfait sur le désir

²Miller, 1998, p.106

insatisfait”³ et qu’ “il n’est pas contradictoire d’affirmer que des actions qui sont généralement approuvées ne sont pas bonnes”⁴. Ayer admet que l’utilitarisme et le subjectivisme pourraient être des projets pour remplacer les notions éthiques actuelles par de nouvelles, mais ces deux doctrines ne sauraient être des analyses de nos notions éthiques actuelles. Il serait donc possible d’inventer un langage où les symboles éthiques seraient définis en termes factuels. Cependant la réduction des jugements éthiques à des jugements non éthiques n’est pas en accord avec notre langage actuel. Puisque les jugements de valeur ne peuvent pas être assimilés à des propositions synthétiques, elles sont invérifiables. Selon Ayer, “les concepts fondamentaux de l’éthique sont inanalysables” et “il n’y a pas de critère par lequel on puisse vérifier la validité des jugements dans lesquels ils se trouvent introduits”⁵. L’idée que les concepts moraux sont inanalysables a été défendue par G.E. Moore dans *Principia Ethica*. Selon Moore, en cherchant à définir les concepts moraux en termes de propriétés naturelles, on commet ce qu’il appelle un sophisme naturaliste. Comme il ne pense pas qu’on puisse les définir en termes non naturels non plus, cela signifie que les concepts éthiques sont inanalysables. Ayer affirme que s’ils sont inanalysables, c’est qu’il s’agit de pseudo-concepts: “la présence d’un symbole éthique dans une proposition n’ajoute rien à son contenu factuel”.⁶ Par exemple, en disant “Vous avez mal agi en volant cet argent” je ne dis rien de plus que “Vous avez volé cet argent”. En disant que l’action est mauvaise, je ne formule pas un autre jugement, je ne fais que manifester ma désapprobation. Les termes éthiques fonctionnent comme un point d’exclamation et ils n’ajoutent rien au sens littéral de la phrase. Un énoncé comme “Voler de l’argent est mal” n’a aucun contenu factuel et n’exprime donc aucune proposition. Il n’a, par conséquent, pas de conditions de vérité.

[...] en disant qu’un certain type d’action est bien ou mal, je ne formule aucun jugement factuel, pas même un jugement sur mon propre état d’esprit. J’exprime simplement certains sentiments moraux. Et celui qui est sensé me contredire ne fait qu’exprimer ses sentiments moraux. Il n’y a donc absolument pas de sens à demander qui de nous a raison. (Ayer, 1956, p.151)

[...] les jugements moraux ne disent rien au sujet du monde. Ce sont de pures expressions de sentiment, et

³Ayer, 1956, p.147

⁴*Ibid*, p.146

⁵*Ibid*, p.150

⁶*Ibid*, p.150

comme telles, elles ne tombent pas dans la catégorie du vrai et du faux. Elles sont invérifiables parce qu'elles n'expriment pas de propositions authentiques. (Ayer, 1956, p. 152)

Puisque les jugements moraux sont des expressions de sentiments, on ne peut pas dire qu'ils sont véritablement des jugements. Le rôle des termes éthiques est de susciter des sentiments et de stimuler l'action.

Il y a une différence entre la position de Ayer et celle d'un subjectiviste classique. Ce dernier ne nie pas que les jugements éthiques soient des propositions. En effet, il s'agit, selon lui, de propositions concernant les sentiments du sujet. Un jugement moral est donc vrai si le locuteur possède les sentiments exprimés et faux s'il ne les a pas. Il faut faire ici la distinction entre l'expression d'un sentiment et l'assertion d'un sentiment. En effet, l'assertion implique l'expression alors que l'inverse n'est pas vrai. Selon Ayer, les jugements moraux sont des expressions de sentiments mais ils n'impliquent pas nécessairement l'assertion de ces mêmes sentiments. En effet, quand je juge que la torture est ignoble, je ne *dis* rien. *J'exprime* ma désapprobation mais je ne dis pas que je désapprouve la torture. On peut établir un parallèle avec la douleur. Lorsque je crie, je ne dis pas que je souffre. Mon cri est une expression de la douleur. Le non-cognitivism échappe ainsi à l'objection qui peut être faite au subjectivisme disant qu'il est absurde de penser que la validité des jugements moraux soit déterminée par la nature des sentiments de son auteur.

La théorie d'Ayer implique néanmoins qu'il est impossible de discuter des questions de valeurs. Pour lui, cela n'est pas véritablement un problème puisqu'il est d'avis que " [...] la discussion n'est pas réellement une discussion sur la valeur mais sur une question de faits"⁷. Lors d'un désaccord, nous chercherions à montrer que notre interlocuteur s'est trompé sur les circonstances du cas en question. Nous cherchons à le convaincre quant à la nature des faits empiriques pour qu'il adopte ensuite la même attitude morale. Cependant si notre interlocuteur a subi un "conditionnement" moral différent du nôtre, il persistera à avoir une évaluation morale différente même après avoir reconnu tous les faits. C'est à ce moment que prend fin l'argument. Un désaccord moral n'est pas la confrontation de croyances contradictoires mais un conflit entre des émotions.

⁷*Ibid*, p.155

Mais nous ne pouvons pas avancer d'argument propre à montrer que notre système (moral) est supérieur. Car notre jugement qu'il en est ainsi est lui-même un jugement de valeur, et par conséquent se trouve en-dehors de l'objet de la discussion. C'est parce que l'argument nous fait défaut quand nous traitons de pures questions de valeur comme distinctes de questions de fait, que nous recourrons finalement à la violence. (Ayer, 1956, p.156)

Ainsi, la discussion de questions morales n'est possible qu'à la condition qu'un système de valeurs soit présupposé.

Si notre interlocuteur s'accorde avec nous pour exprimer la désapprobation morale de toutes les actions d'un type donné *t*, alors nous pouvons obtenir de lui qu'il condamne une action particulière *A*, en montrant que *A* est du type *t*. Car la question de savoir si *A* est ou non du type *t* est une simple question de fait. [...] Ce que nous ne discutons pas et ne pouvons discuter, c'est au sujet de la validité de ces principes moraux. Nous les louons ou condamnons simplement à la lumière de nos propres sentiments. (Ayer, 1956, p.157)

Si on veut faire changer les sentiments éthiques de quelqu'un, on peut attirer son attention sur des faits qu'on pense lui avoir échappés. On peut l'influencer par le choix convenable d'un langage émotif. Si l'autre persiste dans son attitude contraire sans remettre en question de fait, la discussion ne peut plus avancer. Il ne saurait donc y avoir de désaccord moral concernant la validité de principes moraux.

1.2 La notion de fait du positivisme logique

L'argument d'Ayer en faveur de la dichotomie fait-valeur repose sur la thèse de la signification littérale. C'est parce qu'ils ne sont ni analytiques, ni empiriquement vérifiables que les jugements moraux n'ont pas de signification littérale. Il est cependant possible de penser à des énoncés qui ne sont ni analytiques, ni empiriquement vérifiables mais qui possèdent néanmoins une signification selon les critères positivistes. Par exemple, la phrase "Before the Big Bang all the objects in the Universe converged on a single point" est parfaitement signifiante.⁸ Cependant, c'est dans la nature même du Big Bang d'avoir détruit toute trace de ce qui se passait avant et ainsi cet énoncé n'est pas empiriquement vérifiable.

⁸Smith, 1994, p.20

Selon Hilary Putnam, la conviction des positivistes que l'éthique ne constitue pas un domaine du savoir provient du fait que la dichotomie fait-valeur et le dualisme analytique-synthétique se renforcent mutuellement.⁹ En effet, il y a selon eux une division tripartite des jugements: les jugements empiriques qui sont empiriquement vérifiables, les jugements analytiques qui sont vrai ou faux en vertu de règles logiques et les énoncés dépourvus de signification. C'est parce que les jugements évaluatifs ne sont pas factuels, selon la définition positiviste, et parce qu'ils ne sont évidemment pas analytiques, qu'ils ne peuvent pas constituer une forme de connaissance. L'exclusion de l'éthique du domaine de la connaissance repose donc sur la notion de fait et cette exclusion ne peut être valable qu'à la condition d'avoir à disposition une définition exacte de ce qu'est un fait.

In the writings of the positivists, in the cases of both the dualism of analytic and factual statements and the dualism of ethical and factual statements, it is the conception of the "factual" that does all the philosophical work. (Putnam, 2002, p.21)

On peut donc se poser la question de savoir si la notion de fait employée par les positivistes va de soi. Face à l'évolution scientifique, les positivistes ont été amenés à réviser leur conception initiale de ce qui est factuel. Ce faisant, selon Putnam, ils ont détruit la base de la dichotomie fait-valeur.

La notion originale de fait employée par les positivistes, similaire à celle à l'oeuvre dans l'empirisme classique, stipule que tout énoncé factuel doit pouvoir être transformé en énoncé portant sur une expérience sensible du sujet. Ceci a pour conséquence qu'ils doivent être vérifiables et donc, qu'ils puissent être confrontés à l'expérience. Un fait, selon cette conception initiale, est quelque chose qui doit pouvoir être certifié par l'observation. Puisque l'observation constitue un critère déterminant, il s'ensuit que les énoncés portant sur les bactéries ou les électrons sont, soit dépourvus de sens, soit doivent pouvoir être réduits à des termes d'observation. En effet, les bactéries ne sont pas observable au sens positiviste du terme. Au vu des développements scientifiques, l'idée qu'un fait doit être réductible à un énoncé portant sur une expérience sensible ne semble pas très convaincante. C'est pour cette raison que Carnap a été amené à réviser sa conception de ce qui est factuel:

⁹ Putnam. 2002. p.20

In his 1938 "Foundations of Logic and Mathematics" such troublesome terms as "electron" and "charge" do not enter physics through definitions (or even "reductions") but are simply "taken as primitive." As long as the system as a whole enables us to predict our experiences more successfully than we could without them, such "abstract terms" are to be accepted as "empirically meaningful". (Putnam, 2002, p.23)

Les termes théoriques comme "électron" sont admis dans le langage de la science, mais ils sont considérés comme des instruments pour dériver des phrases qui, elles, décrivent réellement des faits empiriques, c'est-à-dire des énoncés d'observation. Les propositions théoriques sont donc indirectement signifiantes du moment qu'elles font partie d'une théorie comprenant des énoncés d'observation bénéficiant d'une confirmation empirique.

Le nouveau critère de signification cognitive établit que, pour autant qu'un système dans son ensemble permette de prédire nos expériences avec plus de succès, c'est-à-dire permette de déduire des énoncés d'observation à partir d'une théorie, ses prédicats théoriques doivent être acceptés comme étant empiriquement signifiants. Si l'on veut déduire quelque chose d'un ensemble de postulats empiriques, il faut en outre disposer d'axiomes logiques et mathématiques. Ces axiomes sont analytiques et n'établissent, donc, aucun fait. Ainsi tout énoncé qui est scientifiquement signifiant n'est pas pour autant un énoncé factuel. La perspective empirique classique affirmant que les énoncés factuels sont à confronter de manière individuelle à l'expérience sensible est ainsi abandonnée. C'est le système scientifique dans son ensemble qui a un contenu factuel.

Par ailleurs, Putnam relève que la conception positiviste est auto-réfutante. En effet, les termes clés de "signification cognitive" et de "non-sens" ne sont ni des termes d'observation, ni des termes théoriques d'une théorie physique, ni des termes logiques ou mathématiques qui sont les seules sortes de termes admis dans le langage de la science. "In short, if it is true that only statements that can be criterially verified can be rationally acceptable, that statement itself cannot be rationally acceptable".¹⁰

Finalement, les travaux de W.v.O. Quine ont remis en question la notion d'analyticité. L'idée qu'il est possible de classer tous les énoncés, y compris les énoncés mathématiques, en tant que "factuel" ou "conventionnel", c'est-à-dire analytique, est un échec selon lui. Quine essaie de faire sens de la notion d'analyticité et de voir comment on peut tracer la limite entre ce qui est

¹⁰Putnam. 1981. p.111

analytique et synthétique. Il commence par examiner la définition de l'analyticité établissant qu'un énoncé est analytique lorsqu'il est vrai en vertu du sens de ses termes, indépendamment de tout fait et se demande quel est le concept de sens qui y est employé. Il faut distinguer le sens de la référence, puisque deux termes peuvent avoir la même extension et référer à un même objet sans avoir le même sens, comme le montre l'exemple de Frege d' "étoile du matin" et d' "étoile du soir". En effet, l'identité entre ces deux termes a été établie par une observation et si "étoile du matin" et "étoile du soir" partageaient le même sens, l'identité entre les deux serait analytique. Quine pense que c'est une erreur de penser que le sens est une entité et que cette erreur découle de l'échec à distinguer le sens de la référence. Selon lui, une théorie de la signification devrait s'occuper de la synonymie des formes linguistiques et de l'analyticité des énoncés. Le sens, en tant qu'obscure entité intermédiaire, devrait être abandonné. Le problème de l'analyticité resurgit donc, puisqu'on ne saurait la définir en faisant appel à la notion de sens.

Quine distingue ensuite deux classes d'énoncés analytiques. La première comprend des énoncés qui sont logiquement vrais tel que "Aucun homme non marié est marié." Cet énoncé demeure vrai quelle que soit l'interprétation que l'on donne de ses composantes mis à part les particules logiques. La seconde classe comprend des énoncés comme "Aucun célibataire n'est marié." Cet énoncé peut être transformé en une vérité logique en remplaçant un synonyme par un autre, dans le cas présent, en remplaçant "célibataire" par "homme non marié". C'est alors vers la notion de synonymie qu'il faut se tourner.

Quine envisage la suggestion selon laquelle la synonymie de deux formes linguistiques consiste en leur interchangeabilité *salva veritate*. La question est de savoir si cette condition est assez forte ou si des expressions non synonymes peuvent être aussi interchangeables. Quine relève que l'interchangeabilité *salva veritate* ne veut rien dire tant qu'elle n'est pas relativisée à un langage. Il constate que l'interchangeabilité *salva veritate* est garante de la synonymie cognitive dans le cadre de langages extensionnels. En effet, dans un langage extensionnel, lorsque deux prédicats réfèrent au même objet, ils sont interchangeables *salva veritate*. Ce langage est constitué de prédicats à une ou à multiple places et d'opérateurs logiques. Cependant, dans un tel langage, le fait que des termes comme "célibataire" et "non marié" aient une même extension n'est pas forcément garanti par la signification des termes. Cela pourrait être un hasard purement contingent comme dans le cas de "créature avec un coeur" et de "créature avec un rein". Ainsi l'accord extensionnel, qui est la meilleure approximation de la synonymie, n'est pas suffisante pour

expliquer l'analyticit  qui requiert que, n cessairement, "c libataire" et "non mari " partagent une m me extension.

So we must recognize that interchangeability *salva veritate*, if construed in relation to an extensional, is not a sufficient condition of cognitive synonymy in the sense needed for deriving analyticity [...] If a language contains an intensional adverb 'necessarily' in the sense lately noted, or other particles to the same effect, then interchangeability *salva veritate* in such a language does afford a sufficient condition of cognitive synonymy; but such a language is intelligible only if the notion of analyticity is already clearly understood in advance. (Quine, 1951, p.30)

Quine se tourne ensuite vers la notion de r gle s mantique. En effet, la difficult  que l'on  prouve   s parer clairement les  nonc s analytiques des  nonc s synth tiques pourrait d couler du caract re vague du langage ordinaire. La distinction deviendrait claire dans le cadre d'un langage artificiel b n ficiant de r gles s mantiques explicites. Or, Quine soutient que, m me dans le cas des langages artificiels, le probl me d'une explication de l'analyticit  demeure. Par exemple, en construisant un langage artificiel contenant une r gle sp cifiant quels  nonc s sont analytiques dans ce langage, on pr suppose une compr hension de ce qu'est l'analyticit . On peut n anmoins construire un autre type de r gle s mantique ne contenant pas le terme "analytique".

A semantic rule of this second type, a rule of truth, is not supposed to specify all the truths of the language; it merely stipulates, recursively or otherwise, a certain multitude of statements which, along with others unspecified, are to count as true. (Quine, 1951, p.33)

On d finit ensuite l'analyticit  en disant qu'un  nonc  est analytique s'il est vrai en vertu de cette r gle s mantique. Le probl me c'est qu'au lieu de faire appel au terme inexpliqu  "analytique", on emploie la phrase inexpliqu e "r gle s mantique". En somme, Quine pense que les r gles s mantiques d terminant les  nonc s analytiques dans un langage artificiel sont int ressantes uniquement lorsque la notion d'analyticit  est pr alablement comprise.

Quine admet que la v rit  d pend g n ralement   la fois du langage et de faits extralinguistiques. Il donne l'exemple de la phrase "Brutus a tu  C sar". Cet  nonc  serait faux si le monde avait  t  diff rent, mais aussi si "tuer" voulait dire "engendrer". C'est de ce constat,  crit Quine, que vient la tentation de penser que la v rit  d'un  nonc  est analysable en une composante linguistique et un autre factuelle. Les  nonc s analytiques seraient ceux o  la

composante factuelle est nulle. Toutefois, à la lumière de son raisonnement, il pense avoir montré qu'il faut se débarrasser de la croyance en une distinction claire entre analytique et synthétique.

But, for all its a priori reasonableness, a boundary between analytic and synthetic statements simply has not been drawn. That there are such a distinction to be drawn at all is an unempirical dogma of empiricists, a metaphysical article of faith. (Quine, 1951, p.34)

Sans distinction entre analytique et synthétique, nous ne disposons plus alors de notion claire de ce qui constitue un fait. C'est pour cette raison, selon Putnam, que l'argument classique en faveur de la dichotomie fait-valeur doit être abandonné. En affirmant que la distinction analytique- synthétique n'est pas pertinente, Quine remet aussi en question le fait que les énoncés puissent être vérifiés individuellement. C'est la science en entier qui doit être confrontée à l'expérience. Lorsqu'il y a un conflit avec l'expérience, on procède à un réajustement du champ scientifique.

No particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of equilibrium affecting the field as a whole. (Quine, 1951, p.40)

Cela implique que nous ne pouvons pas parler du contenu empirique d'un énoncé individuel. Par ailleurs, n'importe quel énoncé peut être tenu pour vrai, pour autant que l'on fasse des ajustements drastiques ailleurs dans le système. Il est donc difficile, selon Quine, de voir où l'on pourrait tracer la limite entre énoncés synthétiques et analytiques. Cependant, Putnam pense qu'il y a des énoncés qui sont analytiques et d'autres vérifiables empiriquement et il estime que Quine est allé trop loin en niant toute pertinence à une distinction entre analytique et synthétique. Ce qu'il retient de l'argument de Quine, c'est que les catégories d'analytique et de synthétique ne sont pas exhaustives. Si elles ne sont pas exhaustives, cela remet en cause la division tripartite des énoncés établie par les positivistes entre énoncés analytiques, énoncés synthétiques empiriquement vérifiables et énoncés dépourvus de signification. Or si cette division ne tient plus, il n'y a plus de raison de penser que les énoncés évaluatifs ne possèdent pas de signification en raison de l'impossibilité de les vérifier. Aucun énoncé ne saurait être vérifié de manière individuelle, c'est la théorie dans son ensemble qui doit être testée.

1.2.1 *La notion de fait est-elle neutre au niveau axiologique?*

Nous avons vu que la dichotomie fait-valeur telle qu'elle est défendue par les positivistes repose sur leur définition de ce qu'est un fait. Les propositions factuelles sont empiriquement vérifiables, contrairement aux énoncés évaluatifs, et ainsi ils possèdent une valeur de vérité. Il y a deux types d'énoncés signifiants: les propositions analytiques et les propositions factuelles. Les énoncés évaluatifs sont eux dépourvus de signification littérale. La notion de fait est aussi liée à la science puisque les faits peuvent également être définis comme étant ce qui est étudié par les sciences de la nature. Pour remettre en question la dichotomie, on peut se demander si les faits sont réellement dépourvus de toute connotation axiologique.

Il existe différents types de valeurs et celles auxquelles on pense le plus fréquemment sont les valeurs éthiques et esthétiques. Mais il ne faut pas oublier les valeurs épistémiques comme la cohérence et la simplicité, par exemple. L'argument de Putnam repose sur le fait que la science présuppose des valeurs épistémiques. Sans disposer, au préalable des valeurs comme la cohérence, nous ne pourrions pas nous engager dans une pratique scientifique. Ces valeurs sont des guides lorsque nous devons nous décider entre deux théories rivales. Cependant, Putnam ne nie pas qu'il y ait une différence entre les valeurs épistémiques et éthiques.

Les valeurs épistémiques sont des indices de vérité. Elles nous permettent de choisir entre différentes théories scientifiques alternatives et nous guident dans notre réflexion sur des recherches passées. Nous faisons d'ailleurs appel à ces mêmes valeurs lorsque nous disons qu'elles nous permettent de décrire le monde d'une manière plus correcte qu'en faisant appel à des oracles.

Lorsqu'une théorie entre en conflit avec ce qui était sensé être un fait, on abandonne parfois la théorie, parfois le fait supposé. La décision est une question pragmatique et c'est donc une question de jugement informel de cohérence, de plausibilité, de simplicité. Il n'est pas toujours le cas, lorsque deux théories entrent en conflit, qu'on attend des données tirées de l'observation pour faire un choix. Il n'y a pas de règle formelle ou de méthode pour prendre une décision. Un grand nombre de théories doivent être rejetées pour des raisons n'ayant rien à voir avec l'observation, simplement parce qu'il est impossible de tester toutes les théories. Putnam donne en exemple le choix qui s'est effectué entre la théorie de la gravitation d'Einstein et la théorie de Whitehead de 1922. Les deux théories s'entendaient sur la relativité spéciale et étaient

capables d'établir des prédictions concernant, entre autres, le caractère non newtonien de l'orbite de Mercure et l'orbite exacte de la lune. C'est la théorie d'Einstein qui a été acceptée au dépend de celle de Whitehead plus de cinquante ans avant qu'on puisse trouver une observation permettant de trancher entre les deux. C'est donc sur d'autres critères que ceux de l'observation que le choix s'est effectué et Putnam pense que ces critères concernent la cohérence et la simplicité.

La justification en termes de plausibilité, de cohérence, de simplicité est présupposée dans l'activité d'acquérir de la connaissance, même dans le cas de la physique. Dans ce cas, ce qui rend une croyance scientifique justifiée, c'est qu'on y est parvenu par une méthode fiable, c'est-à-dire une méthode qui a une grande probabilité d'arriver à une hypothèse vraie. Selon Putnam, la notion de vérité nécessite des critères d'acceptabilité rationnelle et ce sont ces critères qui reposent sur des valeurs épistémiques. Sans le concept de rationalité, la notion de fait serait vide. L'activité de justification d'énoncés factuels présuppose ainsi des jugements de valeur. Si nous voulons éviter le subjectivisme par rapport aux énoncés factuels, nous devons considérer ces jugements de valeur comme étant capables d'être objectifs. Selon Putnam, il faut cesser d'assimiler objectivité et description. Il y a des jugements qui ne sont pas des descriptions mais qui sont néanmoins soumis au contrôle rationnel. Cela implique des critères appropriés en fonction du contexte d'emploi.

I have argued that even when the judgments of reasonableness are left tacit, such judgments are presupposed by scientific inquiry. (Indeed, judgments of coherence are essential even at the observational level: we have to decide which observations to trust, which scientists to trust- sometimes even which of our memories to trust.) I have argued that judgments of reasonableness can be objective, and I have argued that they have all of the typical properties of value judgments. In short, I have argued that my pragmatist teachers were right: "knowledge of facts presupposes knowledge of values." (Putnam, 2002, p.145)

Pour savoir si une théorie scientifique est acceptable, nous regardons si elle manifeste des valeurs épistémiques. Il serait erroné d'analyser la procédure de construction d'une théorie comme une procédure de vérification phrase par phrase. On ne confirme, ni n'infirme une phrase tirée d'une théorie isolément sans tenir compte des autres phrases auxquelles elle est associée. La vérification porte sur la théorie dans son ensemble, c'est tout le système qui doit passer le test de l'expérience. Nous pouvons constater que la science s'efforce de parvenir à une représentation du monde efficace au niveau instrumental, cohérente et le plus simple possible. Nous employons ces critères

pour construire une image théorique du monde empirique et à mesure que cette image se développe, nous révisons nos critères d'acceptabilité rationnelle.

Putnam s'oppose, en outre, à la tentation de tout vouloir expliquer par une perspective physicaliste. Pour lui, le fait que les valeurs ne soient pas réductibles en ces termes ne saurait être un problème. Il pense que le physicalisme ne parvient pas non plus à rendre compte de notions telles que la référence ou la vérité.¹¹ Putnam suppose une interprétation standard de la référence I où "chat" réfère aux chats et "cerise" aux cerises qui soit coextensive, ou identique avec une relation physicaliste R liant les instances de "chats" aux chats. Il imagine ensuite une interprétation non-standard J coextensive avec une relation R' . Cette relation R' se définit en terme de R et des mondes possibles et des permutations employés pour construire J . L'interprétation J "desagrees with I , but which makes the same sentences true in every possible world as I does."¹² La relation R' lie les instances de "chat" avec les cerises. Les relations R et R' sont les deux des relations de correspondance.

The same sentences are 'true' under both correspondances. The actions called for by the R' -truth of a sentence (i.e. the actions which will 'succeed' from the agent's point of view) are the same as the actions for when the sentence is R -true. If R is 'identical with reference'; if R, R' , and all the other relations which assign extensions to our words in ways which satisfy our operational and theoretical constraints are not equally correct; if R, R' and the others are not equally correct because one of them - R - just *is* reference; then that fact itself is an *inexplicable* fact from a physicalist perspective. [...] If there is nothing in the physicalist world-picture that corresponds to the obvious fact that 'cat' refers to cats and not to cherries, then that is a decisive reason for rejecting the demand that all the notions we use must be reduced to physical terms. For reference and truth are notions we cannot consistently give up. (Putnam, 1981, pp.143-144)

C'est pourquoi, selon lui, il est absurde d'exiger que toutes les notions soient réduites à des termes physiques. Pour les positivistes, la connaissance morale est impossible car les termes employés dans le discours moral ne sont pas des termes scientifiques. Putnam s'interroge sur ce que signifie réellement ce manque de scientificité. Il relève que la croyance qu'il existe une telle chose que la

¹¹Par contre les théories causales de la références sont compatibles avec le physicalisme puisqu'il y a un événement qui est le point de départ de la chaîne causale de la signification.

¹²Putnam. 1981. p.217

justice n'est pas à mettre dans la même catégorie que la croyance aux fantômes. Le sens de la justice n'est pas non plus à comprendre comme un sens para-normal qui nous permettrait de percevoir des fantômes. Les personnes affirmant que la justice existe ne sont pas en train de proposer d'ajouter un objet à la liste des objets reconnus par les théories scientifiques. Les termes comme "juste" et "bon" sont des concepts employés dans un discours qui n'est pas réductible à celui de la physique. Mais pour Putnam, cela ne signifie pas que l'éthique soit en conflit avec la physique. Pourquoi le discours éthique serait-il rendu illégitime sous prétexte qu'il n'est réductible à des termes physicalistes alors que le discours sur la référence ne l'est pas? Le langage de la physique est un langage adéquat aux fins de la physique. Il ne l'est pas pour l'éthique. L'investigation en éthique peut néanmoins être objective au sens qu'elle peut être soumise à des critères de rationalité. La science n'épuise pas toute la rationalité. "Talk of 'justice', like talk of 'reference', can be *non*-scientific without being *unscientific*."¹³ De plus, Russ Shafer-Landau, dans *Moral Realism*, relève que le physicalisme lui-même ne peut pas, en tant que théorie, être certifié par la physique.

Physicalism is an ontological claim that limits the occupants of the universe to physical stuff and physical relations. But asserting such limitation is a *metaphysical* undertaking. Physics may in principal, offer complete explanations of physical phenomena, but physics isn't in a position to pronounce on the issue of whether physical phenomena are the only phenomena there are. (Shafer-Landau, 2003, p.81)

Il me semble que Putnam, ici, a une position similaire à celle de Wittgenstein disant qu'il ne faut pas importer des concepts d'un jeu de langage dans un autre. La tentative de vouloir réduire ou même comprendre l'éthique en faisant appel aux termes de la physique serait une entreprise absurde. Il ne faudrait pas vouloir à tout prix regrouper des faits différents dans un même modèle explicatif. Putnam veut remettre en cause l'idée selon laquelle le seul discours légitime est celui réductible aux termes de la physique. Ce qui compte, selon lui, c'est de savoir si des critères de rationalité peuvent s'appliquer. Comme des critères de justification et de rationalité s'appliquent au domaine de l'éthique, il n'y pas de raison de penser que le discours moral est moins légitime que celui de la physique. Il y a cependant un moyen de réconcilier physicalisme et autonomie du discours moral que Putnam n'envisage pas. En effet, si l'on admet

¹³*Ibid.*, p.145

que les propriétés évaluatives surviennent sur des propriétés physiques, il est possible de rendre compte de l'existence de propriétés morales dans un cadre physicaliste. En outre, comme il y a deux niveaux bien distincts, les propriétés de bases et les propriétés survenantes, et que les propriétés survenantes ne sont pas réductibles aux propriétés de bases, le discours moral ne saurait être réduit au langage de la physique. Je reviendrai à ce problème au troisième chapitre.

1.3 La signification émotive

Nous avons vu qu'Ayer ne met pas au même niveau les énoncés métaphysiques et les énoncés éthiques. Si les premiers ne sont que du verbiage, les seconds possèdent une fonction, malgré leur absence de signification littérale. Ils servent à projeter nos émotions ou attitudes sur le monde; ils possèdent, dit Ayer, une signification émotive. Ayer ne veut donc pas éliminer le vocabulaire éthique de notre langage de la même manière qu'il propose de supprimer les discours métaphysiques. Il pense que les énoncés moraux sont utiles à une certaine fin. S'il dispose d'un critère pour distinguer les énoncés possédant une signification de ceux qui en sont dépourvu, le principe de vérification, existe-t-il un critère pour distinguer, parmi les énoncés dépourvus de signification, ceux qui ont une signification émotive de ceux qui n'en ont pas? Alexander Miller ne pense pas que cela soit le cas et que la théorie émotiviste devrait alors considérer que les énoncés éthiques sont, eux aussi, du verbiage.

L'émotivisme suppose qu'un énoncé comme "Le meurtre est un acte abject" a superficiellement la forme d'une affirmation, mais n'est pas une véritable affirmation puisqu'il n'exprime pas une croyance empiriquement vérifiable. Cet énoncé a cependant une signification profonde puisqu'il exprime une désapprobation morale envers le meurtre de la part du locuteur. Toutefois, tous les énoncés ayant une apparence d'affirmation ne possèdent pas de signification émotive. Certains, comme les énoncés métaphysiques, ne sont que pur verbiage. Miller se demande alors si l'émotivisme offre un critère pour distinguer ces deux types d'énoncés. Il propose d'examiner le critère suivant: un énoncé ne possède une signification émotive seulement s'il y a des observations possibles d'occurrences qui constitueraient la manifestation de l'émotion ou attitude pertinente. Le problème vient du fait qu'il est difficile de voir quelles occurrences comportementales pourraient constituer l'expression d'une émotion spécifiquement éthique. Un comportement observable peut manifester une désapprobation, mais il ne peut pas manifester une

désapprobation éthique. Une solution à laquelle pourrait recourir l'émotivisme serait de considérer que le jugement que le meurtre est un acte abject exprime simplement la désapprobation du meurtre, sans que cette désapprobation ait besoin d'être spécifiquement éthique. Nous aurions alors accès à des observations manifestant un comportement de désapprobation.

Miller pense que cette solution pose problème car le jugement que le meurtre est un acte abject est alors dans la même catégorie que le jugement que les anchois sont mauvais.

as ordinarily conceived, the judgment that murder is wrong does have a different sort of content from the judgments concerning the Spice Girls and jellied eels. Isn't this just the elimination of moral discourse by the back door? (Miller, 1998, p.113)

Miller ne pense pas que l'émotivisme puisse invoquer les raisons de l'expression d'un sentiment pour justifier l'existence d'une classe de jugements spécifiquement moraux. En effet, le compte-rendu d'Ayer des disputes morales implique que, s'il est possible de chercher à convaincre son interlocuteur de changer ses croyances concernant les faits, il n'est pas possible de le raisonner par rapport à ses émotions ou attitudes, puisqu'il s'agit d'attitudes non-cognitives. Cela implique que nos sentiments, contrairement à nos croyances, n'ont pas de base rationnelle. Ayer ne peut donc pas distinguer les jugements moraux des jugements de goût en admettant que s'ils expriment le même type de sentiments, ces sentiments sont néanmoins exprimés pour des raisons différentes.

1.4 Les actes de langage

La dichotomie fait-valeur, telle que défendue par l'émotivisme, nous l'avons vu, distingue les énoncés factuels des énoncés évaluatifs. Les premiers servent à décrire des états de choses et expriment des croyances, alors que les seconds servent à exprimer nos émotions ou nos attitudes et ne représentent donc pas des croyances. John Searle a remarqué que cette division est trop rigide et qu'elle ne permet pas de rendre compte des nombreux autres actes de langage.

Selon la conception positiviste, seuls les énoncés descriptifs possèdent une valeur de vérité. Comprendre le sens de ces expressions, c'est savoir sous quelles conditions la proposition qui les contient est vraie ou fausse. Dans le cas des énoncés évaluatifs, ils sont la projection des attitudes ou émotions du locuteur et sont donc subjectifs. Selon Searle, la raison sous-jacente à cette distinction, c'est que, selon le non-cognitivism, les énoncés évaluatifs remplissent une tâche

complètement différente des énoncés descriptifs. Ils ne décrivent pas des états de choses, mais expriment les attitudes ou émotions du locuteur, louent ou condamnent, prescrivent ou recommandent, etc. Si ces énoncés servent à des tâches différentes, il doit y avoir un fossé logique entre eux.

Or, pour Searle, le problème avec cette manière de voir, c'est qu'elle ne permet pas de rendre compte de façon cohérente des notions d'obligation, d'engagement, de responsabilité. Searle constate qu'il existe différents types d'énoncés descriptifs. Certains portent sur des faits bruts tels que "Cette voiture roule à 100 km/h" et d'autres sur des faits institutionnels comme par exemple "X s'est marié", "Y possède cinq dollars".

Though both kind of statements state matters of objective fact, the statements containing words such as "married", "promise", "home run", and "five dollars" state facts whose existence presuppose certain institutions: a man has five dollars, given the institution of money. Take away the institution and all he has is a rectangular bit of paper with green ink on it. (Searle, 1964, p.54)

On peut néanmoins se demander de quel type d'institution il s'agit ici. Pour y répondre, Searle fait appel à la distinction entre deux types de règle ou convention. D'un côté, nous avons les règles régulatrices qui régulent des comportements existants antérieurement comme, par exemple, les règles d'étiquette ou les règles de la circulation. De l'autre, nous avons les règles constitutives qui définissent de nouveaux comportements telles que les règles du jeu d'échecs par exemple. Les institutions dont parle Searle sont des systèmes de règles constitutives. Les faits institutionnels présupposent de telles institutions. De nombreuses formes d'obligation, de droit, de responsabilité sont institutionnels d'une manière similaire. Par exemple, faire une promesse, c'est se placer sous une obligation. Searle en conclut que la conception non-cognitiviste échoue à rendre compte des faits institutionnels. Les faits institutionnels existent au sein de systèmes de règles constitutives et ces systèmes de règles constitutives impliquent des obligations, des engagements et des responsabilités.

Selon Searle, la distinction entre énoncés descriptifs et évaluatifs recouvre au moins deux distinctions. D'une part, il faut distinguer entre différents actes de langage, c'est-à-dire différentes forces illocutoires. D'autre part, il faut distinguer les énoncés possédant une valeur de vérité de ceux qui n'en ont pas. Ainsi, la distinction entre énoncés descriptifs et évaluatifs sert seulement à distinguer deux familles d'actes de langage différentes, décrire et évaluer, mais elle n'est pas très

utile au sens où il existe de très nombreux autres actes de langage.

Par ailleurs, Searle voit un problème dans le fait de chercher à définir le sens d'un terme par l'acte de langage pour lequel il est employé comme le fait l'émotivisme ou le prescriptivisme. Ils pense que les philosophes qui procèdent ainsi oublient que le terme qu'ils cherchent à analyser peut figurer dans de nombreux actes de langages mis à part les assertions.

To point out the association between "good" and commendation is not to answer the question "What does 'good' mean?" but the question "What is it to call something good?" Thus the old philosophical question "What is it for something to be good?" has been confused with the question "What is it to call something good?" (Searle, 1962, p.429)

Ceci est une erreur car une analyse de "bon", par exemple, doit pouvoir rendre compte du fait que le terme conserve sa signification lorsqu'il est employé dans différents actes de langage. "Good' means the same whether I ask if something is good, hypothesize that it is good, or just assert that it is good."¹⁴ L'analyse émotiviste ou prescriptiviste ne fonctionne que dans le cas de l'assertion et ne tient pas compte de tous les contextes dans lesquels les termes qu'elle cherche à analyser peuvent figurer.

1.5 Les attitudes morales

Le non-cognitivism n'a pas seulement des difficultés à tenir compte de la variété de nos actes de langage. Judith Jarvis Thomson, dans *Moral Relativism and Moral Objectivity*, a mis en lumière le fait que le non-cognitivism a également du mal à rendre justice à la diversité et la richesse des concepts moraux. En effet, si les jugements moraux sont des expressions d'attitudes, il faut être en mesure de différencier ces différentes attitudes pour pouvoir rendre compte du fait que les jugements moraux concernant des actions admirables, généreuses ou désirables sont des jugements différents. En effet, si l'on exprimait la même attitude en jugeant que Tom est courageux et en jugeant que Tom est généreux, les deux jugements devraient être équivalents puisque faire un jugement moral, c'est exprimer une certaine attitude. Le problème vient du fait que nos attitudes ne semblent pas avoir une aussi grande diversité que nos concepts moraux. Le

¹⁴Searle, 1962, p.429

non-cognitiviste a deux possibilités face à cette question. Il peut dire que nous avons des attitudes distinctes face à l'admirable, la générosité et la désirabilité. Ou alors, il peut dire qu'il y a deux attitudes fondamentales, à savoir l'approbation et la désapprobation, qui admettent des degrés différents.

Si l'on considère la première alternative, on peut d'abord se demander si réellement nous manifestons des attitudes différentes lorsque nous approuvons la générosité ou le courage d'une personne. Et avons-nous la même attitude à chaque fois que nous jugeons qu'une personne est généreuse? Thomson ne pense pas que cela soit le cas. Elle prend comme exemple la croyance qu'une personne doit faire quelque chose. Dans cette situation, je peux espérer qu'elle fasse cette chose ou je peux être contente qu'elle le fasse. Je pourrais aussi juste me demander si elle va le faire ou n'avoir aucun intérêt pour la question.

[...] we don't always have one pro-attitude whenever we have a belief of the one kind, and a second whenever we have a belief of the second kind, and a third whenever we have a belief of the third kind.
(Thomson, 1996, p.107)

L'émotiviste peut répondre à cela qu'il ne faut pas penser que les attitudes diffèrent au niveau phénoménologique. Le fait d'avoir une pro-attitude liée au courage face à l'action de Tom reviendrait à croire que Tom est courageux. Ce faisant l'émotiviste garanti que les attitudes liées au courage et à la générosité diffèrent. Cependant, Thomson relève que cette thèse ne nous aide pas à résoudre le problème. En effet, si l'on dit que croire que Tom est courageux, c'est manifester l'attitude liée au courage, on peut encore demander quelle est cette attitude. Si on répond qu'avoir cette attitude, c'est croire que Tom est courageux, nous ne sommes pas bien avancés.

La seconde alternative considère que les jugements moraux sont à diviser entre les jugements favorables et les jugements défavorables. Le problème de cette conception, selon Thomson, c'est qu'il faut disposer de notions d'approbation et de désapprobation spécifiquement morales. Il s'agit là du même problème qui a été soulevé par Miller concernant la possibilité d'établir un critère de signification émotive. L'approbation face au courage d'une action n'est pas exactement la même attitude que l'approbation ressentie à l'écoute d'une interprétation de telle symphonie. S'il n'y a pas d'autre moyen de définir l'approbation morale que de dire qu'avoir cette attitude envers une chose, c'est croire un énoncé moral favorable, nous faisons à nouveau face à une circularité. En outre, il devient difficile de différencier le jugement que Tom est courageux

du jugement que Tom est généreux puisque dans les deux cas le jugement consiste en l'approbation du comportement de Tom.

Les problèmes résultant de ces deux réponses possibles sont dus, selon Thomson, à l'incapacité du non-cognitivism à prendre en compte la richesse de la pensée morale et des différentes sortes de jugements moraux. Nous disposons d'un vocabulaire très riche en termes éthiques, mais notre vocabulaire pour décrire nos attitudes est beaucoup plus réduit. Si nos jugements moraux ne servent qu'à exprimer des attitudes, comment rendre justice aux différences, parfois subtiles, apportées par le choix de tel terme plutôt qu'un autre?

1.6 Le problème Frege-Geach

Le non-cognitivism doit, en outre, faire face au problème Frege-Geach. Ce problème, soulevé par Peter Geach, repose sur des contextes où les jugements moraux ne sont pas employés en tant qu'affirmations et donc ne peuvent pas être associés à une attitude exprimée. Par exemple, dans le cas d'un conditionnel, comme "Si c'est mal de mentir, alors je dois m'abstenir d'induire les gens en erreur en gardant le silence.", je n'affirme pas qu'il est mal de mentir. Donc, je ne suis pas en train de désapprouver le mensonge. Si la théorie d'Ayer concernant les énoncés moraux est correcte, il est impossible de concevoir une logique morale. En effet, un énoncé éthique peut figurer dans le discours en tant qu'affirmation, mais, parfois, il n'est pas employé comme une affirmation. "[...] a proposition may occur in discourse now asserted, now unasserted, and yet be recognizably the same proposition."¹⁵ Prenons par exemple l'inférence suivante qui semble faire partie de nos pratiques évaluatives ordinaires:

- (1) C'est mal de mentir
- (2) Si c'est mal de mentir, alors je dois m'abstenir d'induire les gens en erreur en gardant le silence.
- (3) Je dois m'abstenir d'induire les gens en erreur en gardant le silence.

Le même énoncé "c'est mal de mentir" figure dans (1) où il est affirmé et dans (2) où il n'est pas affirmé car il apparaît dans un conditionnel. Pour que les prémisses (1) et (2) soient vraies, il faut

¹⁵Geach, 1965, p.449

non seulement que l'énoncé ait une signification, mais l'énoncé doit aussi avoir la même signification dans les deux prémisses. Pour que ce raisonnement ait un sens, "c'est mal de mentir" devrait signifier la même chose dans (1) et (2). Mais pour Ayer, (1) n'a pas de signification et exprime une désapprobation du mensonge alors que (2) n'exprime pas de sentiment. Si on accepte l'émotivisme, on ne peut donc pas effectuer ce type d'inférence morale.

[...] the very same sentence can occur in an "if" clause; and to such occurrences the anti-descriptive theories will not apply. For example, in saying, "If what the policeman said is true, then..." I am not confirming or agreeing with what the policeman said; in saying "If he hit her, then..." I am not ascribing the act to him, and still less giving some moral or legal verdict about him; in saying "If that looks red, then..." I am not even tentively asserting that the thing is red. (Geach, 1965, p.462)

This possibility of varying use, however, cannot be appealed to in cases where an ostensibly assertoric utterance of "*p*" and "If *p*, then *q*" can be teamed up as premises for a *modus ponens*. Here, the two occurrences of "*p*" by itself and in the "if" clause, must have the same sense if the *modus ponens* is not to be vitiated by equivocation; and if any theorist alleges that its ostensibly assertoric occurrence "*p*" is really no proposition at all, it is up to him to give an account of the role of "*p*" that will allow of its standing as a premise. (Geach, 1965, p.463)

Le compte-rendu standard de la validité des *modus ponens* fait appel à la notion de vérité. Il n'y a pas de monde possible où les prémisses sont vraies et la conclusion est fausse. Mais puisque le non-cognitiviste nie que les énoncés moraux ont une valeur de vérité, il ne peut pas accepter cette façon d'expliquer la validité des *modus ponens* moraux.

1.6.1 *La solution de Blackburn*

Simon Blackburn a tenté de trouver une solution expressiviste au problème soulevé par Geach. Il a offert une première réponse dans *Spreading the Word*, puis est revenu à cette question dans un essai de 1988 publié à nouveau dans *Essays in Quasi-Realism* ainsi que dans *Ruling Passion*.

Ce que Blackburn veut montrer, c'est qu'il y a des attitudes qu'on ne saurait avoir simultanément de manière cohérente. Par exemple, il serait aberrant de promouvoir *p* tout en promouvant non-*p*. Dans un *modus ponens* contenant des énoncés moraux, on ne peut pas

approuver les deux prémisses et désapprouver la conclusion. Si une personne avait une telle sensibilité, on ne pourrait pas l'approuver car elle serait fracturée.

The 'cannot' here follows not (as the realist explanation would have it) because such a sensibility must be out of line with the moral facts it is trying to describe, but because such a sensibility cannot fulfill the practical purposes for which we evaluate things. (Blackburn, 1984, p.195)

Pour Blackburn, lorsque je dis "Si je ne dois pas mentir, alors je ne dois pas induire les gens en erreur en gardant le silence.", j'exprime mon approbation envers une sensibilité morale combinant la désapprobation du mensonge et la désapprobation d'induire les gens en erreur en gardant le silence. Mon emploi du conditionnel sert donc bien à exprimer une attitude, mais il s'agit d'une approbation envers une sensibilité morale dans son ensemble.

Blackburn introduit des opérateurs permettant de formaliser les attitudes d'approbation et de désapprobation qu'il note (H!, B!) pour "hourah" et "boo". L'exemple proposé plus haut serait alors formalisé comme suit.

(1') B! (Mentir)

(2') H! [B!(Mentir)]; [B! (Induire en erreur en gardant le silence)]

(3') B!(Induire en erreur en gardant le silence)

Un agent qui approuverait les deux prémisses sans approuver la conclusion échouerait à avoir une combinaison d'attitudes que lui-même approuverait. En effet, il lui manquerait la combinaison de désapprobation du mensonge et de désapprobation d'induire en erreur en gardant le silence alors qu'il approuve cette combinaison. C'est le fait que cet agent, manifestant un ensemble d'attitudes fracturé, commettrait une erreur logique qui permet à Blackburn de dire que sa formulation employant des opérateurs d'approbation ou de désapprobation est un exemple de *modus ponens*.

What I have done here is to explain how conditionals can be regarded as ways of following out implications, although it is not imperative that the commitments whose implications they trace have have 'truth-conditions'. (Blackburn, 1984, p.195)

L'argument de Blackburn repose donc sur le fait qu'une personne ayant un ensemble d'attitudes fracturé est coupable d'une erreur logique. Les objections à la solution de Blackburn vers lesquelles je vais maintenant me tourner, celles de G.F. Schueler et de Mark van Roojen, nient que cela soit le cas. Ces auteurs pensent que Blackburn emploie une notion de l'incohérence logique qui est trop large et qui permet ainsi d'englober ce qui n'est que de l'incohérence pragmatique.

Dans son article "Modus Ponens and Moral Realism", Schueler critique la réponse de Blackburn et explique qu'il ne pense pas que le compte-rendu de Blackburn permette de préserver le *modus ponens* en tant que règle d'inférence valide. En effet, il pense que le statut du conditionnel n'est pas claire, mais qu'il ne s'agit pas d'une instance de la forme logique $P \rightarrow Q$. Schueler reprend un exemple donné par Blackburn:

(A) Le courage est une chose intrinsèquement bonne.

(H) Si le courage est une chose intrinsèquement bonne, alors les jeux organisés devraient faire partie du curriculum des écoles.

(C) Les jeux organisés devraient faire partie du curriculum des écoles.

Il note ensuite que

"In short, Blackburn's idea is that, though *H* appears to represent a hypothetical connecting the sentences labeled *A* and *C* above, in fact it is a claim about one attitude "involving" another. But on this reading the logical form of the reasoning from *A* and *H* to *C* will just be P, R, Q , that is, not modus ponens. So Blackburn has not shown that, on his version of anti realism, the reasoning which goes from *A* and *H* to *C* can be shown to be valid as an instance of modus ponens. (Schueler, 1988, p.495)

Schueler ne pense pas non plus que l'on puisse parler d'incohérence logique entre tenir *A* et *H* mais nier *C*. On peut dire qu'une personne possédant cet ensemble d'attitudes commet une erreur ou qu'elle échoue à faire ce qu'elle devrait faire. Or ceci n'est pas équivalent à l'incohérence logique.

As these words are standardly used I am involved in logical inconsistency if I believe (or assert, etc.) A set of sentences which are such that, as a matter of logic, not all of them can be true. But the whole point of Blackburn's antirealist account of moral language is that in uttering sentences such as *A*, *H* and *C* we are not doing these sort of things, we are "expressing attitudes or commitments" and the like. And it is just false to

say that violating or failing to keep a commitment involves any *logical* inconsistency in the standard sense of these words. (Schueler, 1988, p.496)

Blackburn est par la suite revenu sur sa réponse au problème soulevé par Geach. Dans *Ruling Passion*, il explique qu'une théorie expressiviste est capable de rendre justice à des *modus ponens* employant des termes moraux. Dans le cas d'un *modus ponens* factuel, une personne qui affirmerait les prémisses, mais rejetterait la conclusion manifesterait un déficit de compréhension. Il pense qu'il en va de même lorsque les prémisses ont un contenu moral.

the person represents themselves as tied to a tree of possible combinations of beliefs and attitudes, but at the same time represents themselves as holding a combination that the tree excludes. So what is given at one moment is taken away at the next, and we can make no intelligible interpretation of them. (Blackburn, 1998, p.72)

Lorsque nous employons des énoncés moraux dans des contextes non assertés comme les conditionnels, nous nous attachons à un arbre d'engagements. Dans le cas d'un *modus ponens* où l'on accepte les prémisses et non la conclusion, les engagements se liraient comme un engagement à croire que *p*, un engagement soit à croire que non-*p* ou soit à croire que *q* et un manque d'engagement à croire que *q*. Cet ensemble d'engagements est incohérent et, selon Blackburn, incohérent d'un point de vue logique.

Dans son article "Expressivism and Irrationality", van Roojen discute lui aussi de la réponse offerte par Blackburn au problème Frege-Geach. Il relève que toutes ses réponses impliquent une interprétation, à savoir que les phrases employant une terminologie non affirmée expriment des attitudes d'ordre supérieur concernant les jugements évaluatifs. Blackburn soutient que les conditionnels réfèrent à des attitudes qui semblent faire partie d'un tout plus large. Il s'agit d'affirmations qui semblent faire un énoncé factuel à propos d'états de choses mais qui font en fait un énoncé factuel à propos des attitudes envers ces états de choses. La phrase "Si je dois m'abstenir de mentir, alors je dois m'abstenir d'induire les gens en erreur en gardant le silence" énonce qu'une attitude de désapprobation du mensonge implique une attitude de désapprobation envers le fait d'induire les gens en erreur en gardant le silence. La validité de ce type d'inférence est expliquée par le fait que dans des circonstances où l'on exprime une attitude envers un certain

état de choses et que l'on énonce aussi que cette attitude implique une autre attitude, on serait logiquement incohérent si l'on ne possédait pas l'attitude impliquée. Si l'on désapprouve le mensonge et que l'on croit qu'une telle désapprobation implique la désapprobation d'induire les gens en erreur en gardant le silence, on ne peut pas ne pas désapprouver l'acte d'induire les gens en erreur en gardant le silence sans être logiquement incohérent.

La réponse de van Roojen repose, également, sur le questionnement de l'emploi fait par Blackburn de la notion de "logiquement incohérent". Normalement l'incohérence logique implique des croyances qui se contredisent et il n'est pas question d'une telle incohérence dans l'explication fournie par Blackburn. Blackburn semble soutenir qu'il y a des attitudes qui ne peuvent pas être exprimées ensemble de manière cohérente et qu'il est donc irrationnel d'avoir les deux à la fois. Néanmoins, selon van Roojen, cette sorte d'incohérence n'est pas logique. Il pense que tout compte-rendu expressiviste qui cherche à expliquer comment des jugements évaluatifs peuvent être incohérents avec d'autres jugements doit étendre la notion d'incohérence au-delà de la notion habituelle s'appliquant à des jugements dont les contenus ont une valeur de vérité. Les expressivistes tels que Blackburn pensent que leur notion d'incohérence leur permet de soutenir la même chose qu'une interprétation réaliste des expressions évaluatives. En effet, le réaliste n'a pas de problème pour rendre compte d'une contradiction entre "Mentir est toujours une bonne chose" et "Mentir n'est pas toujours une bonne chose". L'expressiviste, lui, doit disposer d'une traduction qui tient ces jugements pour incohérents. Cependant, la notion de contradiction employée par les expressivistes va au-delà de la conception orthodoxe puisque l'expressivisme permet à des jugements n'ayant pas de conditions de vérité incompatibles d'être contradictoires.

Le projet d'expliquer la validité logique de ce type d'inférence morale devient plus facile en étendant la notion d'incohérence logique. Plus le nombre d'énoncés comptant comme incohérents est grand, plus on est en mesure de donner des explications de la validité logique de ces inférences. Van Roojen, néanmoins, relève qu'il y a une distinction entre cohérence logique et cohérence pragmatique et que l'expressivisme est incapable de conserver cette distinction. Il note également que, par ailleurs, l'incohérence pragmatique ne concerne pas que les jugements évaluatifs. "Il pleut" et "Je ne crois pas qu'il pleut" ne sont pas logiquement contradictoires mais ne sont pas cohérents au niveau pragmatique.

Selon van Roojen, le problème général vient du fait que les implications logiques entre des

contenus et les connexions rationnelles entre les attitudes exprimant ces mêmes contenus ne sont pas liés de manière directe. P ou Q et non (P) impliquent toujours Q mais croire que P ou Q et croire que non (P) peut être ou ne pas être une raison de croire que Q. Il ne faut pas, selon van Roojen, expliquer l'implication entre contenus en se basant sur des connexions rationnelles entre les attitudes exprimant ces contenus. Il n'est pas toujours irrationnel d'accepter tous les membres d'une ensemble de croyances incohérentes. Par exemple, s'il faut un très grand nombre de croyances pour générer l'incohérence et, s'il n'y a pas de raison de se concentrer sur une croyance en particulier et de la considérer comme moins probable, il est possible qu'il soit rationnel d'accepter toutes ces croyances. Van Roojen termine son article sur ces mots:

We thus have two reasons for pessimism about the noncognitivist's approach to logic. By extending what is to count as inconsistency, the approach threatens to class consistent views as logically inconsistent. And relatedly, rational connections between beliefs and logical connections between their contents are two different things, which impinge upon one another but are yet distinct. These considerations strengthen the suspicion broached by earlier critics of Blackburn's accounts, that logic is one thing, pragmatic incoherence and the rationality of beliefs and desires another. Attempts to explain one in terms of the other will distort our account of which propositions are consistent with which, or our account of which inferences are rationally warranted, or very likely both. (Van Roojen, 1996, p.335)

Le problème avec la solution de Blackburn au problème Frege-Geach, c'est qu'il est difficile de voir pourquoi, dans le cas d'un *modus ponens* moral, une personne qui approuverait les prémisses mais rejetterait la conclusion serait coupable d'une erreur de logique et non d'une erreur simplement pragmatique ou morale. On peut penser que si je suis d'accord avec les prémisses "C'est mal de mentir" et "Si c'est mal de mentir, alors je ne dois pas induire les gens en erreur en gardant le silence", mais que je désapprouve qu'il est mal d'induire les gens en erreur en gardant le silence, j'ai violé un principe moral. En effet, en refusant la conclusion, je serais prête à induire quelqu'un en erreur en gardant le silence, alors que ceci va à l'encontre d'un de mes principes moraux puisque j'approuve la prémisse "Si c'est mal de mentir, alors je ne dois pas induire les gens en erreur en gardant le silence".¹⁶

Par ailleurs, Blackburn donnerait sûrement un compte-rendu standard des *modus ponens*

¹⁶Cette idée que l'erreur est à situer sur un plan moral plutôt que logique est une idée de Bob Hale (1986) et elle est exposée dans l'ouvrage d'Alexander Miller. 2003. pp.63-64

non moraux affirmant qu'il est impossible que la conclusion soit fautive, alors que les prémisses sont vraies. Or cela signifie que Blackburn différencie la validité des *modus ponens* moraux de la validité des *modus ponens* non moraux. C'est seulement dans le cas où il contient des opérateurs moraux que le conditionnel "si...,alors..." reçoit une explication expressiviste. Blackburn ne peut donc pas offrir un compte-rendu uniforme des opérateurs logiques ou de la validité des inférences dans lesquelles ces opérateurs figurent. Cette absence de traitement uniforme implique, selon Miller, des ambiguïtés *ad hoc*. Miller pense que les occurrences de "si...,alors..." sont systématiquement ambiguës dans la solution de Blackburn puisqu'elles reçoivent une fonction sémantique différente selon qu'elles contiennent des opérateurs moraux ou pas. Or il n'y pas d'autre raison, selon lui, pour postuler cette ambiguïté que de sauver l'expressivisme d'une objection.

1.6.2 La solution de Gibbard

La réponse d'Allan Gibbard au problème Frege-Geach tente justement d'éviter de postuler des ambiguïtés *ad hoc*.¹⁷ Il veut offrir un compte-rendu uniforme des opérateurs logiques et de la validité des arguments. Pour se faire, il introduit la notion de monde factuel-normatif. Un tel monde est entièrement cohérent et ne contient aucune incertitudes que ce soit du point de vue factuel ou normatif. Un monde factuel-normatif comprend un système de normes et pas simplement des normes. Un jugement normatif dépend d'une pluralité de normes que nous considérons posséder une certaine force. Nous acceptons de ce fait que certaines normes puissent en supplanter d'autres. Pour rendre compte du poids et des priorités des différentes normes, Gibbard considère qu'un système de normes est un système de permissions et d'obligations qui s'applique à un ensemble de circonstances actuelles ou hypothétiques. Un système de normes N contient les prédicats "N-interdit", "N-optionnel", "N-obligatoire". "N-interdit" signifie "interdit par le système de normes N". Ces prédicats sont descriptifs, selon Gibbard, puisqu'on peut assigner une valeur de vérité aux énoncés dans lesquels ils figurent. La question de savoir si un système de normes donné permet telle action est une question factuelle. Un système de normes est considéré comme complet si toutes les possibilités sont couvertes par les trois prédicats "N-

¹⁷ Pour le compte-rendu de la position de Gibbard, je m'appuie sur Miller, 2003, pp.95-107

interdit”, “N-optionnel” et “N-obligatoire”.

Gibbard définit ensuite la notion de validité du *modus ponens* en disant qu’un argument est valide si chaque monde factuel-normatif où les prémisses sont acceptées est un monde où la conclusion l’est également. Reprenons notre exemple:

- (1) C’est mal de mentir
- (2) Si c’est mal de mentir, alors je dois m’abstenir d’induire les gens en erreur en gardant le silence.
- (3) Je dois m’abstenir d’induire les gens en erreur en gardant le silence.

Pour Gibbard, les jugements moraux sont des jugements à propos de quels sentiments moraux il est rationnel de ressentir. Les sentiments peuvent être appropriés ou non et les jugements moraux portent sur le caractère approprié de la culpabilité ou du ressentiment. On peut alors traduire l’argument précédent comme suit:

- (1a) Il est rationnel de ressentir de la colère envers un mensonge.
- (2a) S’il est rationnel de ressentir de la colère envers un mensonge, alors il est rationnel de ressentir de la colère envers l’acte d’induire les gens en erreur en gardant le silence.
- (3a) Il est rationnel de ressentir de la colère envers l’acte d’induire les gens en erreur en gardant le silence.

La validité de cet argument s’explique selon Gibbard par le fait qu’il n’y a pas de monde factuel-normatif où les prémisses sont vraies mais pas la conclusion.

- (1b) Ressentir de la colère envers les mensonges est N-permis.
- (2b) Si ressentir de la colère envers les mensonges est N-permis, alors ressentir de la colère envers l’acte d’induire les gens en erreur en gardant le silence est N-permis.
- (3b) Ressentir de la colère envers l’acte d’induire les gens en erreur en gardant le silence est N-permis.

Si l’argument de (1b) et (2b) à (3b) est valide, alors l’argument de (1a) et (2a) à (3a) l’est

également. En donc l'argument de (1) et (2) à (3) est aussi valide.

Cependant, on peut aussi se demander si le type de cohérence employée dans l'explication de Gibbard est une cohérence logique et non morale ou pragmatique. L'explication de Gibbard est-elle en mesure de prouver qu'une personne acceptant les prémisses, mais refusant la conclusion a commis une erreur de type logique? Si j'accepte l'existence d'un système de normes où il est permis de ressentir de la colère envers les mensonges, et où ceci implique qu'il est permis de ressentir de la colère envers l'acte d'induire les gens en erreur en gardant le silence, en quoi mon refus d'un système de normes où il est n'est pas permis de ressentir de la colère envers l'acte d'induire les gens en erreur en gardant le silence témoigne d'un faute logique. Mon défaut pourrait plus s'apparenter à de l'inconstance ou à un échec au niveau pratique ou moral.

Blackburn a aussi souligné une difficulté à la solution de Gibbard au problème Frege-Geach dans un article intitulé "Gibbard on Normative Logic". En effet, une solution expressiviste à ce problème doit être en mesure de prendre au sérieux l'aspect propositionnel des énoncés moraux, c'est à dire la surface du langage morale, sans invoquer de propriétés allant au-delà de l'explication expressiviste du discours moral.

A solution to the Frege-Geach problem must explain, and make legitimate, the propositional surface. But it must do this without invoking properties of evaluative discourse that go beyond the expressivist starting point. The constraint is quite delicate. For example it would not be met if we simply invoked the fact that evaluative judgements can be negated: negation is itself in need of a theory -it is itself part of the propositional surface, and the wrong side of the Fregean abyss. The constraint would not be met if we allow ourselves a notion of consistency and inconsistency, since that too is on the wrong side. All these notions must be built, if they can be, on the more primitive basis. (Blackburn, 1993b, p.61)

Or, Blackburn pense que Gibbard prend pour acquis la notion de système de normes, alors que cette notion devrait être construite à partir de notions plus primitives. De plus, Gibbard emploie la notion de cohérence des croyances normatives lorsqu'il définit ce qu'est un monde factuel-normatif. Blackburn pense que Gibbard n'est pas autorisé à employer les notions de cohérence ou d'incohérence pour définir les systèmes de normes. Ces notions devraient au contraire recevoir une explication dans un cadre expressiviste.

It is as if in response to Frege's charge that primitive representations cannot be negated we decided that what

they represent are systems of facts that themselves already stand in logical relations. (Blackburn, 1993b, p.62)

Gibbard affirme, par ailleurs, que les prédicats “N-interdit” ou “N-obligatoire” sont descriptifs puisque la question de savoir si une action est permise ou donnée dans un système de normes donné est une question factuelle. Cela signifie qu’un prédicat comme “N-interdit” concerne la relation entre un système de normes et une action, c’est-à-dire une relation entre des règles et des actions qui suivent ou ne suivent pas ces règles. Or, note Miller, ceci est une question normative puisque les règles nous disent comment nous devrions nous comporter. Si les règles et les normes sont elles-mêmes normatives, Gibbard devrait donner un compte-rendu expressiviste du fait de suivre une règle. Le prédicat “N-interdit” serait alors un prédicat normatif et non plus descriptif. La question est alors de savoir si une explication expressiviste du fait de suivre une règle est convaincante et si elle permet de conserver la solution de Gibbard au problème Frege-Geach.

1.7 Conclusion

Nous avons vu que le non-cognitivism fait face à de sérieux problèmes et que, donc, ses arguments en faveur de la dichotomie fait-valeur s’en trouvent fortement affaiblis. D’abord, la notion de fait, telle qu’elle employée par l’émotivisme, ne va pas de soi. Si celle-ci est remise en question, c’est l’ensemble de la théorie qui doit être remise en cause, puisqu’elle repose justement sur une division tripartite entre jugements factuels, jugements analytiques et jugements dépourvus de signification. Or, du moment que nous ne disposons plus de critères clairs pour établir ce qui doit compter comme énoncés factuels, il devient difficile de soutenir que les énoncés moraux sont dépourvus de signification cognitive en raison de notre incapacité à les vérifier empiriquement.

L’émotivisme ne parvient pas non plus à fournir de critère permettant de distinguer les énoncés moraux qui, s’ils ne possèdent pas de signification littérale, possèdent néanmoins une signification émotive, des énoncés totalement dépourvus de signification, comme les énoncés métaphysiques. En l’absence d’un tel critère, l’émotivisme devrait alors admettre que les jugements moraux ne sont que du verbiage et prôner l’élimination du discours éthique.

Le non-cognitivism a aussi du mal à rendre compte de la richesse de notre langage. Il peine à rendre compte de la diversité des actes langages. En effet, notre langage possède bien

d'autres actes de langage mis à part décrire et évaluer ou prescrire. Ceux-ci trouvent difficilement leur place au sein d'une théorie non-cognitiviste. En outre, le non-cognitivism a du mal à rendre compte de la richesse de notre vocabulaire moral. Si les jugements moraux servent à exprimer des émotions et attitudes, il faudrait que ces dernières soient aussi variées que nos concepts et prédicats moraux, ce qui ne semble pas être le cas. Par ailleurs, la solution consistant à diviser les attitudes morales en attitudes soit d'approbation, soit de désapprobation, n'est pas très prometteuse puisqu'elle ne permet pas de comprendre la spécificité de la désapprobation morale. Comme le note Miller, s'il n'y a pas de différence qualitative entre mon jugement que les anchois sont mauvais et mon jugement que le meurtre est abject, n'élimine-t-on pas la morale d'une manière détournée?

Le problème Frege-Geach est une sérieuse épine dans le pied du non-cognitiviste. Si un compte-rendu expressiviste des *modus ponens* moraux n'est pas disponible, cela implique que le non-cognitivism interdit toute inférence morale, pratique pourtant commune lors de nos réflexions éthiques. Nous avons examiné deux solutions possibles à ce problème. La première, celle de Blackburn, cherche à construire une logique des attitudes. Blackburn a bel et bien réussi à montrer que certaines combinaisons d'attitude généraient une incohérence pragmatique, mais cette incohérence serait du même type que la phrase "Il pleut, mais je ne le crois pas". En effet, l'incohérence entre des attitudes a à peu près la forme de "Le meurtre est mal, mais je ne le crois pas". Or cette sorte d'incohérence n'est habituellement pas comprise comme une incohérence logique. La phrase "Il pleut, mais je ne le crois pas" est d'habitude considéré comme cohérente mais auto-réfutante. On comprend généralement l'incohérence logique en faisant appel aux notions de vérité et de fausseté, notions auxquelles évidemment Blackburn ne peut pas recourir. La solution de Blackburn emploie donc une notion d'incohérence logique qui est trop large. Ainsi, pour que ce type d'inférence morale soit valide, il faudrait que les jugements moraux possèdent un contenu passible d'être vrai ou faux. Nous sommes donc en mesure de penser que les jugements moraux représentent des états de choses et ne servent pas uniquement à exprimer des états affectifs. Concernant la solution offerte par Gibbard, il aussi permi de douter qu'il emploie une notion de cohérence logique et non de cohérence pragmatique. Nous avons aussi vu que Blackburn a exprimé des doutes quant au fait que sa solution soit véritablement expressiviste.

Pour toutes ces raisons, il est permis de penser que le réalisme moral est une position méta-éthique plus prometteuse. Il est compatible avec la forme déclarative de nos énoncés moraux

et n'a pas de problèmes avec les contextes non assertés. Néanmoins, les non-cognitivistes soulignent que si les jugements moraux reportent des états de choses, il est difficile de comprendre comment ils pourraient nous motiver à agir. Or, il semble bien que les jugements moraux possèdent la faculté de nous motiver. Les non-cognitivistes soulignent que si des faits moraux pouvaient être motivants, ils seraient bizarres au niveau métaphysique. Je vais donc maintenant me tourner vers les questions liées à la motivation exercée par les jugements moraux.

2. Non-cognitivism et motivation

Selon les non-cognitivistes, les énoncés factuels et évaluatifs diffèrent aussi en fonction de leur lien avec la motivation. En effet, pour le non-cognitivism, les jugements moraux sont nécessairement motivants, alors que les jugements factuels ne possèdent pas un tel lien avec la motivation. Cela signifie que, du moment qu'une personne formule un jugement moral, elle est par là même, motivée à promouvoir ou empêcher ce qu'elle approuve ou désapprouve. L'internalisme considère que l'éthique est essentiellement pratique. En effet, nous jugeons généralement qu'une personne n'agissant pas en fonction de ses convictions morales est hypocrite, ce qui tendrait à prouver qu'il y a une connexion nécessaire entre émettre un jugement moral et être motivé en fonction de celui-ci. En outre, un changement dans les jugements moraux d'une personne apporte une modification au niveau de ses motivations. Si cet agent ne change pas de comportement, on ne pensera pas qu'il a réellement altéré ses jugements moraux. Or, d'après la théorie humienne de la motivation, les croyances seules ne suffisent pas à être motivantes. Si les jugements moraux sont nécessairement motivants, cela implique que les jugements évaluatifs n'expriment pas des croyances et ne représentent donc aucun fait. C'est pour cette raison qu'ils n'ont pas de valeur de vérité. Les jugements moraux, selon le non-cognitivism, ont plus à voir avec des désirs qui, eux, ont une efficacité au niveau de la motivation: ils sont l'expression d'états affectifs. En somme, l'argument non-cognitivistique repose sur deux thèses, à savoir la thèse internaliste et la théorie humienne de la motivation. Il a été formulé comme suit par Russ Shafer-Landau:

1. Necessarily, if one sincerely judges a action right, then one is motivated to some extent to act in accordance with that judgement. (*Motivational Judgement Internalism*)
2. When taken by themselves, beliefs neither motivate nor generate any motivationally efficacious states. (*Motivational Humeanism*)
3. Therefore moral judgements are not beliefs. (*Moral Non-cognitivism*) (Shafer-Landau, 2003, p.121)

Par contre, selon le cognitivism, les jugements moraux ont une fonction représentationnelle. Il prend au sérieux la forme descriptive des énoncés moraux. Les jugements moraux expriment des croyances morales et, de ce fait, ont des conditions de vérités. Le cognitivistique a deux possibilités. D'une part, il peut défendre une position externaliste et penser qu'il est possible qu'une personne formant un jugement moral sincère ne soit pas motivée par ce

dernier. Il peut admettre que les jugements moraux sont souvent liés à la motivation, mais nier qu'il s'agisse d'un lien de nécessité conceptuelle. D'autre part, il peut être internaliste et défendre que les jugements moraux sont intrinsèquement motivants. Il doit alors montrer que des croyances seules peuvent suffire à la motivation.

Ainsi, pour qu'une distinction entre jugements évaluatifs et énoncés descriptifs basée sur le lien avec la motivation puisse être maintenue, il faut montrer qu'en aucun cas une croyance seule peut être motivante ou alors, que le lien entre jugement évaluatif et motivation n'est pas contingent.

2.1 L'externalisme

Le problème de la motivation exercée par les jugements moraux est une difficulté à laquelle doit faire face le réaliste. En effet, il semble bien que les jugements moraux ont une dimension pratique dont sont dépourvus les jugements factuels. Le réaliste doit donc être en mesure de donner une explication à cela.

J.L Mackie a développé un argument apparenté contre le réalisme moral. Il constate que s'il existait des valeurs morales, celles-ci auraient la faculté de motiver. Comme il ne voit pas quelles sortes d'entités pourraient posséder une telle force motivationnelle, il en conclut que les valeurs n'existent pas. Il pense néanmoins que nos jugements moraux expriment des propositions portant sur des propriétés morales. Comme ces propriétés n'existent pas, nos jugements moraux sont systématiquement faux., "although most people in making moral judgments implicitly claim [...] to be pointing to something objectively prescriptive, these claims are all false".¹⁸ Le discours moral est, en ce sens, similaire aux énoncés portant sur les licornes ou les sirènes

Plato's Forms give a dramatic picture of what objective values would have to be. The Form of the Good is such that knowledge of it provides the knower with both a direction and an overriding motive; something's being good both tells the person who knows this to pursue it and makes him pursue it. An objective good would be sought by anyone who was acquainted with it, not because of any contingent fact that this person, or every person, is so constituted that he desires this end, but just because the end has to-be-pursuedness somehow built into it. Similarly, if there were objective principles of right and wrong, any wrong (possible) course of action would have not-to-be-doneness somehow built into it. (Mackie, 1977, p.40)

¹⁸Mackie 1977. p.35

On peut, cependant, remettre en question l'idée de Mackie selon laquelle le fait saisir une exigence morale induit une motivation surpassant tous les autres facteurs. On peut aussi douter que la motivation émane des propriétés morales elles-mêmes et penser, plutôt, que ce sont les jugements moraux qui nous fournissent des raisons d'agir, raisons qui peuvent entrer en conflit avec d'autres raisons. Un compte-rendu plausible de la motivation doit donc laisser de la place aux circonstances qui peuvent troubler la motivation découlant d'un jugement moral particulier.

L'internalisme permet de saisir la dimension pratique de l'éthique. La thèse internaliste énonce que si un individu juge sincèrement qu'il devrait *F*-er, il a alors une raison ou une motivation à *F*-er. Nous pensons, par exemple, généralement qu'une personne n'agissant pas conformément aux jugements moraux qu'elle prononce est soit hypocrite, soit irrationnelle. Nous constatons également qu'une personne changeant de convictions morales, sera motivée à accomplir d'autres actes que par le passé. Si, par exemple, une personne devient convaincue que le traitement que subissent les animaux destinés à la consommation humaine est barbare, on devrait normalement s'attendre à ce qu'elle devienne végétarienne. Michael Smith, dans *The Moral Problem*, souligne la dimension pratique des jugements moraux en constatant que:

Moral judgements seem to be, or imply opinions about the reasons we have for behaving in certain ways, and, other things being equal, having such opinions is a matter of finding ourselves with a corresponding motivation to act. (Smith, 1994, p.7)

La nécessité à l'oeuvre entre jugements moraux et motivation est, selon l'internaliste, une nécessité conceptuelle. Cela signifie qu'on ne saurait attribuer un jugement moral à un agent à moins qu'il soit motivé de manière appropriée. Puisque les jugements moraux motivent nécessairement, si l'on accepte la théorie humienne, ils ne peuvent pas être des croyances et donc ne représentent aucun état de choses. Comme je l'ai déjà mentionné, il y a deux manières de réfuter cet argument. On peut mettre en doute la première prémisse et défendre une position externaliste ou remettre en question la deuxième prémisse et s'opposer à la théorie humienne de la motivation. Je commencerai par considérer l'externalisme qui soutient qu'il est conceptuellement possible qu'un jugement moral sincère ne soit pas motivant. Dans ce débat, j'examinerai les arguments de Russ Shafer-Landau, de David Brink et de Sigrun Svavarsdottir en

faveur de l'externalisme ainsi que les objections formulées par Michael Smith. Shafer-Landau rejette à la fois l'internalisme et la théorie humienne de la motivation, contrairement à Brink, qui accepte cette dernière. Shafer-Landau pense que les jugements moraux sont intrinsèquement motivants, c'est-à-dire qu'il peuvent motiver sans la présence d'un désir préexistant, mais qu'ils ne sont pas nécessairement motivants. Selon Brink, nos jugements moraux sont motivants en vertu de faits psychologiques contingents. Ces faits dépendent des points de vue moraux des agents ainsi que de leurs désirs et attitudes. Brink admet que la plupart des gens désirerons agir en fonction de leurs jugements moraux, mais il nie que ce lien entre les jugements moraux et la motivation soit un lien de nécessité. Svavarsdottir, pour sa part, explique la motivation en faisant appel au désir de faire ce qui est moralement valable ou requis. Smith, quant à lui, accepte l'internalisme et la théorie humienne de la motivation mais défend néanmoins le réalisme moral. Il pense qu'il existe des raisons normatives découlant de ce que nous voudrions désirer. Smith pense que certains désirs sont rationnels. Si nous croyons avoir une raison normative pour accomplir une certaine action, alors rationnellement nous devrions faire cette action. Lorsque nous jugeons avoir une raison normative, alors, si nous sommes rationnels, nous serons motivés en conséquent.

L'externaliste nie qu'il existe un lien de nécessité entre jugement moral et motivation. Cependant l'indifférence morale de l'agent postulée par l'externalisme n'a pas besoin d'être systématique. Il suffit à l'externaliste de montrer, qu'à l'occasion, il est possible qu'un agent ne soit pas motivé par un jugement moral qu'il fait sincèrement.

Shafer-Landau envisage plusieurs scénarios exemplifiant cette possibilité. En premier lieu, les demandes morales pourraient être trop exigeantes en certains cas, ce qui aurait tendance à plutôt démotiver l'agent que le pousser à agir en fonction de ces principes moraux. L'agent n'aurait alors pas la volonté nécessaire pour agir de manière à mettre en pratique ses convictions morales.

Many have held that moral demands can be so strenuous as to be unsatisfiable in particular cases. It seems conceptually possible for an individual who holds this view to fail to be appropriately motivated. [...] When considering non-moral cases, it seems conceivable for agents to lose all motivation if they judge that accomplishment to be impossible. (Shafer-Landau, 2003, p.149)

Il existe aussi des cas où la prudence entre en conflit avec les exigences morales. En quel cas, il est possible que l'agent se plie à des considérations de prudence tout en continuant à avoir les mêmes jugements moraux. En effet, il est possible que la moralité et la rationalité soient bien distinctes l'une de l'autre et que leurs exigences respectives ne convergent pas toujours parfaitement. Si l'on considère, par exemple, que le point de vue moral est un point de vue impartial et que la raison pratique est, par contre, instrumentale ou prudentielle, il y aura des cas de figure où les demandes morales et rationnelles divergeront. Shafer-Landau donne, comme exemple de cas de conflit entre prudence et exigences morales, le cas d'un soldat croyant qu'il est de son devoir de se battre pour son pays. Il est, malgré tout, conscient des horreurs provoquées par la guerre et peut, par ce fait, perdre sa motivation au combat.

He recognizes the convenience of modifying his moral views. If he embraces pacifism, or at least came to morally oppose the present war, then he could with good conscience justify his unwillingness to fight. But he cannot do this. He thinks of himself a coward precisely because of the lack of motivation to do his perceived duty. By his own account he is entirely unmoved by that prospect of doing what he morally ought to do. (Shafer-Landau, 2003, p.150)

On peut également penser que les personnes dépressives ne peuvent pas être motivées par leurs jugements moraux car elles se trouvent dans une situation où elles manquent de motivation tout court. Il n'en demeure pas moins qu'elles continuent à former les mêmes jugements moraux qu'avant leur dépression. C'est l'état dépressif qui rompt le lien entre leurs jugements moraux et la motivation qui accompagnerait ceux-ci en temps normaux.

We can imagine an agent who views such obligations as perfectly reasonable, and as achievable by ordinary persons. But she is not such a person, or does not see herself as one. She is clinically depressed, apathetic, relatively affectless, or otherwise listless to an extreme degree. Such a person can issue non-moral evaluative claims and entirely fail to be moved by them, all the while wishing she were capable of such motivation. (Schafer-Landau, 2003, p.150)

De même, une personne ayant subi un traumatisme cérébral peut avoir une lésion lui ayant supprimé sa motivation tout en laissant intact ses jugements sur des questions morales. Cette personne continuerait à croire aux mêmes choses et à faire les mêmes jugements pratiques qu'avant son accident mais serait incapable d'être motivée de façon appropriée. La motivation

pourrait donc même être localisée dans certaines parties du cerveau, si des personnes souffrant de troubles similaires ont subi des lésions aux mêmes régions du cerveau. Antonio Damasio dans *L'erreur de Descartes*, a décrit le cas de Phineas Gage qui a été victime d'un accident au cours duquel une barre métallique lui a transpercé la tête¹⁹. Il a non seulement survécu à cet accident, mais ses facultés mentales comme l'attention, la perception, la mémoire, le langage et l'intelligence n'ont pas été altérées. Par contre, son comportement a fortement changé. Après l'accident il a perdu le sens des responsabilités sur les plans social et personnel. Damasio écrit:

Une certaine partie du système de valeurs persiste et peut être utilisée en termes abstraits, mais elle n'est pas mise en jeu dans les circonstances de la vie réelle. Autrement dit, il semble bien que, lorsque Phineas Gage avait besoin d'agir dans le contexte de la vie quotidienne, son processus mental de prise de décision n'était que peu influencé par les valeurs qu'il avait appris à reconnaître autrefois. (Damasio, 1995, p.29)

L'internalisme peut prendre en compte ces exemples en spécifiant que le lien entre jugements moraux et motivation ne tient que pour des agents rationnels, c'est-à-dire ni dépressifs, ni souffrant de faiblesse de la volonté, etc. En effet, les cas de figure évoqués plus haut sont des contre-exemples uniquement à une thèse forte de l'internalisme. Voici la formulation qu'en donne Smith:

If an agent judges that it is right for her to ϕ in circumstances C, then she is moved to ϕ in C. (Smith, 1994, p.61)

Smith est d'avis que cette forme d'internalisme est trop forte justement parce qu'elle nie la possibilité des cas de faiblesse de la volonté, dépression, etc qui font en sorte que l'agent perd sa motivation tout en conservant les mêmes convictions morales. L'internalisme faible, lui, admet qu'il y a une connexion conceptuelle entre les jugements moraux et la volonté, mais conçoit cette connexion comme étant défaisable.

If an agent judges that it is right for her to ϕ in circumstances C, then either she is motivated to ϕ in C, or she is practically irrational. (Smith, 1994, p.61)

¹⁹Damasio, 1995, pp.19-32

Cependant, selon l'externaliste, il est possible de concevoir des cas où des agents ne souffrant pas de troubles au niveau de la motivation ne sont pas motivés par leurs jugements moraux. Il s'agit du cas de l'amoraliste qui, bien qu'il reconnaisse les exigences morales, y demeure totalement indifférent. Si cette possibilité existe, cela montrerait que même des formes plus faibles de l'internalisme sont erronées.

L'internaliste pense que l'externaliste se trompe lorsqu'il affirme qu'il est possible de ne pas être motivé par ses jugements moraux. Dans les exemples donnés par ce dernier, nous n'aurions pas affaire à des jugements véritablement moraux. Par exemple, un amoraliste qui demeurerait indifférent à ses jugements moraux emploierait les termes moraux entre guillemets et ceux-ci n'auraient ainsi pas le même sens que lors d'un emploi standard. La notion d'emploi entre guillemets repose sur une distinction entre usage et mention. C'est Hare qui a, en premier, évoqué la possibilité que l'amoraliste fasse un emploi entre guillemets des termes moraux.

Thus it is possible to say "You ought to go and call on the So-and-sos" meaning by it no value-judgment at all, but simply the descriptive judgement that such an action is required in order to conform to a standard which people in general, or a certain kind of people not specified but well understood, accept. (Hare, 1952, p.164)

L'amoraliste, en employant les termes moraux ne ferait que les mentionner comme dans le cas du discours indirecte où l'on rapporte les propos d'une personne sans forcément endosser ce que celle-ci a dit. L'amoraliste serait ainsi indifférent aux standards moraux que les autres acceptent mais que lui-même ne considérerait pas comme moraux. Lorsqu'il dirait "Telle action est juste", il ne dirait pas que cette action est juste mais qu'accomplir cette action revient à agir en conformité avec ce que les autres jugent être juste. Puisqu'il ne considérerait pas les standards comme moraux, il ne formerait pas de jugements véritablement moraux en employant les termes y servant généralement. L'amoraliste ne serait en fait qu'un amoraliste apparent. En effet, il n'est pas motivé, non pas par ce qu'il pense être des considérations morales, mais par ce que les autres considèrent comme des considérations morales. Il ne ferait donc que parodier, en quelque sorte, le langage moral. Mais en est-il vraiment ainsi? Ne pourrait-on pas construire un exemple d'une personne non motivée par ses jugements moraux mais employant les termes moraux de manière standard? Selon Svavarsdottir, il est possible de penser qu'un agent capable d'entretenir une

discussion morale, de remarquer des inconsistances entre différents points de vue éthiques, de relever des différences pertinentes entre plusieurs cas, etc puisse néanmoins échouer à être motivé par ses jugements moraux. Lors de ses discussions, il n'emploierait pas les termes moraux d'une autre manière que son interlocuteur. S'il est le cas que des éliminativistes ou des nihilistes pensant qu'il n'y a pas de faits moraux font effectivement un emploi entre guillemets des termes moraux, il ne faudrait pas généraliser ce phénomène à toutes les personnes qui ne sont pas motivées par leurs jugements moraux.²⁰

Un argument pour défendre l'idée que l'amoraliste ferait un emploi entre guillemets des termes moraux consiste à établir une analogie entre jugements portant sur les couleurs et jugements moraux. David Brink, dans son article "Moral Motivation", prend en compte ce défi porté à l'externalisme tel qu'il a été proposé par Michael Smith dans *The Moral Problem*. L'idée est qu'il est nécessaire d'avoir des expériences visuelles des couleurs pour posséder les concepts de couleurs. Ainsi, une personne aveugle qui utilise des termes de couleurs, même si elle les emploie correctement, ne possède pas les concepts de couleur et, donc, ne porte pas véritablement de jugements sur les couleurs. Smith imagine le cas d'une personne aveugle dès la naissance qui ferait appel à une méthode fiable pour employer les termes de couleur et qui serait donc capable d'employer ces termes de manière similaire à une personne voyante.

However, despite the facility such a blind person has with colour language, many theorists have thought that we should still deny that she possesses colour concepts or mastery of terms. For, they say, the ability to have the appropriate visual experiences under suitable conditions is partially constitutive of possession of colour concepts and mastery of colour terms. And what such theorists thereby commit themselves to saying is that despite her facility with colour terms, such a blind person does not *really* make colour judgements at all. (Smith, 1994, p.69)

Selon Smith, de la même manière, le fait d'être motivé par ce que l'on juge éthique est en partie constitutif de la possession des concepts moraux. Ainsi l'amoraliste ne possède pas de concepts moraux puisqu'il n'est pas motivé par ses jugements. Il ne formule donc pas de jugements véritablement moraux.

Smith cherche à montrer que l'externaliste ne peut pas faire appel à la possibilité de l'amoraliste pour défendre sa position. Le fait d'interpréter les jugements de l'amoraliste comme

²⁰Svavarsdottir, 1999, pp.188-191

des jugements moraux sans apporter d'arguments supplémentaires reviendrait juste à accepter l'externalisme. Ce n'est pas un argument sa faveur. Smith établit une fois encore le parallèle avec les jugements de couleur. Il imagine un objecteur qui nierait que la capacité d'avoir des expériences visuelles soit en partie constitutive de la maîtrise des termes de couleur et qui affirmerait que l'on ne prend pas au sérieux les aveugles disant que l'herbe est verte et les fraises rouges.

Suppose the objector insists that since blind people can reliably use colour terms in this way, it just follows that they have full mastery of colour terms. Would the objection be a good one? I do not think so. For the objection simply assumes the conclusion it is supposed to be arguing for. It assumes that blind people have mastery of colour terms, something that those who think mastery requires the capacity to have the appropriate visual experiences under the appropriate conditions deny. (Smith, 1994, p.70)

Smith pense que Brink présuppose, de manière similaire, que l'emploi fiable des termes moraux dont fait preuve l'amoraliste est une marque de sa maîtrise de ces concepts et que, donc, Brink présuppose qu'une motivation appropriée n'est pas une condition à la maîtrise de ces concepts. On pourrait répondre à Smith, que l'attribution de concepts de couleur à la personne aveugle dépend aussi de quelle théorie du langage on adopte. Ainsi, un partisan d'une théorie wittgensteinienne du langage selon laquelle, le sens d'un terme c'est son usage, dira vraisemblablement que, si l'aveugle est capable d'employer les termes de couleurs de façon conforme à celle des voyants, il maîtrise l'emploi de ces concepts et, donc, formule des jugements de couleurs en bonne et due forme.

Brink répond à Smith que c'est plutôt l'internalisme qui lui semble peu plausible car il nie quelque chose qui lui semble possible.²¹ Il se défend de simplement faire appel à la possibilité de l'amoraliste sans apporter d'autres arguments pour soutenir l'externalisme. Selon lui, l'indifférence de l'amoraliste s'explique par le fait qu'on puisse douter de l'autorité de la morale. En effet, la motivation est liée aux croyances portant sur nos raisons d'agir. Or, selon lui, il est possible que des exigences morales manquent d'autorité rationnelle, puisqu'il dissocie moralité et raison pratique. Ainsi il est possible de former un jugement moral sans être motivé si le contenu de ce dernier ne coïncide pas avec ce qu'exige la raison pratique.

²¹Brink, 1997, p.23

Par ailleurs, Brink doute que l'analogie entre jugements de couleur et jugements moraux soit décisive. Pour commencer, cela ne va pas de soi que certaines expériences visuelles sont constitutives des concepts de couleur. Une personne née aveugle peut tout de même avoir appris un certain nombre de choses concernant les couleurs de ses semblables lui permettant d'employer les termes de couleur d'une manière faisant sens. Brink relève que même si l'emploi fait par un aveugle des termes de couleurs dépend du fait qu'il fait partie d'une communauté de langage de voyants, cela ne montre pas que son emploi des termes de couleur est un emploi entre guillemets. Dans ce cas, même si l'amoraliste n'est pas motivé de la même manière que le moraliste, cela ne signifie pas qu'il ne forme pas des jugements moraux.

Ensuite, même en acceptant que les personnes aveugles dès la naissance ne possèdent pas de concepts de couleur, il n'en va pas de même pour celles ayant perdu la vue suite à un accident. Lorsqu'une personne ayant perdu la vue récemment dit que tel objet est rouge, on peut difficilement nier qu'elle effectue un jugement de couleur. De même, selon Brink, "the principled amoralist need not be congenitally amoral."²² L'amoraliste pourrait très bien avoir été un moraliste dans le passé et en être venu à une indifférence face aux exigences morales suite à une remise en question de l'autorité de la morale, par exemple. Si cette personne continue à faire les mêmes jugements employant des termes moraux, pourquoi devrait-on dire que son emploi de ces termes est devenu un emploi entre guillemets. De même, les personnes ayant perdu la motivation suite à une dépression ou un dommage neurologique continuent à employer les mêmes termes moraux qu'avant. Elles ont les mêmes aptitudes cognitives et peuvent faire les mêmes discriminations et raisonnements moraux qu'avant. Elles ne sont simplement plus motivées et on ne saurait donc établir un changement de sens dans leur utilisation des termes moraux.

En outre, on ne peut pas nier que des personnes souffrant de troubles de la vision sélectifs, comme une incapacité à percevoir certaines nuances par exemple, puissent formuler des jugements de couleur. L'indifférence de l'amoraliste pourrait être pareillement sélective. S'il possède des conceptions de la morale et de la raison pratique faisant en sorte que leurs exigences ne soient pas toujours identiques, il est possible qu'il ne se conforme pas à certaines demandes morales sans que l'absence de motivation soit systématique. C'est seulement lorsque les exigences de la morale et de la prudence diffèrent que l'agent ne serait pas motivé par ses jugements moraux.

²²*Ibid.* p.24

Par ailleurs, Brink souligne que l'expérience visuelle n'est pas la même chose que la motivation. Si l'on pense que les expériences visuelles sont des états cognitifs alors l'analogie entre couleurs et morale montrerait que des états cognitifs sont en partie constitutifs des jugements moraux. Cela ne serait donc pas pertinent pour établir que la motivation est constitutive des jugements moraux. Il faudrait alors prendre pour hypothèse que les expériences visuelles sont des états précognitifs sur lesquels reposent les jugements représentationnels. Cela montrerait alors que les jugements de couleurs et les jugements moraux nécessitent des états précognitifs. Cette manière de voir présente plusieurs problèmes selon Brink. En premier lieu, il n'est pas convaincu que les expériences visuelles soient des états précognitifs plutôt que représentationnels. Ensuite, même si elles étaient des états précognitifs, cela ne signifierait pas qu'elles sont des états pratiques tels que des désirs.

Finalement, Brink ne voit pas pourquoi nous aurions des raisons de penser que les jugements moraux sont comme les jugements de couleur à moins d'être déjà convaincu par l'internalisme. On pourrait, par exemple, penser que les jugements moraux sont comme les jugements concernant la loi ou l'étiquette; ces derniers ne semblant pas nécessiter de composante non-cognitive. Brink est d'accord que les jugements moraux vont de pair normalement avec une motivation à accomplir ou empêcher ce que l'on approuve ou désapprouve. Il se différencie de l'internaliste dans la mesure où il ne pense pas que nous ayons affaire ici à une vérité conceptuelle. Pour lui, la morale a une dimension pratique uniquement lorsque ses exigences sont compatibles avec les demandes de la raison pratique. Ainsi la motivation morale aurait une importance secondaire puisqu'elle dépendrait en fin de compte de la raison pratique. Les questions importantes concerneraient alors les liens entre le point de vue moral et celui de la raison pratique.

Ce que l'externaliste tente de prouver, c'est que le lien entre jugement moral et motivation n'est pas un lien nécessaire, mais un lien contingent. L'externaliste ne nie pas qu'il y ait un lien particulier entre les jugements moraux et la motivation, ce qu'il conteste c'est que celui-ci soit nécessaire. Il y a ainsi des cas où l'efficacité à motiver des jugements moraux est effacée par d'autres facteurs. Néanmoins, l'externaliste doit être en mesure d'expliquer comment un changement au niveau de la motivation succède à une altération des jugements moraux s'il n'y a pas de lien nécessaire entre les deux.

Selon Svavarsdottir, le poids de l'explication repose plutôt sur l'internalisme qui devrait

être capable d'expliquer pourquoi il faudrait penser que le fait que les jugements moraux motivent nécessairement constitue une vérité conceptuelle²³. Si la thèse internaliste exprime une vérité conceptuelle, il doit y avoir au moins un concept qui rende l'explication externaliste incohérente. Il y aurait deux candidats pour un tel concept. D'une part, les concepts employés dans les jugements moraux tels que "bon", "juste". D'autre part, il y aurait le concept de jugement moral. Cependant, il ne faudrait pas que la thèse internaliste ait déjà été employée pour choisir le compte rendu approprié de ces concepts. Pour Svavarsdottir, l'intuition internaliste ne saisit pas correctement les concepts moraux mais exprime plutôt une conviction morale qui fait en sorte qu'il est difficile pour un agent de concevoir que quelqu'un puisse rester indifférent à ses jugements moraux. Si l'internalisme rend justice au fait que bien souvent nous sommes motivés en fonction de nos jugements moraux et qu'on s'attend d'une personne changeant de jugements moraux qu'elle soit motivée différemment, son erreur est de penser que ce qu'il ne peut pas envisager est inconcevable. Ce n'est pas parce qu'il y a souvent un lien entre motivation et jugements moraux que ce lien est nécessaire. En outre, l'intuition internaliste traduit un grand optimisme. En effet, selon celle-ci, du moment que l'on parvient à faire partager ses jugements moraux à autrui, sa motivation à agir est immédiatement garantie.

Selon Svavarsdottir, les personnes motivées par leurs jugements moraux possèdent bien un état conatif qui manque aux personnes non motivées par leurs jugements moraux. Cet état conatif pourrait être le désir d'être moral par exemple. Ce qui est nié, c'est que le fait de former un jugement moral soit suffisant à être motivé à agir en fonction de celui-ci. Il y aurait quelque chose de plus entre le jugement moral et la motivation. La présence d'un désir permettrait aussi d'expliquer les différences au niveau de l'intensité de la motivation entre des personnes possédant des jugements moraux similaires. "It is because the strength of their desire to do what is morally valuable or required differs; this is what underlies differences in their degree of moral commitment."²⁴ Ce que l'externaliste refuse, c'est que le désir fasse partie du contenu du jugement évaluatif. Il pense qu'il est possible de former un jugement moral sans avoir un tel désir.

David Brink est d'avis que l'externalisme est capable de rendre justice à la dimension pratique de l'éthique. Il admet que les jugements moraux ont normalement une incidence sur la

²³Svavarsdottir, 1999, p.181

²⁴Svavarsdottir, 1999, p.218

motivation, mais il ajoute que la motivation morale repose sur l'autorité rationnelle des exigences morales. "Morality will be practical in this sense just insofar as its requirements have rational authority."²⁵ Il différencie ensuite éthique et raison pratique et c'est pour cela que des exigences morales n'ont pas forcément l'autorité rationnelle. En effet, si l'on associe la moralité à un point de vue impartial et la raison pratique à une conception centrée sur l'agent, c'est-à-dire liée à une conception instrumentale ou prudentielle de la raison pratique, leurs exigences respectives ne coïncideront pas forcément. L'amoraliste se pose des questions sur l'autorité des exigences morales. Si ce que lui dicte la raison pratique n'est pas similaire aux exigences morales, il est possible qu'il forme un jugement moral sans être motivé en fonction du contenu de ce dernier.

On pourrait alors reformuler la position internaliste pour dire que celle-ci est d'avis que les exigences morales et les exigences de la raison pratique convergent et qu'il n'est pas concevable qu'elles puissent diverger. Cela implique que soit l'on conçoit la morale et la raison pratique en termes impartiaux, soit en termes centrés sur l'agent. On ne peut pas concevoir chacune en des termes différents. La question qui se pose alors est de savoir ce qui empêche de tenir à la fois une conception impartiale de la morale et une conception centrée sur l'agent de la raison pratique.

I see no reason to deny the possibility of the principled amoralist. If the rational authority of morality is an open question, then it's possible to make moral judgements without being motivated to act. If so, the (principled) amoralist is conceivable, even if her indifference can be shown to rest on mistaken conceptions of morality, rationality, or other auxiliary issues. If so, the externalist solution to the puzzle about moral motivation is the most plausible. But notice that if the externalist solution is correct, the noncognitivist solution is not only unnecessary but unavailable. For if it is possible to be unmoved by moral judgements, they cannot essentially express pro-attitudes, as the noncognitivist claims. (Brink, 1997, p.21)

Michael Smith a, cependant, formulé une autre objection à l'externalisme. Il rejette l'idée que la motivation soit toujours médiatisée par un désir d'être moral. Il pense que lorsqu'on vient en aide à un proche, par exemple, notre motivation est directement de l'aider. Nous n'agissons pas de la sorte en nous disant que nous désirons être moral et que l'action morale consiste, en ce cas, à venir en aide à cette personne. L'externaliste serait donc forcé de dire qu'une personne vertueuse n'a qu'un désir dérivé d'accomplir telle bonne action particulière, désir qui serait dérivé d'un désir

²⁵Brink, 1997, p.30

d'être moral. Le fait de n'agir que par désir d'être moral ne rendrait pas compte du comportement de la personne vertueuse qui serait directement motivée à faire ce qui est bien dans une situation particulière. En effet, selon Smith, le comportement d'une personne qui est motivée par la pensée de faire le bien, et non par la pensée qu'il faut sauver son proche, par exemple, ne saurait être qualifié de vertueux.

For commonsense tells us that if good people judge it right to be honest, or right to care for their children and friends and fellows, or right for people to get what they deserve, then they care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, not just one thing: doing what they believe to be right, where this is read *de dicto* and not *de re*. Indeed, commonsense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue. (Smith, 1994, p.75)

Smith pense qu'une bonne personne a une préoccupation directe et non dérivée pour le bien-être d'autrui, par exemple. Cette préoccupation est dite *de re*. Cette préoccupation ne saurait, selon Smith, est dérivée d'une préoccupation basique, *de dicto*, à faire le bien. Postuler un désir basique d'être moral aliénerait la personne vertueuse des fins qui sont véritablement poursuivies par l'éthique. Comme l'internalisme postule un lien conceptuel entre jugement moral et motivation, il est en mesure d'expliquer pourquoi la personne vertueuse est motivée à faire ce qui est bien.

David Brink répond à cette objection qu'il accepte qu'il soit plausible que l'éthique enjoigne de se préoccuper de soi et des siens. Ce qu'il ne comprend pas, c'est pourquoi le désir d'être moral et le fait de se soucier de soi et des siens sont incompatibles. Il voit là, plutôt, un conflit entre différentes motivations. Il admet qu'il serait pathétique qu'un mari ayant le choix entre sauver sa femme et sauver un étranger, doive passer par une longue délibération morale avant de se décider à sauver sa femme, mais il ajoute:

But that doesn't show that it is absurd for one's special attachments to be regulated in an appropriate way by moral concerns. For insofar as moral requirements have rational authority, it would be inappropriate to act on special concern for his wife where this was morally impermissible. (Brink, 1997, p.28)

C'est seulement dans des cas où il est permis de douter que le fait de sauver son proche soit moralement permis, comme par exemple dans le cas où ce choix entraînerait la mort de dizaines

d'innocents, que l'agent moral s'engagerait dans une délibération morale. Cependant, Brink envisage qu'on pourrait répondre à l'externaliste que, si je me soucie des miens uniquement parce que je désire agir moralement, la valeur que j'accorde à mes proches est dérivée de mon désir d'être moral et que, donc, il ne s'agirait que d'une valeur instrumentale. Or, Brink ne pense pas que le fait d'avoir des obligations spéciales envers nos proches, régulées par un désir d'être moral et des croyances morales, signifie que nous désirons le bien de nos proches en vertu du désir d'être moral.

The externalist needn't suppose that one cares about intimates in order to be moral; rather, when special concerns is morally regulated, one cares about one's intimates only in ways that are morally permissible. Here, morality acts as a constraint or filter on emotions, not as a goal at which emotions are directed. (Brink, 1997, p.29)

En outre, même si être moral représentait le but ultime, cela ne signifierait pas que l'attention particulière que nous portons à nos proches est instrumentale. On pourrait penser que une telle attention est à valoriser parce qu'elle est moralement requise.

It is not that these forms of special concern causally bring about dutiful action; it is, rather, that they are part of one's duty. So that if one were to value one's intimates for the sake of what's morally required, one would be according special concern contributory value that is intrinsic, not instrumental, value. (Brink, 1997, p.29)

James Dreier a lui aussi proposé une solution externaliste au problème soulevé par Smith dans un article intitulé "Dispositions and Fetishes: Externalist Models of Moral Motivation". Il propose de postuler des désirs de second ordre qui seraient le désir de valoriser les choses qui sont moralement bonnes. Le désir de sauver son proche serait alors causé par la croyance que c'est la bonne chose à faire. Toutefois, une fois que le désir de sauver son proche est généré, il n'est plus conditionné par la croyance que le fait de sauver son proche est la bonne chose à faire. L'agent agit alors par désir de sauver son proche. Les motivations de premier ordre sont causées par des motivations de second ordre, mais une fois qu'elles sont en place, elles ne dépendent plus des motivations de second ordre. Smith pense que nous devrions être motivés directement par les désirs de premier ordre sous peine de trahir la moralité. Or Dreier est d'avis qu'il n'y a rien de suspect dans le désir de second ordre de vouloir promouvoir le bien.

It is common ground that a person ought to want to be moved by right-making characteristics of actions, and that she ought to want this even if she is not sure what the right-making features of actions are. (Dreier, 2000, p.636)

Dreier soutient que nous n'avons pas besoin d'attribuer trop d'importance au désir de second ordre. S'il joue initialement un rôle causal en générant la motivation de premier ordre, il perd son importance une fois cette motivation formée. Dreier donne l'exemple d'un homme qui en vient à apprécier la justesse à oeuvre afin de mettre un terme à l'expérimentation animale. Une fois sa motivation en place, il se soucie authentiquement, et non de manière instrumentale, au sujet de l'expérimentation animale. Le fait de mettre un terme à ces expériences lui tient à coeur et il se préoccupe des moyens d'y parvenir. Il forme alors un nouveau désir de second ordre, à savoir le désir de second ordre de continuer à vouloir mettre un terme à l'expérimentation animale. Ce nouveau désir de second ordre est un désir *de re*, selon la terminologie de Smith, puisqu'aucune considération concernant la volonté de faire le bien ou d'être moral y figure.

Nous avons vu, pour résumer, que l'internalisme explique la dimension pratique de l'éthique en postulant un lien de nécessité conceptuelle entre jugements moraux et motivation. C'est ce lien qui expliquerait pourquoi, lorsque nous changeons d'avis sur une question éthique, nous sommes motivés à agir différemment et pourquoi nous ne prenons pas au sérieux les engagements moraux des personnes incapables de les traduire en actes. Cependant, l'externalisme ne nie pas qu'il existe un lien spécial entre les jugements moraux et la motivation. Ce qu'il remet en cause, c'est le fait que cette connexion soit nécessaire. Nous avons vu que les exemples de faiblesses de la volonté, de dépressions et de dommages neurologiques sont des contre-exemples à la thèse forte de l'internalisme. Pour contrer la thèse internaliste faible, l'externaliste fait appel à la possibilité de l'amoraliste, c'est-à-dire à la possibilité d'une personne jugeant sincèrement qu'il en son devoir de faire telle action mais qui n'est pas motivée à accomplir cette action. Nous avons vu que l'amoraliste n'a pas besoin de n'être jamais motivé par ses jugements moraux. Il suffit que parfois il ne le soit pas. Brink, en opposant les exigences morales et prudentielles, décrit une façon selon laquelle une personne, sincèrement convaincue de ce qui constitue son devoir, peut néanmoins ne pas être motivés à l'accomplir, si cela va à l'encontre de ses intérêts prudents.

2.2 La théorie humienne de la motivation

L'argument non-cognitivistique repose, outre la thèse internaliste, sur la théorie humienne de la motivation selon laquelle seuls les désirs ou les états conatifs ont la faculté de motiver. Des croyances seules ne peuvent donc pas être intrinsèquement motivantes. La motivation requiert toujours la présence d'un désir préexistant. Comme on peut constater un lien entre les jugements moraux et la motivation, c'est cette théorie qui permet d'affirmer que les énoncés évaluatifs ne peuvent pas exprimer des croyances et, par conséquent, qu'ils ne représentent pas des états de choses. La question qu'on peut alors se poser, pour remettre en question la conception non-cognitivistique, c'est de savoir s'il est vraiment impossible de trouver des croyances qui sont motivantes.

Selon Shafer-Landau, qui défend une position anti-humienne, les désirs ne sont pas les seuls états intrinsèquement motivants. Il pense qu'il existe des croyances capables de motiver par elles-mêmes et plus précisément que ces croyances sont les croyances évaluatives.

Evaluative beliefs ascribe value (moral or non-moral) to acts or states of affairs. Believing something desirable, taking oneself to have a reason to act, judging an option admirable or good, may prompt action even in the absence of antecedent desires. (Shafer-Landau, 2003, p.122)

Shafer-Landau trouve aussi des exemples où la théorie humienne semble une explication peu plausible. Il prend comme exemple le cas ordinaire où nous formons une croyance sur la base d'une certaine raison de croire. Il paraît superflu de nous attribuer un désir de former des croyances ou de faire des inférences. En effet, nous cherchons certaines raisons de croire, ou alors elles se présentent à nous, et nous en inférons une conclusion. Il pose un même diagnostic concernant les actes mentaux comme se souvenir, se faire du souci etc. Il s'agit bien d'actions, mais nous ne désirons pas faire ces choses-là. "Desires are not implicated in the most natural and straightforward explanation of such actions."²⁶ Un défenseur de la théorie humienne pourrait objecter à cela que croire n'est pas à proprement parler une action, ou que se faire du souci, c'est être dans un certain état émotionnel et non une action. Pour trancher, il faudrait voir quelle théorie de l'action est présupposée par chacun et déterminer laquelle paraît la plus pertinente, mais je

²⁶Shafer-Landau, 2003, p.124

n'entrerai pas dans ce débat ici.

Shafer-Landau cite, en outre, le cas où l'on s'attribue faussement un désir. Il est clair que nous faisons ce genre d'erreur et que l'explication la plus simple consiste à dire que nous avons une croyance fautive à propos de nos désirs. Afin d'illustrer une telle situation, Shafer-Landau imagine le cas d'une personne convaincue qu'elle désire devenir avocat. Il suppose que le père et le grand-père de cette personne étaient aussi avocats et que tous s'attendaient à ce qu'elle le devienne aussi. Cette personne ne s'épanouit pas à la faculté de droit, découvre suite à un emploi d'été qu'elle adore la menuiserie, quitte l'université et n'y pense plus par la suite. En réfléchissant rétroactivement, cette personne réalise qu'elle n'a jamais voulu devenir avocat.

Our action may sometimes involve us in ongoing aggravation or tedium, extreme physical pain, emotional contortions, suffering of another's stupidity or arrogance. Very few of us desire to experience such things. The Humean must explain our willingness to put up with these things by citing a further desire- a desire for some goal that we believe attainable via the unpleasantness. But this is impossible for those cases in which we mistake our desires. (Shafer-Landau, 2003, p.125)

Si nous supportons de telles souffrances, c'est parce que nous croyons désirer le but qu'elles permettraient d'atteindre. Dans le cas où nous nous trompons sur nos désirs, nous ne pouvons pas faire appel au désir d'atteindre ce but pour expliquer notre comportement. L'explication la plus simple consiste alors à dire que ce sont les fausses croyances qui nous ont motivé. La réponse d'un humien à ce genre d'exemple consistera à voir s'il n'y a pas d'autres désirs possibles qu'on pourrait invoquer pour expliquer le comportement. Dans l'exemple proposé par Shafer-Landau, on pourrait avancer que cette personne avait le désir de ne pas décevoir sa famille, par exemple.

Michael Smith, dans son article "The Humean Theory of Motivation" avance plusieurs arguments en faveur de la théorie humienne de la motivation. Philip Pettit a commenté cet article dans un texte intitulé "Humeans, Anti-Humeans and Motivation".

Smith commence par donner une formulation formelle du principe sous-tendant la théorie humienne. ϕ et ψ réfèrent à des verbes d'action.

(P1) R à t constitue une raison motivante pour l'agent A à ϕ ssi il y a un ψ tel que R à t consiste

en un désir de A à ψ et en une croyance que s'il ϕ -ait, il ψ -rait.²⁷

Ce principe implique que toute motivation trouve sa source en la présence d'un désir pertinent et d'une croyance concernant les fins et les moyens. Par exemple, si je désire faire des documentaires et que je crois que, pour cela, il me faut une caméra, alors je désirerai m'acheter une caméra.

Smith différencie ensuite les raisons motivantes des raisons normatives. Elles ont en commun de justifier le comportement de l'agent du point de vue de l'agent. Il y a un lien a priori entre citer les raisons que possède l'agent de se comporter d'une certaine manière et donner une justification partielle de sa manière d'agir. Mais les raisons motivantes et normatives établissent le lien entre justification et action de manières différentes.

Concernant la raison motivante il écrit "The distinctive feature of motivating reason to ϕ is that in virtue of having such a reason an agent is in a state that is *potentially explanatory* of his ϕ -ing."²⁸ Un agent peut avoir une raison motivante sans que celle-ci l'emporte sur d'autres raisons. En outre, les raisons motivantes sont disponibles à l'agent puisque celles-ci comprennent les buts de l'agent. En effet, note Smith,

[the] agent's motivating reasons are, as we might put it, *psychologically real*, for it would seem to be part of our concept of what it is for an agent's reasons to have the potential to explain his behaviour that his having those reasons is a fact about *him* (Smith, 1987, p.38)

Par contre, concernant les raisons normatives, Smith relève que:

It is to say that there is some normative requirement that he [the agent] ϕ 's. It is therefore to justify his ϕ -ing from the perspective of the normative system that generates that requirement. (Smith, 1987, p.39)

Ainsi les raisons normatives ne sont pas forcément accessibles à l'agent puisqu'il est possible que ce dernier échoue à adopter le point de vue du système normatif.

²⁷Smith, 1987, p.36. Je traduis

²⁸ Smith, 1987, p.38

Michael Smith souligne que le principe (P1) lie des raisons motivantes à la présence de désirs et de croyances et qu'il ne concerne pas les raisons normatives. C'est pour cette raison que Smith rejette l'objection de Thomas Nagel à la théorie humienne de la motivation telle qu'il l'a formulée dans *The Possibility of Altruism*. D'après lui, l'objection de Nagel ne prend pas en compte le fait qu'il n'est question que de raisons motivantes dans la théorie humienne.

Nagel se concentre sur la motivation prudentielle. Celle-ci fait en sorte que si l'agent reconnaît qu'il désirera *F*-er dans le futur, il dispose alors d'une raison maintenant de promouvoir le fait de *F*-er plus tard. Il faut donc qu'il désire *F*-er maintenant. La théorie humienne, pour rendre compte de ce phénomène, explique la motivation prudentielle par le fait que les agents possèdent un désir actuel de promouvoir leurs intérêts futurs. Nagel répond à cela qu'il est possible que je désire maintenant quelque chose pour le futur que je ne m'attends pas à désirer le moment venu. Je ne disposerai alors d'aucune raison de promouvoir ce quelque chose même si mon désir actuel me fournit maintenant une telle raison. Il est aussi possible que je m'attende à avoir un certain désir dans le futur. Dans le futur, je disposerai d'une raison pour promouvoir ce quelque chose, mais je ne dispose pas de cette raison maintenant puisque je n'ai pas encore le désir approprié. Nagel note que:

A system with consequences such as this not only fails to require the most elementary consistency in conduct over time, but in fact sharpens the possibilities of conflict by grounding an individual's plotting against his future self in the apparatus of rationality. (Nagel, 1970, p.40)

Nagel s'oppose à l'idée que la prudence s'explique par le fait que mes intérêts futurs me fournissent des raisons d'agir parce que je dispose d'un désir actuel de promouvoir ces intérêts.

A person's future should be of interest to him not because it is among his present interests, but because it is *his future*. He already stands in a far stronger and more important relation to his future and the desires he can expect to experience than could possibly be established by any desire which might assail him in the present. (Nagel, 1970, pp.42-43)

Pour lui, la présence d'un désir prudentiel n'est pas nécessaire pour faire le lien avec son futur soi parce que cette connexion est garantie par les conditions formelles de la raison pratique. Il ne pense pas que les désirs soient les principales sources de motivation. Il pense que le désir de

satisfaire ses désirs futurs est justifié et motivé par des raisons fournies par le fait qu'on s'attend à avoir ces désirs. Ce sont ces raisons qui sont à la base de la motivation prudentielle, selon lui, et non des désirs. L'influence des raisons peut s'étendre à travers le temps parce qu'il y a une raison à promouvoir ce pour quoi il y a ou il y aura une raison. Mais nous reviendrons à cette conception rationaliste plus tard.

Ce que Smith reproche à Nagel c'est qu'il amalgame raison motivante et raison normative dans ses exemples.

Nagel's objection to the Humean's claim that an agent may have a *motivating* reason now to promote his ϕ -ing in the future, despite the fact that he believes that he will have no motivating reason to ϕ then, is that it would be *irrational* to do so; that is, that he now has no reason from the perspective of rationality to do so. But this is to conflate the claim that an agent has a motivating reason to ϕ with the claim that he has a normative reason from the perspective of rationality to ϕ . The Humean is only the first claim, not the second. (Smith, 1987, p.42)

Smith s'oppose, en outre, à une conception phénoménologique des désirs. En effet, contrairement aux sensations, les désirs ont un contenu propositionnel. De plus, si nous considérons les désirs à long terme, nous ne sommes pas enclins à penser que la personne perd ses désirs lorsqu'elle n'éprouve pas la sensation qui y correspond. Smith relève que de nombreux opposants à la théorie humienne partent d'une conception phénoménologique des désirs. Smith cherche donc une conception des désirs qui prennent au sérieux le fait qu'il s'agit d'états ayant un contenu propositionnel. Smith défend une conception dispositionnelle des désirs:

we should think of the desire to ϕ as that state of a subject that grounds all sorts of his dispositions: like the disposition to ϕ in conditions C, the disposition to ϕ in conditions C', and so on (where in order for conditions C and C' to obtain, the subject must have, *inter alia*, certain beliefs). (Smith, 1987, p.52)

D'après cette conception, les désirs ont un contenu phénoménologique dans la mesure où une de leur disposition est de produire une certaine sensation, sous certaines conditions.

La différence importante entre désirs et croyances serait une différence quant à la direction d'ajustement, notion que Smith emprunte à Anscombe. Alors que nous cherchons à ajuster le monde à nos désirs, nous cherchons à ajuster nos croyances au monde. Or, être dans un état auquel le monde doit s'ajuster équivaut à avoir un but. Ainsi, pour Smith, désirer c'est avoir un but.

Comme avoir une raison motivante revient à avoir un but, on peut ainsi établir que le fait d'avoir une raison motivante revient à avoir un désir. En effet, Smith pense qu'on peut comprendre ce que c'est que d'avoir une raison motivante pour autrui en lui imputant un certain but. Le principe (P1) serait donc impliqué par les trois prémisses suivantes:

- (1) Having a motivating reason *is, inter alia*, having a goal
- (2) Having a goal *is* being in a state with which the world must fit
and
- (3) Being in a state with which the world must fit *is* desiring. (Smith, 1987, p.55)

Smith a donc montré que les désirs sont constitutifs de la motivation.

Philip Pettit a écrit une réponse à l'article de Smith. Il pense que Smith n'a pas porté son attention sur le point crucial départageant les humiens des anti-humiens. La question principale ne serait pas de savoir si les raisons motivantes impliquent toujours des désirs, mais de savoir si elles impliquent toujours la présence d'états non-cognitifs. Or Smith présume, sans examen supplémentaire, que les désirs sont des états non-cognitifs. Pettit voit deux manières de nier cela.

The first, a minority preference, is to hold that desires or, more precisely degrees of desire, are an agent's estimates of objective goodness or utility. This line is not generally favored, because it runs against the common assumption that it is only beliefs which are capable of having cognitive status. (Pettit, 1987, p.531)

La seconde manière de donner un contenu cognitif aux désirs est d'établir que les désirs présents dans les raisons motivantes y figurent conséquemment à la présence de certaines croyances. Les désirs hériteraient du statut cognitif de ce que Pettit nomme des croyances désidératives. Ces croyances désidératives portent sur ce qui est bon, approprié ou utile. Le désir que *p* serait, alors, impliqué par la croyance que *p* est bon, utile, etc. Le désir que *p* pourrait demeurer insensible à la perception qu'il n'est pas le cas que *p*, mais il sera sensible au fait que *p* n'est pas bon ou pas approprié. Le désir semble alors dépendre de la raison de l'agent puisque c'est sa raison qui lui indique que *p* n'est pas approprié.

In order for the Humean to establish his point of view he needs to be able to resist at least the possibility that desires can inherit cognitive status in this way. (Pettit, 1987, p.531)

Pettit suggère quelques pistes selon lesquelles un humien pourrait répondre à ses objections. Il pourrait nier qu'il existe des croyances désidératives. En effet, la seule preuve en faveur de l'existence de croyances désidératives est l'occurrence d'assertions désidératives. Or ces assertions sont mieux expliquées en tant qu'expressions de désirs. Une autre façon serait de nier que les croyances désidératives sont des états cognitifs puisqu'il n'y a pas de faits désidératifs et donc pas de perception désidérative.

Shafer-Landau a également émis des objections à la théorie de Smith. Pour commencer, il n'est pas d'accord sur le fait que posséder une raison motivante équivaut à avoir un but. D'après Smith, être motivé reviendrait à déterminer une fin et à réfléchir sur les moyens de la réaliser. Or, pour Shafer-Landau la motivation n'est pas toujours dirigée vers une fin. En effet, un agent pourrait être motivé par la croyance qu'accomplir une certaine action est de son devoir, par exemple. Ainsi toutes les motivations ne seraient pas calquées sur le modèle fin-moyens, puisque d'après les anti-humiens, le fait de voir que quelque chose est bon ou juste est suffisant à motiver l'agent. Il s'agit là d'une motivation déontologique.

The mark of what we might call deontological motivation is the absence of such an end state from the agent's understanding of what is going on. She believes that an action has a certain feature (being right or obligatory, for instance), and is moved to act on that basis alone. She needn't conceptualize some state of affairs and take the realization of that state as a goal to be attained by her action (Shafer-Landau, 2003, p.135)

Shafer-Landau est d'avis que cette analyse anti-humienne fonctionne même dans le cadre d'une morale téléologique. En effet, ce qui compte ce n'est pas de savoir si la morale en fin de compte implique une structure téléologique, mais de savoir si l'agent la conçoit comme telle. Or même si la morale avait une structure téléologique, il serait fallacieux, selon Shafer-Landau, de penser que tous les agents cherchant à faire le bien vont penser en terme de réalisation d'un état de choses positif. Ainsi, même si la morale a une structure téléologique, beaucoup ne le voient pas et donc de nombreux agents moraux n'ont pas l'intention de parvenir à une certaine fin lorsqu'ils accomplissent le bien. Ceci montrerait qu'un agent peut être motivé par une croyance évaluative seule. C'est la croyance évaluative qui va générer le désir pertinent à la motivation. "Beliefs are,

all by themselves, sufficient to yield motivated or derived desires, which, together with the evaluative beliefs, generate motivation.”²⁹

Shafer-Landau pense que lorsqu’il est question de délibération et de réflexion concernant ce que nous devrions faire, la motivation est expliquée par nos croyances. Ces croyances concernent ce que nous désirons, ce que pensons être désirable, important, bon, etc ainsi que les moyens d’y parvenir. Néanmoins, Shafer-Landau admet que toutes les actions intentionnelles ne sont pas délibératives et qu’il y a donc des fois où la motivation ne repose pas sur des croyances. Par exemple, si je prends un verre de jus parce que j’ai soif, cet action est déclenchée par un désir. Ainsi la théorie humienne est parfois la description correcte de la motivation.

What the anti-Humean denies is that the explanation of why we form intentions to act, and so act, will not always require citation of desire. In fact, invoking desires as explanations of action seems appropriate in only a minority of cases, namely, those that do not involve deliberation. Deliberative acts can be explained just by citing the combination of evaluative beliefs or beliefs about one’s desires, with means-end beliefs. Since many instances of moral motivation do involve deliberation, we have good reason to analyse such cases as the anti-Humean suggests. (Shafer-Landau, 2003, p.140)

Le débat entre humiens et anti-humiens ne me paraît pas amener d’argument décisif en faveur de l’un ou l’autre et il me faudrait sûrement approfondir les théories en philosophie de l’esprit concernant les croyances et les désirs pour pouvoir affiner ma compréhension de cet enjeu. Il me semble néanmoins, que même dans les cas où l’on a l’impression que c’est une croyance à propos de ce qui est bien ou à propos de nos désirs qui motive, on peut trouver un désir plus fondamental à l’arrière-plan. En effet, même si c’est ma croyance que cette action est de mon devoir qui me pousse à agir en conséquent, il faut que j’aie, tout de même, le désir préalable d’agir selon mon devoir.

2.3 Les théories rationalistes de la motivation

Certaines théories rationalistes de la motivation soutiennent que les jugements moraux sont intrinsèquement motivants tout en acceptant le cognitivisme. Selon ces théories, la

²⁹Shafer-Landau, 2003, p.137

motivation ne requière pas de désirs et des croyances peuvent donc être motivantes. Par exemple, Nagel avance une conception prudentielle de la motivation qui affirme qu'un agent est motivé par ce qu'il reconnaît promouvoir ses intérêts. Puisque le comportement est considéré comme intentionnel, l'action est le produit des croyances et ne dépend pas de désirs. Les désirs viendraient en tant que conséquences mais ne seraient pas une condition nécessaire à la motivation.

Shafer-Landau défend également que, lorsqu'il y a délibération et réflexion, des croyances évaluatives additionnées de croyances concernant les moyens et les fins sont suffisantes à la motivation. Ces croyances peuvent néanmoins générer des désirs dérivés. Shafer-Landau distingue en fait désirs et pro-attitudes. Les désirs sont un état mental plus particulier et les pro-attitudes désignent le fait que l'agent soit motivé. La classe des désirs est très diversifiée et c'est le fait d'être intrinsèquement motivants qui unifie cette classe. Les désirs comprennent les appétits, les souhaits, les aversions ainsi que les "passions calmes". La classe des désirs est intrinsèquement motivante, comme l'illustre l'exemple où je me sers un verre d'eau pour calmer ma soif. Par contre, les pro-attitudes ne désignent pas vraiment des états mentaux, mais une manière d'attribuer de l'intentionnalité à un agent. Nous attribuons aux agents des désirs pour indiquer qu'ils ont été motivés par un ensemble de croyances. Étant donné que les agents veulent faire ce qu'ils font de manière intentionnelle, Shafer-Landau admet qu'on pourrait croire que c'est le désir qui motive. Mais il se demande ce qui a causé l'intention en premier lieu. Il est d'avis que ce qui forme l'intention est un ensemble de croyances, dont la croyance que l'on a un désir et la croyance que la meilleure manière de satisfaire ce désir est d'agir d'une certaine façon.

Brink évoque la théorie rationaliste de la motivation dans son article concernant la motivation morale pour lui trouver des objections.³⁰ Il considère que la motivation requière des désirs ultimes faisant partie de la psychologie d'arrière-plan et soutient, par conséquent, la théorie humienne. Sans ces désirs, les croyances morales ne mèneraient nulle part. Pour illustrer son propos, Brink donne cet exemple-ci. Je crois qu'il pleut. Je veux sortir et désire rester au sec. Je crois qu'un parapluie me protégera de la pluie et qu'il y a un parapluie dans l'armoire. Brink accepte que le désir d'aller chercher le parapluie dans l'armoire est une conséquence de mes croyances. Par contre, il pense que la motivation repose sur un autre désir, à savoir mon désir de sortir et de rester sec. Selon lui, si l'on fait appel à ma croyance qu'il pleut pour expliquer mon action c'est seulement parce qu'on considère que mon désir de rester sec va de soi et qu'il fait

³⁰Brink, 1997, pp.13-15

partie de mon arrière-plan psychologique. Brink pense qu'il en va de même pour la motivation morale. Si on peut faire appel à des croyances, comme ma croyance qu'il faut être équitable, pour expliquer mon don à une ONG d'entraide, c'est parce qu'on prend pour acquis mon désir d'être équitable. Sans ce désir plus fondamental, ma croyance morale n'aurait aucun effet. Ainsi une croyance que la morale exige une certaine action ne pourra motiver que dans la mesure où elle est conjointe à un désir d'être moral.

We can agree that moral beliefs are sufficient to motivate the virtuous person, but this is only because a virtuous person is someone with a certain well-developed psychological profile that is structured by various cognitive and conative states. The virtuous person has beliefs about what justice, rights, mutual respect, friendship, and compassion require of agents and has standing aims to be just, to respect the rights and dignity of others, to be a true friend, and to display kindness. When the virtuous person adds to her background psychological states beliefs about what these moral categories require in particular circumstances, she may be motivated to act, but this will be in virtue of her cognitive and conative background and not simply because of her newly acquired cognitive states. (Brink, 1997, p.15)

Il existe une autre forme de rationalisme qui nie que la motivation soit possible sans la présence de désirs ou pro-attitudes, mais qui pense que certaines croyances impliquent des pro-attitudes. Ces croyances seraient les croyances normatives. C'est la croyance que je possède une raison d'agir qui générerait un désir d'accomplir cette action. Si la croyance que les choses devraient être de telle manière représente bien un état de choses, cette croyance possède néanmoins une dimension pratique puisque l'agent aimerait que le monde soit conforme au contenu de sa croyance. En croyant que les choses iraient mieux si elles étaient différentes, l'agent produit le désir de faire en sorte que les choses soient différentes.

On this version of rationalism, pro-attitudes are not merely consequentially ascribed because the action is intentional; the pro-attitudes are psychologically prior to the action and play an ineliminable role in generating and, hence, explaining action. But these pro-attitudes are consequential or dependent on the belief that one should perform the action. (Brink, 1997, p.15)

Brink note que deux conditions doivent être remplies pour que cette version du rationalisme soit vraie. Premièrement, une croyance qu'il y a une raison pour accomplir telle action doit générer

une pro-attitude envers cette action. Deuxièmement, la croyance qu'une action est moralement requise doit impliquer la croyance que cette action dispose d'une raison d'agir. Comme nous l'avons déjà vu, Brink ne pense pas que les jugements moraux expriment nécessairement ce qu'il est rationnel d'accomplir. Il est possible que les exigences de la morale et les exigences de la raison pratique entrent en conflit. La seconde condition n'est donc pas remplie selon lui.

2.4 Conclusion

Nous avons vu que l'argument non-cognitivistique repose sur la validité de deux thèses: la thèse internaliste et la théorie humienne de la motivation. En examinant la thèse internaliste, nous avons vu qu'elle faisait face à un certain nombre de difficultés. Premièrement, les cas de faiblesses de la volonté, de dommages neurologiques, de dépression montrent qu'il y a des situations où le lien habituel entre jugements moral et motivation est brisé. Ainsi une thèse forte de l'internalisme serait erronée. Deuxièmement, la figure de l'amoraliste illustre la possibilité d'un agent capable de former des jugements moraux authentiques, mais dénué de motivation. Il n'est pas nécessaire que l'amoraliste ne soit jamais motivé par ses jugements moraux, ce qui semblerait effectivement peu plausible. Il suffit qu'il existe certaines circonstances où la motivation exercée par les jugements moraux soit effacée au profit d'autres facteurs. Brink propose l'exemple du conflit entre éthique et prudence. L'amoraliste pourrait ainsi être un agent qui face à des considérations prudentielles, met en doute l'autorité des exigences morales et qui n'est pas irrationnel.

Comme je l'ai déjà mentionné, les arguments contre la théorie humienne de la motivation ne me semblent pas décisifs. Même s'il est possible d'expliquer le comportement moral en faisant référence à nos croyances, il est possible que la présence d'un désir fondamental d'être moral soit nécessaire en arrière-plan.

Néanmoins, il est suffisant que l'internalisme soit faux pour que l'argument non-cognitivistique échoue. En effet, il suffit qu'une des deux prémisses soit fausse pour que la conclusion, à savoir la thèse selon laquelle les jugements moraux ne sont pas des croyances, le soit également. L'externalisme soutient que les jugements moraux n'ont pas de lien nécessaire avec la motivation. On ne peut donc pas déduire que les jugements moraux ne sont pas des croyances. En effet, c'est la connexion nécessaire avec la motivation, en lien avec la théorie humienne, qui permet au non-cognitivistique de dire que les jugements moraux ne sont pas des croyances. Rien

n'interdit alors de penser que les jugements moraux représentent des états de choses. Cependant, pour que cela soit possible, il faut qu'il y ait quelque chose à représenter. En effet, s'il n'existe ni propriété, ni fait moral, on voit mal quel contenu pourraient avoir des croyances morales. Il est donc temps de se tourner vers la question du réalisme moral.

3. Le réalisme moral

La dichotomie fait-valeur, au niveau métaphysique, affirme que les faits et les valeurs appartiennent à des domaines ontologiques différents. Il n'y aurait rien dans le monde qui corresponde aux valeurs et donc, pas d'états de choses axiologiques. Au niveau sémantique, il y a deux possibilités lorsque l'on nie l'existence de faits moraux. D'une part, on peut considérer que les jugements de valeurs n'ont pas de valeurs de vérité et fonctionnent comme des exclamations ou sont des projections de nos émotions sur le monde. C'est la position non-cognitviste. D'autre part, on peut penser qu'ils sont systématiquement faux, car nous parlons comme s'il y avait des propriétés dans le monde auxquelles référerait nos termes de valeur, alors qu'il n'en est rien. Il s'agit de la théorie de l'erreur de J.L. Mackie développée dans *Ethics*. Mackie note que si les valeurs existaient, elles auraient la propriété d'être motivante. Comme il ne voit pas quelles sortes d'entités pourraient posséder une telle caractéristique, il conclut que les valeurs n'existent pas. Il pense, néanmoins, que nous cherchons à attribuer des propriétés à des objets lorsque nous énonçons des jugements moraux. Cependant, nos énoncés moraux échouent systématiquement à représenter des états de choses, puisqu'il n'y a rien dans le monde qui correspondent à nos termes évaluatifs. La différence entre Mackie et le non-cognitvisme, c'est que Mackie pense qu'un énoncé comprenant un terme ne référant à rien dans le monde est faux, alors que le non-cognitvisme est d'avis qu'un tel énoncé ne possède pas de signification.

Toutefois, s'il existait des faits moraux, il n'y aurait plus de raison de maintenir une séparation profonde, au niveau ontologique, entre faits et valeurs. Le réalisme moral défend que notre langage moral est descriptif et qu'au moins certains de nos jugements moraux sont vrais, parce qu'il existe une réalité morale qui les rend tels. Le réalisme moral se divise en deux familles: les naturalistes et les non-naturalistes. D'après les non-naturalistes, bien que les propriétés morales soient réelles, elles ne sont pas réductibles ou identiques aux propriétés naturelles. Le non-naturalisme affirme que les propriétés et faits moraux sont *sui generis*. Du fait que les non-naturalistes refusent que les propriétés évaluatives soient réductibles à des propriétés naturelles, ils prennent très au sérieux l'autonomie de l'éthique. Le réalisme moral naturaliste se divise aussi en deux familles: les réductionnistes et les non réductionnistes. Les naturalistes réductionnistes, comme leur nom l'indique, pensent que les propriétés morales sont réductibles à des propriétés naturelles. Pour éviter l'argument de la question ouverte, ils soutiennent que l'identité entre propriétés axiologiques et propriétés naturelles est du même type que l'identité entre le concept d'eau et la structure moléculaire H₂O. Il s'agit d'une identité synthétique qui se

découvre de manière *a posteriori*. L'investigation, dans le domaine moral, se ferait par la réflexion en éthique normative. C'est la meilleure théorie morale qui nous indiquera à quelles propriétés naturelles équivaut telle propriété morale. Par exemple, si l'utilitarisme de l'acte hédoniste s'avérait être la meilleure théorie, nous pourrions définir le bien en tant que plaisir ou absence de souffrance. Le réalisme naturaliste non réductionniste est représenté par le réalisme de Cornell qui soutient que les propriétés morales surviennent sur des propriétés naturelles ou sont réalisées par des propriétés naturelles. La différence entre ces deux formes de réalisme moral naturaliste porte sur une compréhension divergente de ce que signifie pour des propriétés morales d'être des propriétés naturelles. Les réalistes réductionnistes pensent qu'il y a une identité entre les propriétés morales et naturelles, alors que les réalistes non réductionnistes pensent que les propriétés morales sont constituées par des propriétés naturelles. Il ne s'agit pas d'une relation d'identité car une propriété morale particulière, dont on aurait déterminé la base naturelle, aurait aussi pu être réalisée par d'autres propriétés. La propriété d'injustice, pour prendre un exemple de David Brink, est réalisée dans une situation particulière par certaines conditions économiques et sociales.³¹ Mais la propriété d'injustice peut également être réalisée par d'autres facteurs économiques et sociaux. Ceci semble indiquer que les propriétés morales peuvent être réalisées par un grand nombre, peut-être un nombre infini, de propriétés naturelles. La thèse non réductionniste suppose donc que l'identité n'est pas une relation qui peut être obtenue entre des termes qui sont infiniment disjonctifs.

Dans ce chapitre, je commencerai par examiner l'argument de la question ouverte, proposé par G.E. Moore, à l'encontre du naturalisme éthique pour voir s'il est possible de soutenir que les propriétés évaluatives sont réductibles à des propriétés naturelles. Je me tournerai ensuite vers la théorie de la survenance. C'est en faisant appel à la survenance que les défenseurs du réalisme moral expliquent généralement comment des propriétés morales peuvent être réelles sans avoir à postuler l'existence d'entités mystérieuses, non détectables par nos sens. Si les propriétés morales surviennent sur des propriétés physiques, elles sont compatibles avec une conception du monde naturaliste. Cependant, Simon Blackburn a noté que la survenance pose un problème pour le réalisme moral. Je regarderai donc les arguments apportés par Blackburn ainsi que les réponses possibles à son objection. Terrence Horgan et Mark Timmons ont aussi soutenu que les réalistes naturalistes ne peuvent pas offrir un compte-rendu naturaliste de la relation de survenance.

³¹Brink, 1989, p.158

Enfin, pour que les faits moraux soient considérés comme des faits à part entière, il faut qu'ils puissent jouer un rôle explicatif. Je regarderai donc s'il est possible qu'ils puissent avoir un pouvoir causal et s'il est possible de donner des explications de nos observations en se passant de toute référence aux faits moraux.

3.1 L'argument de la question ouverte

G.E. Moore est célèbre pour avoir tenté de montrer, dans *Principia Ethica*, que les propriétés évaluatives ne sont pas naturelles grâce à l'argument de la question ouverte. Il pense que toute tentative de définir "bien" en termes de propriétés naturelles est un échec. Si le bien était identique à une propriété naturelle telle que "maximise l'utilité totale", par exemple, alors la phrase "Cette action n'est pas bonne même si elle maximise l'utilité totale" devrait être un énoncé auto-contradictoire et pas seulement faux. En effet, "cette action est bonne" devrait faire partie de la signification de "cette action maximise l'utilité totale". Cependant, cette phrase n'est pas auto-contradictoire. On peut toujours se demander si telle action qui maximise l'utilité totale est vraiment bonne. Alors que si l'on prend l'énoncé analytique "Tous les célibataires sont non mariés", il est absurde de se demander "Est-ce que vraiment ce célibataire n'est pas marié?". Si l'on se pose la question, c'est que l'on n'a pas compris la signification de "célibataire". Dans le cas de "être bon" et "maximiser l'utilité totale", ce n'est pas parce que quelqu'un se demande si une action est bonne alors qu'elle maximise l'utilité qu'il faut en conclure qu'il manifeste une incompréhension de termes ou une confusion conceptuelle. Ces deux propriétés "être bon" et "maximiser l'utilité totale" pourraient être des propriétés liées mais elles n'entretiennent pas, selon Moore, une relation d'identité. Moore en conclut qu'on ne saurait établir de relation d'identité entre un énoncé moral et un énoncé factuel. Pour toute définition x qu'on pourrait donner au terme évaluatif "bon" ou "juste", etc, il sera toujours possible de se demander si x est vraiment bon ou juste. S'il n'est pas possible de définir les termes évaluatifs par des termes naturels, cela signifie que les propriétés évaluatives ne sont pas réductibles à des propriétés naturelles. Moore pense que les propriétés morales sont des propriétés simples, c'est-à-dire non analysables, et non naturelles. La difficulté avec sa théorie éthique consiste à comprendre ce que pourraient être ce type de propriétés non naturelles. Moore défend l'intuitionnisme et considère que les énoncés moraux ont des conditions de vérité mais son argument a aussi été repris par les

non-cognitivistes. Les non-cognitivistes assument que, puisque les propriétés morales ne sont pas définissables en termes de propriétés naturelles, elles ne font pas partie du monde naturel. Comme ils refusent de postuler l'existence d'entités non naturelles, ils en concluent que les valeurs n'existent pas et que, par conséquent, nos énoncés moraux ne réfèrent à rien.

Il existe plusieurs réponses à cet argument de Moore et l'une des voies pour remettre en question son argument consiste à mettre en doute sa notion d'analyse conceptuelle. Michael Smith remarque que l'argument de la question ouverte suppose qu'il est impossible qu'une analyse conceptuelle soit à la fois vraie et informative. Prenons un concept C qui peut être analysé en termes d'un concept C*. Si cette analyse est informative, la question de savoir si C* est C sera une question sensée. Or, selon Moore une analyse ne peut être correcte qu'à la condition qu'il ne fasse pas sens de se demander si C* est C. Ainsi l'argument de Moore implique qu'une analyse n'est correcte que si elle n'est pas informative. Le problème, c'est qu'il y a des analyses informatives comme l'analyse dispositionnelle des couleurs ou l'analyse du savoir en termes de croyances vraies justifiées.

The point that I am making here does not require the assumption that we can correctly analyse the first in terms of the second in each case. The point is rather that the mere fact that it is an open question whether we can [...] is not already enough to show that we cannot. (Smith, 1994, p.36)

Smith pense qu'en acquérant un concept C, nous acquérons un ensemble de dispositions inférentielles et de dispositions concernant le jugement. Ces dispositions lient les faits exprimés en termes du concept C avec d'autres sortes de faits. Un énoncé de toutes ces dispositions constitue un ensemble de platitudes entourant le concept C. Une analyse de ce concept est alors une tentative d'articuler ces platitudes. L'analyse de C en termes de C* est correcte si la connaissance de C* nous fournit une connaissance des platitudes entourant C. Ces dispositions ne sont, en effet, pas forcément transparentes à la personne maîtrisant le concept C. La maîtrise d'un concept nécessite un savoir-comment (*know-how*), alors l'analyse d'un concept requière un savoir-que (*know-that*) concernant notre savoir-comment. Il est donc possible d'avoir la capacité de maîtriser un concept sans avoir la connaissance propositionnelle décrivant cette capacité.

En outre, si Moore suppose qu'il ne peut pas y avoir une identité analytique entre propriétés morales et naturelles, il n'envisage, cependant, pas la possibilité d'une identité

synthétique entre ces propriétés. Une identité synthétique est découverte *a posteriori* de manière empirique. L'identité entre l'eau et H₂O est l'exemple paradigmatique d'une telle identité. Des scientifiques ont découvert à un moment donné la structure moléculaire de l'eau, mais l'eau a toujours été identique à H₂O. Hilary Putnam, dans *Reason, Truth and History* relève que si l'on rejette la possibilité d'identités synthétiques, cela remet en cause des découvertes scientifiques telle celle qui a établi que la température équivaut à l'énergie cinétique moléculaire moyenne.³² En effet, on pourrait utiliser l'argument de Moore pour montrer que la température n'est pas identique à l'énergie cinétique moléculaire moyenne. L'énoncé "x a une température *T*, mais n'a pas l'énergie cinétique moléculaire moyenne *E*", alors que *E* est la valeur de l'énergie cinétique moléculaire moyenne correspondant à *T*, est un énoncé faux, mais il ne manifeste pas une auto-contradiction du la part du locuteur. Il est possible de se demander si la température équivaut vraiment à l'énergie cinétique moléculaire moyenne ou si l'eau est véritablement constitué de H₂O et une personne qui se pose de telles questions ne manifeste pas un échec à comprendre la signification des termes. Or, d'après l'argument de Moore, s'il est possible de se poser ces questions, on ne saurait postuler de relations d'identité entre ces propriétés. Il faudrait alors conclure qu'il n'y a pas d'identité entre l'eau et H₂O, ni entre la température et l'énergie cinétique moléculaire moyenne, mais plutôt que ces propriétés sont corrélées.

Selon Putnam, Moore mélange les notions de propriété et de concept. Il pourrait y avoir des concepts différents correspondant à la même propriété.

There *is* a notion of property in which the fact that two *concepts* are different (say 'temperature' and 'mean molecular kinetic energy') does not at all settle the question whether the corresponding *properties* are different. (And discovering how many fundamental physical properties there are is not discovering something about *concepts*, but something about the *world*.) (Putnam, 1981, p.207)

C'est par une investigation empirique que nous pouvons découvrir les identités synthétiques et non pas par analyse conceptuelle. Le concept de "bon" n'est certainement pas synonyme d'un concept naturel mais il ne faut pas conclure trop hâtivement qu'être bon n'est pas la même propriété qu'une propriété *P* naturelle ou fonctionnaliste.

³²Putnam. 1981. p.207

Pour défendre son argument, Putnam fait appel à l'idée qu'il existe des nécessités métaphysiques qui sont découvertes empiriquement, notion qu'il emprunte à Kripke dans *Naming and Necessity*. Il y a un monde possible où les gens ont des sensations de froid et de chaud et où ces sensations sont expliquées par un mécanisme différent que l'énergie cinétique moléculaire moyenne. Ce monde possible n'est pas un monde où la température n'est pas identique à l'énergie cinétique moléculaire moyenne, mais un monde où un autre mécanisme que la température fait en sorte que certains objets sont ressentis comme froid ou chaud. Puisque, dans le monde actuel, nous avons accepté l'identité synthétique entre température et énergie cinétique moyenne, il ne peut pas y avoir de monde possible où la température n'est pas de l'énergie cinétique moyenne. En acceptant qu'il n'y a pas de monde possible où la température n'est pas identique à l'énergie cinétique moléculaire moyenne, on doit accepter qu'il s'agit d'une vérité nécessaire, même si nous ne pouvons pas la connaître *a priori*. En effet, une proposition qui est vraie dans tous les mondes possibles est une proposition nécessaire. Un énoncé établissant une relation d'identité synthétique est un énoncé empirique mais néanmoins nécessaire: nous avons découvert l'essence de la température par une investigation empirique.

L'argument de Moore cherche à établir que puisque un énoncé comme "Cette action n'est pas bonne même si elle maximise l'utilité totale" n'est pas auto-contradictoire, il ne saurait y avoir de propriété naturelle identique à la propriété de la bonté. Mais la possibilité de vérités nécessaires que l'on découvre de manière empirique bloque l'argument. Tout ce qu'on peut tirer de l'argument de Moore, c'est que le concept de bon n'est pas synonyme de "maximise l'utilité totale". La non-synonymie des termes n'indique pas la non-identité des propriétés. Il est donc possible qu'on découvre un jour, de manière empirique, une identité entre la propriété d'être bon et certaines propriétés naturelles. L'argument de Moore ne suffit donc pas à montrer que tout réalisme naturaliste est faux et ne s'oppose en fait qu'au naturalisme éthique analytique qui cherche à établir des identités analytiques entre propriétés morales et naturelles. Un réaliste naturaliste peut penser que les locuteurs n'ont pas autorité sur la signification des termes moraux et qu'il est donc possible qu'ils ne sachent pas reconnaître des identités synthétiques impliquant des termes moraux. En effet, nous avons vu que les identités synthétiques n'impliquent pas la synonymie.

Moral and nonmoral terms can pick out the same property even if those terms are not synonymous. These will be synthetic property-identity claims, which are to be justified by a dialectical moral inquiry. For

example, hedonistic utilitarianism is not false because 'is one's duty' and 'maximizes pleasure' are not synonymous. This particular property-identity claim is a synthetic claim, to be assessed on substantive moral grounds. (Brink, 2001, p.162)

Cependant, même si l'on doute que l'investigation empirique nous permette un jour d'établir des relations d'identité entre propriétés descriptives et propriétés morales. Mais ce doute ne devrait pas nous mener à rejeter le réalisme. D'une part, cet argument ne menace que le réalisme naturaliste réductionniste et l'on pourrait alors argumenter en faveur d'un réalisme non naturaliste. Mais on peut aussi sauver le réalisme naturaliste en adoptant la théorie de la survenance. Celle-ci stipule que si une propriété morale survient sur des propriétés naturelles, la moindre altération au niveau des propriétés de base aura une incidence sur la propriété survenante. Le réalisme moral naturaliste non réductionniste soutient que les propriétés morales sont constituées par des propriétés naturelles, mais non que les propriétés morales sont identiques à des propriétés naturelles. Nous serions alors en mesure de concevoir une certaine dépendance des valeurs par rapport aux faits sans qu'il s'agisse d'une relation d'identité, ni d'élimination des valeurs au profit des faits.³³ Je vais donc me tourner maintenant vers la théorie de la survenance.

3.2 La survenance

D'après la théorie de la survenance, les valeurs ne sont pas réduites à des faits naturels mais elles dépendent et varient avec ces derniers. La notion de survenance est d'abord apparue chez des auteurs de philosophie morale comme R.M. Hare ou Bernard Williams.

Let us take that the characteristic of "good" which has been called its supervenience. Suppose that we say "St. Francis was a good man." it is *logically impossible* to say this and to maintain at the same time that there might have been another man placed exactly in the same circumstances as St. Francis, and who behaved in exactly the same way but who differed from St. Francis in this respect only, that he was not a good man (Hare, 1952, p.145)

Elle a, néanmoins, été particulièrement féconde en philosophie de l'esprit. En effet, elle permet d'expliquer ce que sont les propriétés mentales, c'est-à-dire des propriétés non physiques, tout en

³³Ogien, 1999, p.141

admettant une description du monde physicaliste. Je commencerai donc par regarder la définition de la survenance donnée par Jaegwon Kim dans *Philosophy of Mind*.

La théorie de la survenance établit que si deux choses ou événements sont exactement pareils sous tous leurs aspects physiques, elles ne peuvent pas différer au niveau mental. Il ne peut donc pas y avoir de différence au niveau mental sans qu'il y ait une différence au niveau physique. Kim énonce trois caractéristiques de la relation de survenance:

(1) If N is a neural state on which mental property M supervenes, then N is a sufficient condition for the occurrence of M (that is, as a matter of law, whenever an organism is in N , it is in M); (2) M can have multiple supervenience bases, N_1, N_2, \dots , each of which is sufficient to give rise to M (this is analogous to the functionalist idea of multiple realizability of the mental); (3) M is distinct from each of its many bases, N_1, N_2, \dots (Kim, 1998, p.149)

Kim voit cependant une difficulté pour la notion de survenance: comment celle-ci aide-t-elle à résoudre le problème de la causalité mentale? La survenance permet-elle une causalité entre esprit et corps tout en évitant une réduction du mental au physique? Si je dis que la douleur à mon épaule cause ma grimace, la théorie de la survenance soutient que la douleur a une base sur laquelle elle survient, c'est-à-dire un état neuronal N . Si N est la base de la douleur, la présence de N suffit pour qu'il y ait douleur. Donc si la douleur est une cause suffisante à ma grimace, il semble que N soit suffisant pour causer ma grimace et donc que N soit la cause de la douleur. Faut-il penser que la grimace ait deux causes distinctes, à savoir la douleur et l'état neuronal N ? Cela ne semble pas crédible à Kim. "For that would mean that the wincing is *causally overdetermined*; in fact it would follow that all cases of mental causation are cases of overdetermination, which is difficult to believe."³⁴ Penser que la douleur est un chaînon intermédiaire entre N et la grimace ne semble pas une solution plus prometteuse. En effet, la douleur et l'occurrence de N sont simultanées. Le problème est donc de faire entrer la douleur dans l'histoire causale. Selon le modèle de la causalité survenante (*supervenient causation*) la douleur est une cause de la grimace puisque la douleur consiste en un état survenant sur l'état

³⁴Kim, 1998, p.150

neuronal N . Le processus causal au niveau neuronal est fondamental et le processus causal au niveau mental survient sur ce niveau fondamental.

the pain supervenes on N and the wincing supervenes on the contraction of a certain group of muscles; the pain's causation of wincing consists in its supervenience base N 's causing the muscle contraction, which is the supervenience base of the wincing. The latter causal relation is fully physical, and we are supposing here that there is no special problem about physical causation. Further, we may speak of the pain as a supervenient cause of the wincing as well, since its supervenience base causes it. In general, then, an instantiation of a property is a supervenient cause of an event in virtue of the fact that its supervenience base causes that event. (Kim, 1998, p.151)

Le modèle de la causalité survenante s'applique à toutes les relations macrocausales. Le fait de chauffer l'eau est la cause de l'eau qui bout. L'augmentation de la température survient sur certains processus moléculaires et ces processus sont la cause de l'éjection de molécules d'eau dans l'air. Le macrophénomène, l'eau qui bout, survient sur ce phénomène moléculaire. Si le mental a le même statut causal que ces phénomènes macrophysiques, il possède la même pertinence au niveau causal.

L'ontologie présente dans les théories de la survenance présente le monde en tant que constitué de couches différentes.

It is usually supposed that there is a bottom level consisting of what according to microphysics are the most basic particles out of which all matter is composed. Above this bottom level there are levels consisting of atoms, of molecules, of cells, of organisms, and so on. Entities belonging to a given level are supposed to be characterized by properties distinctive to that level (although lower-level properties may apply to them also); thus, basic properties like electrical conductivity and temperature are properties of aggregates of molecules. (Kim, 1998, p.221)

Les entités des niveaux supérieurs sont entièrement composés d'entités des niveaux inférieurs. Kim propose une analogie dans le domaine esthétique:

Suppose you are working on an art project, say, a piece of marble sculpture. Your aim is to create a work with certain aesthetic qualities (e.g., beauty, drama, expressiveness), but the work you do on a slab of marble is laborious physical work, cutting, drilling, chipping, and chiseling at a piece of hard material substance. You are giving it a physical shape and texture; that is, you are endowing the piece of marble with a set of

physical properties. When the work is finished, there is no *further* aesthetic work to do, no further step of attaching beauty and other desired aesthetic properties to the material object you have created. Once the physical work is done, the whole project is done. This is so because the physical properties of the object wholly determine its aesthetic properties. And we may justify the attribution of an aesthetic property to it on the basis of the physical properties on which it supervenes. (Kim, 1998, p.222)

Ainsi deux oeuvres qui sont indiscernables du point de vue physique doivent également l'être du point de vue esthétique. La relation de survenance est une relation de dépendance entre niveaux supérieurs et inférieurs sans que l'on puisse réduire l'un à l'autre. "Although aesthetic properties supervene on physical properties in the sense explained, we do not believe that beauty, say, can be given a physical definition or reduction; the aesthetic character of a work of art is distinct from, and irreducible to, its physical properties."³⁵ On pourrait donc penser qu'il en va de même pour les propriétés éthiques. Si Socrate est une bonne personne, quiconque identique à lui du point de vue des propriétés naturelles doit aussi être une bonne personne. Si un certain meurtre est moralement répréhensible, un autre meurtre identique au niveau descriptif sera également moralement répréhensible.

3.2.1 *L'objection de Blackburn*

Néanmoins, selon Blackburn, la théorie de la survenance pose des problèmes au réalisme moral. Je vais examiner ici l'argument offert par Blackburn dans deux essais, "Moral Realism" et "Supervenience Revisited" ainsi que différents commentaires. Blackburn note, dans "Moral Realism", que d'après la théorie de la survenance, les propriétés morales ne peuvent pas varier à moins d'un changement au niveau des propriétés naturelles sur lesquelles elles surviennent. Il s'agit là d'une nécessité conceptuelle.

It seems to be a conceptual matter that moral claims supervene upon natural ones. Anyone failing to realize this, or to obey the constraint, would indeed lack something constitutive of competence in the moral practice. And there is good reason for this: it would betray the whole purpose for which we moralize, which is to choose, commend, rank, approve, or forbid things on the basis of their natural properties. (Blackburn, 1993, p.137)

³⁵Kim, 1998, p.223

Une personne faisant des évaluations morales différentes de deux situations sans mentionner aucune différence naturelle entre ces deux situations ne maîtriserait pas les concepts moraux. Deux situations ayant une même description naturaliste complète doivent recevoir une même évaluation morale. La survenance interdit les mondes mixtes. Il n'y a pas de monde possible où un objet *a* possède un ensemble de propriétés naturelles *N* et aussi une propriété morale *M* et où un objet *b* possède l'ensemble de propriétés *N*, mais pas *M*.

Par contre, les propriétés naturelles d'un objet ou action ne déterminent pas nécessairement les propriétés morales de cet objet ou action. Il est tout à fait possible d'imaginer un monde possible où une autre propriété morale que celle de notre monde actuel survienne sur la même base naturelle. Cela signifie qu'aucun ensemble de vérités non morales ne saurait impliquer la vérité d'une évaluation morale particulière.

(E) There is no moral proposition whose truth is entailed by any proposition ascribing naturalistic properties to its subject. (Blackburn, 1993, p. 116)

Cette condition repose sur la loi de Hume selon laquelle on ne saurait dériver un "devoir être" d'un "être". Puisqu'il y a eu de nombreuses controverses concernant la loi de Hume, il est permis d'affaiblir (E).

(E') There are some moral propositions which are true, but whose truth is not entailed by any naturalistic facts about their subject. (Blackburn, 1993, p. 120)

Le problème, selon Blackburn, c'est qu'il faut expliquer l'interdiction des mondes mixtes. En effet, s'il s'agit d'une vérité conceptuelle qu'une propriété morale résultant d'un ensemble de propriétés naturelles soit toujours une propriété survenante à cet ensemble, dans un monde donné, comment expliquer qu'il soit conceptuellement possible pour cette configuration de propriétés naturelles de ne pas réaliser cette propriété morale en question? Pourquoi serait-il possible qu'il y ait un monde où une propriété morale particulière ne survient pas sur le même ensemble de propriétés naturelles que dans le monde actuel, alors qu'il ne saurait y avoir de monde possible où un objet *a* possède l'ensemble de propriétés naturelles *N* et la propriété morale *M*, et où un objet *b* possède *N*, mais pas *M*? Si aucune description naturaliste n'implique une évaluation morale particulière, il devrait être possible pour le même état de choses non moral d'être parfois

bon et parfois mauvais. Mais c'est justement la possibilité de tels mondes mixtes qu'interdit la survenance.

Imagine a thing *A*, which has a certain set of naturalistic properties and relations. *A* also has a certain degree of moral worth: say, it is very good. This, according to the realist, reports the existence of a state of affairs: *A*'s goodness. Now the existence of this state of affairs is not entailed by *A* being as it is in all naturalistic respects. This means, since all the propositions involved are entirely contingent, that the existence of this state of affairs is not strictly implied by *A* being as it is in all naturalistic respects. That is, it is logically possible that *A* should be as it is in all naturalistic respects, yet this further state of affairs not exist. (Blackburn, 1993, p.118)

Si une telle possibilité existe, il devrait être possible que *A* soit pareil au niveau de ses propriétés naturelles et que l'état de choses, à savoir la bonté de *A*, cesse d'exister. L'existence de faits naturels ne saurait donc garantir, au niveau logique, selon Blackburn, l'état de choses moral. Si le réaliste accepte la thèse selon laquelle des vérités non morales ne peuvent pas impliquer une évaluation morale particulière, alors il lui est difficile d'expliquer l'interdiction des mondes mixtes postulée par la survenance. James Klagge reformule le problème posé au réaliste comme suit:

There are some moral properties with instantiations that are not entailed by the existence of the naturalistic properties with which they are co-instantiated. Yet the persistence of those co-instantiated natural properties, or their duplication in another object, entails the persistence or duplication of the moral property. (Klagge, 1984, pp.372-373)

Pour Blackburn, l'interdit de mondes mixtes va à l'encontre de la notion même de possibilité. Dans "Supervenience Revisited", il imagine le cas d'un monde possible w_1 qui est G^*/F , c'est à dire où F survient sur G^* . Il y a un autre monde w_2 qui est comme w_1 , mis à part qu'il est G^*/O . Blackburn se demande ce qui nous interdit d'avoir un monde possible mixte qui commence par être G^*/F et qui après un moment devient G^*/O . Le problème est que si nous autorisons des mondes possibles mixtes, la thèse de la survenance est falsifiée.

Blackburn ne pense pas que ceci pose problème à sa propre théorie éthique. En effet, selon lui, nous avons des émotions ou attitudes en vertu de propriétés naturelles. Des personnes différentes peuvent avoir des standards d'évaluation différents ce qui implique que les propriétés

naturelles ne contraignent pas nécessairement tout le monde à adopter la même attitude morale envers un même objet ou action. Néanmoins lorsqu'une personne a adopté une certaine attitude morale envers un objet, elle aura la même attitude envers un autre objet possédant des propriétés naturelles similaires pour des raisons de cohérence.

Toutefois, selon Shafer-Landau, les expressivistes ne disposent pas de bonne explication de la survenance. Pour eux la survenance ne peut être qu'une contrainte portant sur des attitudes ou des actions: lorsque nous faisons face à des situations identiques du point de vue descriptif, nous devons exprimer une attitude identique ou prescrire des actions identiques. Il leur faut alors expliquer la provenance de cet impératif. Pourquoi devrions-nous projeter les mêmes attitudes sur des choses que nous croyons identiques du point de vue naturel? Blackburn l'expliquerait en disant que nous devons avoir cette cohérence parce que cela rend la vie meilleure. Cependant Shafer-Landau relève que cette cohérence est trop contingente pour justifier une nécessité conceptuelle. L'autre possibilité, pour l'expressiviste, c'est de dire que par définition le comportement moral requiert de la cohérence au niveau des attitudes. Pour que des actions soient considérées comme des actions morales, il faut agir avec ce genre de cohérence. Cependant, l'expressiviste doit alors expliquer pourquoi la morale requiert cette cohérence au niveau des attitudes. Shafer-Landau renvoie donc à Blackburn la demande d'explication de la relation de survenance: "To the extent that the supervenience of the moral upon the descriptive requires explanation, non-cognitivists have no plausible way to satisfy this requirement."³⁶

Face au problème soulevé par Blackburn, Shafer-Landau est d'avis que nous devrions abandonner la thèse du manque d'implication, à savoir la thèse disant qu'aucun ensemble de vérités naturalistes ne peut impliquer une évaluation morale particulière.³⁷ Il souligne que la survenance implique une relation de dépendance asymétrique et que s'il est vrai qu'une propriété morale n'implique aucune propriété naturelle, on peut penser que des caractéristiques non-morales déterminent les traits moraux d'un état de choses. Shafer-Landau pense que même s'il est conceptuellement possible qu'un même ensemble de propriétés naturelles réalise des propriétés morales différentes, cela n'est pas métaphysiquement possible. Il explique la covariance des

³⁶Shafer-Landau, 2003, p.89

³⁷Shafer-Landau, 2003, p.85

propriétés morales et naturelles en termes de constitution. C'est en vertu de ses constituants physiques qu'une chose possède une caractéristique morale.

These latter exhaustively constitute the former. That is why any two situations that are descriptively identical will be morally identical as well, and why any change of moral status must be explicable by reference to a change in descriptive features. (Shafer-Landau, 2003, p.86)

If constitution relations are to explain a supervenience relation, then the general supervenience claim will say that, necessarily, if any grouping of base properties B^* underlies (because it constitutes) property S , then (in that world) anything else that is also B^* must be S . This claim is a conceptual truth, *no matter what sorts of properties are substitutes for the variables B^* and S* . (Shafer-Landau, 2003, p.88)

Shafer-Landau souligne qu'il y a deux manières de comprendre la survenance, l'une faisant référence à la nature réaliste de la relation de survenance, l'autre non.

(S1): if a concatenation of descriptive properties once underlies a moral one, then it must (in that world) always do so.

(S2): if a concatenation of descriptive properties once underlies a moral one, then it must (in that world) always do so, and do so realistically, i.e. independently of any beliefs about these relations. (Shafer-Landau, 2003, p.87)

(S2) n'est pas une vérité conceptuelle, donc Blackburn ne pourrait pas faire appel à cette conception de la survenance pour attaquer le réalisme moral. En effet, les anti-réalistes comprennent parfaitement le concept de survenance alors même qu'ils nient qu'une telle relation soit à interpréter de manière réaliste. Si (S2) n'est pas une vérité conceptuelle, le fait qu'une thèse du manque d'implication soit vraie au niveau conceptuel ne poserait pas de problème puisque la théorie de la survenance et la thèse du manque d'implication n'apparaîtraient pas au même niveau de modalité. La réponse réaliste au niveau de la nécessité métaphysique suffirait à répondre au problème des mondes mixtes possibles.

Si (S1) n'est pas une vérité conceptuelle, il n'y a pas de problème pour le réaliste puisqu'une fois encore, les deux théories ne figureraient pas à un même niveau de modalité. Par contre, si (S1) est une vérité conceptuelle, le problème des mondes mixtes possibles refait surface. Cependant, Shafer-Landau constate que cela poserait également problème pour des domaines où

nous sommes généralement tous des réalistes. En effet, les relations de survenances entre états mentaux et états physiques, entre faits chimiques et atomiques seraient soumises à la même difficulté. Dans ces domaines, il est aussi peu plausible qu'il existe des vérités conceptuelles exprimant une relation d'implication liant les propriétés de base aux propriétés survenantes. Or si le problème des mondes mixtes possibles mène à une position anti-réaliste en éthique, il devrait faire de même dans les domaines des faits mentaux, chimiques, etc. Toutefois, Blackburn n'est pas un anti-réaliste dans ces domaines là et son argument est sensé porter sur l'éthique uniquement. Il lui faudrait donc expliquer, si l'on suit Shafer-Landau, pourquoi son argument ne nous mène pas à un anti-réalisme global. Toutefois, Terrence Horgan et Mark Timmons soutiennent qu'on ne saurait montrer que la relation de survenance ne devrait pas poser de problèmes en éthique en soulignant qu'elle est acceptée dans s'autres domaines où le réalisme est la position dominante. Je regarderai cet argument plus en détail dans la prochaine section.

Finalement, pour terminer cet examen de l'argument de Blackburn, je me tournerai vers les arguments de James Dreier qu'il a développé dans un article intitulé "The Supervenience Argument Against Moral Realism". Il veut montrer que l'argument de Blackburn n'est pas vraiment un péril pour le réalisme moral.

Il commence par distinguer survenance faible et survenance forte. La survenance faible établit que:

[...] necessarily, for each moral property instantiated by some object *o*, there is a natural property that *o* has, which the moral property always *as a matter of fact* accompanies. Strong supervenience says that necessarily, for each moral property instantiated by some object *o*, there is a natural property that *o* has, which the moral property always *as a matter of necessity* accompanies. (Dreier, 1992, p.14)

Dreier se penche ensuite sur les sortes de nécessités à l'oeuvre dans ces deux formes de survenance. La nécessité portant sur la formule entière, c'est-à-dire le "nécessairement" figurant au début de phrase, est une nécessité analytique. Cela signifie que, quelle que soit le type de survenance liant le moral et le naturel, cette relation existe en vertu des concepts moraux. Personne, capable de maîtriser les concepts éthiques, ne niera la relation de survenance entre ces deux domaines. Cela implique que la thèse de la survenance faible est vraie. Pour savoir si la thèse de la survenance forte est vraie, il faut regarder quelle type de nécessité est impliquée à l'intérieur de la formule. Si on considère qu'il s'agit d'une nécessité métaphysique, il semble bien que la

thèse de la survenance forte soit vraie. Si on juge que Socrate est un homme bon, toute personne possible qui est comme Socrate d'un point de vue naturaliste, sera un homme bon. Par contre, si cette nécessité est analytique, Dreier pense que Blackburn nierait que la thèse de la survenance forte soit vraie. En effet, cela impliquerait qu'il y a une description purement naturaliste de Socrate impliquant analytiquement que Socrate est un homme bon. Or, Blackburn nie qu'une description naturaliste implique une évaluation morale particulière. Aucune description naturaliste de Socrate n'implique analytiquement la phrase "Socrate est un homme bon". Si on résume, Blackburn soutient que la thèse de la survenance faible et la thèse de la survenance forte comprenant une nécessité métaphysique sont vraies. Par contre, la thèse de la survenance forte comprenant une nécessité analytique est fausse. Dreier reformule alors l'argument de Blackburn en distinguant deux prémisses: la première énonce que la thèse de la survenance forte métaphysique est vraie et la seconde énonce que la thèse de la survenance forte analytique est fausse. Or, le réaliste serait incapable d'expliquer la conjonction de ces deux prémisses.

Dreier distingue ensuite différentes formes de réalisme moral et regarde si l'argument de Blackburn fonctionne contre ces différentes variantes. Dans le cas du réalisme réductionniste, l'argument de Blackburn prendrait pour prémisse la négation de cette théorie. Le réaliste réductionniste ne peut pas expliquer pourquoi la thèse de la survenance forte analytique est fausse. Or, souligne Dreier, la thèse de la survenance forte analytique est proche de la thèse centrale du réalisme réductionniste qui affirme qu'il y a des descriptions naturalistes qui, lorsqu'elles sont satisfaites, garantissent analytiquement la possession de certaines propriétés morales. L'argument de Blackburn contre le réalisme réductionniste ne serait pas décisif puisqu'il présuppose la fausseté de la thèse qu'il veut justement réfuter.

Dreier pense que, par contre, le réalisme non réductionniste peut expliquer pourquoi la thèse de la survenance forte métaphysique est vraie et la thèse de la survenance forte analytique fausse. Si les propriétés morales sont réalisées ou constituées par des propriétés naturelles, alors deux objets ne peuvent pas différer dans leurs propriétés naturelles sans différer dans leurs propriétés morales. Mais, d'un autre côté, une évaluation morale particulière ne peut pas être déduite analytiquement d'une description naturaliste. D'une manière similaire, on ne peut pas déduire analytiquement "Louis a bu de l'eau" d'une description faisant référence à de l' H_2O . En effet, l'identité entre eau et H_2O n'est pas établie en vertu de la signification des termes, mais de manière empirique. Or, les naturalistes non réductionnistes soutiennent qu'il n'y a pas d'identité

analytique entre les termes moraux et des descriptions naturalistes, mais plutôt des identités de type synthétique. Cependant, Blackburn n'accepterait sûrement pas ce genre d'explication pour sauver le réalisme remarque Dreier. Blackburn pense que l'explication de pourquoi la thèse de la survenance forte métaphysique est vraie nous pousse à adopter l'anti-réalisme. Blackburn, dans "Supervenience Revisited", soutient que l'anti-réaliste est dans la meilleure position pour expliquer l'interdiction des mondes mixtes possibles. L'anti-réaliste l'explique en termes de nos croyances. Nous croyons que deux choses ne peuvent pas être identiques d'un point de vue physique sans être de la même espèce et nous pensons que toute personne compétente pensera de même. Toutefois, nous supposons également que des personnes compétentes pourraient ne pas être d'accord que si quelque chose est constitué d' H_2O , c'est de l'eau, car reconnaître cette identité suppose un niveau supérieur de connaissance spécialisée. En faisant une conjonction de ces deux croyances,

[...] then we have 'competently possible' worlds of the two kinds and the ban on mixed worlds. But this has now been explained purely by the structure of beliefs that can coexist with competence. There is indeed no further inference to a metaphysical conclusion about the status of wateriness, because the explanation which, in the other cases, that inference helps to provide is here provided without it. To put it another way, we could say that in the moral case as well, when we deal with analytically possible worlds, we are dealing with beliefs we have about competence: in this case the belief that the competent person will not flout supervenience. But this belief is only explained by the further, anti-realist nature of moralizing. If moralizing were depicted further, moral aspects of reality, there would be no explanation of the conceptual constraint, and hence of our belief about the shape of competent morality. (Blackburn, 1993, p.143)

Dreier pense que cette défense de l'anti-réalisme implique qu'aucun réaliste ne pourrait jamais expliquer aucune connexion analytique entre des concepts. Pour illustrer son point, il imagine quelle réponse un réaliste pourrait offrir au défi d'expliquer la connexion nécessaire entre être un célibataire et ne pas être marié. Le réaliste l'explique en disant que la signification de "célibataire" implique logiquement "non marié", étant donné la signification de "non marié". C'est un fait à propos de notre langage et de la manière dont nous employons les mots qui explique cette connexion nécessaire. Cependant, Dreier imagine une objection à la Blackburn soutenant qu'une explication faisant appel à des faits linguistiques, plutôt qu'à des faits métaphysiques se commet à l'anti-réalisme. Une réponse que les réalistes non réductionnistes pourraient offrir à Blackburn consisterait à dire que nommer une propriété naturelle fait partie de la signification du

mot “bon”, de la même manière que nommer une espèce naturelle fait partie de la signification du mot “eau”. Quiconque pensant que deux choses peuvent être identiques dans leurs propriétés physiques et différer quant au fait qu’elles sont de l’eau ne maîtrise pas le concept “eau”. Ainsi quiconque soutenant que deux choses peuvent être identiques dans leurs propriétés naturelles et différer quant à leur bonté morale, ne maîtriserait pas la signification de “bon”. Il s’agit là d’un fait concernant nos pratiques linguistiques. Toutefois, être constitué d’eau et être bon sont des propriétés réelles. Si la thèse de la survenance forte métaphysique est elle-même une vérité conceptuelle, elle doit bien sûr être expliquée en faisant référence à la manière dont nous parlons et pensons la moralité, plutôt qu’en faisant référence à des faits naturels ou moraux. Cela n’implique cependant pas qu’il faille adopter l’anti-réalisme.

Nous voyons ainsi que le réaliste peut répondre à l’argument de Blackburn. Il peut souligner que si l’argument de Blackburn conduit à adopter l’anti-réalisme en éthique, il devrait en être de même pour d’autres entités comme les propriétés mentales ou macrophysiques. Dreier a aussi souligné que l’argument de Blackburn présuppose la fausseté du réalisme naturaliste réductionniste et n’est donc pas un argument décisif contre ce dernier. Il a également montré comment un réaliste naturaliste non réductionniste peut répliquer à l’argument de Blackburn.

3.2.2 *L’explication de la relation de survenance*

J.L. Mackie a soulevé une autre difficulté découlant de la théorie de la survenance pour le réalisme. Il pense que le fait d’invoquer la survenance pour expliquer le lien entre propriétés morales et naturelles revient à postuler un lien de conséquence mystérieux entre les deux.

Another way of bringing out this queerness is to ask, about anything that is supposed to have some objective moral quality, how this is linked with its natural features. What is the connection between the natural and fact that an action is a piece of deliberate cruelty- say causing pain just for fun- and the moral fact that it is wrong? It cannot be an entailment, a logical or semantic necessity. Yet it is not merely that the two features occur together. The wrongness must somehow be “consequential” or “supervenient” [...] It is not sufficient to postulate a faculty that sees wrongness: something must be postulated which can see at once the natural features that constitute cruelty, and the wrongness, and the mysterious consequential link between the two. (Mackie, 1977, p.41)

Selon Mackie, donc, le réaliste doit être en mesure d'expliquer la relation établissant que les propriétés morales surviennent sur les propriétés naturelles. Il pense, cependant, que cette relation n'est pas explicable en termes naturalistes. L'objection de Mackie a été reprise par Terrence Horgan et Mark Timmons qui défendent que le réalisme moral est incapable de fournir une explication naturaliste acceptable de la relation de survenance entre propriétés morales et propriétés naturelles. Ils n'acceptent pas que la relation de survenance soit une relation brute et inexplicable. En l'absence d'explication valide, ils concluent, comme Mackie, que les propriétés morales sont bizarres d'un point de vue métaphysique.

Les réalistes peuvent répondre que nous acceptons que les propriétés mentales surviennent sur les propriétés physiques ou que les faits macrophysiques surviennent sur des faits microphysiques, sans remettre en doute l'existence des propriétés mentales, ni des faits macrophysiques. Si nous acceptons la survenance entre le moral et le physique, nous ne disposerions pas de raisons pour mettre en question l'existence des faits et propriétés morales. Néanmoins, Horgan et Timmons n'acceptent pas cette réponse du réalisme moral naturaliste consistant à montrer que la survenance en morale ne devrait pas poser problème puisque nous l'acceptons dans d'autres domaines. Ils pensent que la survenance entre propriétés morales et propriétés naturelles est plus difficile à expliquer que celle liant propriétés macrophysiques et microphysiques ou entre propriétés mentales et physiques. En effet, ils soutiennent que l'explication de la survenance à l'oeuvre entre propriétés mentales et propriétés physiques n'est pas disponible pour un réaliste moral naturaliste. La cible de Horgan et Timmons n'est pas le naturalisme réductionniste qui a déjà été mis à mal par l'argument de la question ouverte de Moore. Ils s'opposent à ce qu'ils appellent le "réalisme moral nouvelle vague" qui est une version du réalisme moral basé sur le naturalisme sémantique causal. David Brink et Nicholas Strugeon sont des représentants de ce courant aussi appelé le réalisme de Cornell. Horgan et Timmons considèrent que le réalisme moral et le naturalisme faisant appel aux théories de la référence directe, dont la théorie causale, sont incompatibles.

L'explication de la relation de survenance entre propriétés mentales et physiques adoptée par Horgan et Timmons est une explication sémantique. Ils établissent un parallèle entre nos concepts d'espèces naturelles et nos concepts mentaux. Même si nous ne connaissons pas la nature essentielle d'une espèce naturelle particulière, nous pensons qu'elle a une constitution microphysique. Nous savons conceptuellement que quelque chose possédant cette constitution est

une instance de cette espèce naturelle. Horgan et Timmons cherchent à appliquer ce modèle à la survenance entre mental et physique.

Le fonctionnalisme en philosophie de l'esprit peut prendre deux formes, le fonctionnalisme analytique ou le psychofonctionnalisme. Le fonctionnalisme affirme en gros que l'essence des propriétés mentales est révélée par le rôle causal joué par ces propriétés. Le fonctionnalisme analytique lie le rôle fonctionnel d'un état mental particulier à la signification du terme. Les généralisations de la psychologie du sens commun opèrent comme des contraintes sémantiques. Le psychofonctionnalisme, quant à lui, affirme que les termes d'attitudes propositionnelles réfèrent de manière rigide à des propriétés fonctionnelles dont l'essence est saisie, non pas par la psychologie du sens commun, mais par une théorie psychologique empirique plus élaborée. Nous ne savons pas encore de quoi aura l'air cette théorie, mais nos termes et concepts impliquant des attitudes propositionnelles sont causalement régulés par des propriétés fonctionnelles. C'est par investigation empirique que nous découvrirons ce que sont ces propriétés fonctionnelles. Nous ne pouvons pas les connaître en nous basant uniquement sur le sens des termes mentaux considérés. Horgan et Timmons soutiennent que nous n'avons pas besoin de savoir quel rôle spécifique joue une propriété mentale, mais seulement que les propriétés mentales jouent un rôle causal essentiel. Cela fait partie de nos concepts mentaux. Étant donné que ce sont les propriétés physiques qui déterminent quel rôle causal est instantié, ce sont les propriétés physiques qui déterminent quelle propriété mentale est instantiée. Admettons que P soit une propriété physique et M une propriété mentale:

P is a member of a system of physicochemical properties which together more or less realize the pattern of causal relations characterized by the psychological generalizations which semantically constrain mental terms and concepts. Furthermore, P itself occupies the role in that system which, according to those psychological generalizations, constitutes the M-role. Hence, whenever someone instantiates P, he must instantiate M as well. (Horgan & Timmons, 1992, p.238)

Nous avons là, pensent Horgan et Timmons, une explication de la survenance entre mental et physique qui ne peut pas être transposée dans le domaine moral. Il n'y aurait pas, selon eux, de contraintes sémantiques régulant les termes et concepts moraux qui pourraient être combinées à des faits physiques afin de déterminer une relation de survenance morale objective.

David Brink, qui est un représentant du réalisme moral visé par Horgan et Timmons, pense que les propriétés morales sont des propriétés fonctionnelles caractérisées par leur rôle causal dans les activités humaines. Brink écrit que:

The realist might claim that moral properties are those which bear upon the maintenance and flourishing of human organisms. Maintenance and flourishing presumably consist in necessary conditions for survival, other needs associated with basic well-being, wants of various sorts, and distinctively human capacities. [...] The physical states which contribute to or interfere with the satisfaction of these needs, wants, and capacities are the physical states upon which, on this functionalist theory, moral properties ultimately supervene. (Brink, 1984, pp.121-22, cité par Timmons, 1999, p.54)

Les propriétés morales réfèrent ainsi à des propriétés fonctionnelles qui surviennent sur des propriétés physiques. Selon Brink, nous pouvons découvrir l'essence de ces propriétés fonctionnelles par le biais de l'investigation morale. Ce sont les différentes théories morales qui expliquent la manière dont les propriétés morales surviennent sur les propriétés naturelles. Les termes moraux saisissent certaines propriétés fonctionnelles dont l'essence est révélée par certaines théories morales normatives. La théorie morale capable d'expliquer au mieux la survenance morale sera celle capable de fournir un compte-rendu de nos croyances morales et non-morales.

Si on accepte l'idée qu'il existe des vérités empiriques nécessaires, la relation de survenance n'est plus une relation inexplicable au niveau métaphysique pour le naturaliste et il n'est plus besoin de postuler un lien de conséquence entre les propriétés naturelles et les propriétés morales. Il n'y aurait rien d'étrange dans cette relation de survenance. Brink relève que nous ne pensons pas que les propriétés mentales sont étranges alors qu'elles surviennent sur le physique, pas plus que nous questionnons l'existence des faits sociaux, biologiques ou macroscopiques et il ne voit pas pourquoi on ne pourrait pas faire appel à la survenance des propriétés morales sur des propriétés physiques pour expliquer l'existence de propriétés morales.

Horgan et Timmons notent que Brink se base sur le psychofonctionnalisme en philosophie de l'esprit pour développer son fonctionnalisme moral. Ainsi, d'après Brink, nous découvrirons de manière empirique les généralisations captant l'essence fonctionnelle des propriétés morales. Ces généralisations appartiendront à la théorie morale qui rendra compte au mieux de nos croyances morales et non morales. Les termes moraux réfèrent des propriétés fonctionnelles dont

l'essence sera révélée par les généralisations d'une théorie morale synthétique. Ce que Horgan et Timmons nient, c'est que les propriétés morales soient identiques à des propriétés fonctionnelles.

Horgan et Timmons imaginent une expérience de pensée pour montrer que les termes moraux ne fonctionnent pas comme des termes d'espèces naturelles. Ils reprennent, en la modifiant, l'expérience de pensée de la Terre Jumelle imaginée par Hilary Putnam. L'expérience de pensée de Putnam met en scène une Terre Jumelle identique à notre Terre, mis à part le fait que l'eau n'y est pas constitué de H_2O , mais de XYZ. La question que Putnam se pose est de savoir si le terme "eau" veut dire la même chose sur Terre et sur la Terre Jumelle. Putnam conclut que "eau" doit désigner de l' H_2O dans tous les mondes possibles. Il ne pense pas que "eau" veuille dire la même chose sur Terre et sur Terre Jumelle. Les termes "eau" ont une signification différente parce que la signification de ces termes n'est pas déterminée par le concept ou les descriptions associées aux termes respectifs, mais plutôt par la nature du monde dans lequel les Terriens et les habitants de la Terre Jumelle se trouvent. Pour Putnam, peut être appelé "eau" tout ce qui possède une relation d'équivalence avec une instance de liquide référée en tant qu'eau dans le monde actuel. Le terme "eau" fonctionne donc comme un désignateur rigide³⁸, c'est-à-dire qu'il désigne la même propriété dans tous les mondes possibles. Il n'est, ainsi, pas logiquement possible que l'eau soit autre chose que de l' H_2O . Le terme "eau" désigne la propriété d'être de l' H_2O , non pas en vertu du concept que le terme exprime, mais grâce à sa connexion causale avec cette propriété. Comme l'identité entre l'eau et H_2O est une nécessité synthétique qui a été découverte de manière empirique, un locuteur n'a pas besoin de savoir que l'eau, c'est de l' H_2O , pour être un locuteur compétent. En effet, avant la découverte de la constitution chimique de l'eau, les gens étaient des locuteurs compétents du terme "eau".

Horgan et Timmons imaginent une Terre Jumelle Morale. Ils postulent que les jugements moraux des Terriens sont causalement régulés par des propriétés fonctionnelles dont on peut caractériser l'essence en termes des généralisations d'une théorie morale normative unique. Ils suppose que cette théorie est la théorie conséquentialiste. Sur, la Terre Jumelle Morale, par contre, la théorie dont les généralisations fournissent l'essence des propriétés fonctionnelles régulant les

³⁸Notion que Putnam emprunte à Saul Kripke. Pour ce dernier, les noms propres sont des désignateurs rigides. Par exemple, la référence de "Aristote" n'est pas déterminée par une description généralement associée à ce nom. Nous pourrions, par exemple, découvrir qu'Aristote n'a pas été l'élève de Platon, mais Aristote n'en demeurerait pas moins la même personne. La référence du nom "Aristote" a été fixée lorsqu'Aristote a été baptisé ainsi.

jugements moraux est la théorie déontologiste. Les termes moraux sont donc causalement régulés par des propriétés naturelles différentes sur Terre et sur la Terre Jumelle Morale. La question qu'on doit alors se poser est de savoir si les Terriens et les habitants de la Terre Jumelle emploient le même terme lorsqu'ils disent que quelque chose est bon ou juste, etc. Doit-on conclure, comme dans l'expérience de pensée de Putnam, qu'ils emploient des termes différents et que "bon" sur Terre ne veut pas dire la même chose que "bon" sur la Terre Jumelle? Horgan et Timmons pensent que les Terriens et les habitants de la Terre Jumelle n'emploient pas des termes ayant une signification différente, mais plutôt qu'ils ont des croyances morales différentes et qu'ils diffèrent dans leurs théories morales. En effet, nous ne sommes pas prêts d'admettre que les conséquentialistes et les déontologistes ne disent pas la même chose en employant les termes moraux. Nous pensons plutôt qu'ils s'accordent sur les termes mais pas sur leur application. La question de savoir quelles sont les propriétés naturelles régulant les termes moraux est, selon Horgan et Timmons, une question ouverte et aucune investigation empirique ne permettra de le déterminer. Ils soutiennent donc qu'il n'y a pas de contrainte sémantique dans l'emploi des termes moraux selon laquelle les propriétés morales sont réalisées par des propriétés physiques. L'idée que le moral survient sur le physique ne serait pas construite dans nos termes et concepts moraux. Comme, pour expliquer la relation de survenance entre le mental et le physique, le fonctionnalisme faisait appel à une contrainte sémantique additionnée à des faits physiques, le réaliste moral ne peut pas faire valoir un argument similaire.

Nick Zagwill propose une objection à l'argument de Horgan et Timmons en niant que le fonctionnalisme en philosophie de l'esprit explique la survenance. En effet, pour que leur argument fonctionne, il faut accepter au préalable que l'instantiation d'un certain rôle causal déterminant une propriété mentale soit une contrainte sémantique. En supposant que ce sont des états physiques qui instancient ce rôle causal, nous devons parvenir à la conclusion que le mental est déterminé par le physique. Zagwill note que cet argument ne peut établir la survenance qu'une fois la première prémisse établie. Or le fonctionnalisme, qu'il soit analytique ou psychologique, est la thèse selon laquelle les propriétés mentales sont déterminées par leur rôle causal. Or ceci ne constitue pas une explication du fait qu'elles le soient. Zagwill pense qu'on ne peut pas établir de faits modaux, tel que la survenance, en faisant appel à des faits sémantiques. Même si nos concepts mentaux impliquaient que certaines conditions physiques réalisent les propriétés

mentales, il faudrait quand même pouvoir justifier le fait de posséder ces concepts. Zagwill est d'avis que les faits modaux sont une chose et les concepts une autre.

The fact that our concepts involve us in a *commitment* to certain determination relations between properties does not tell us that there really are such determination relations. How could an appeal to concepts explain anything about those modal facts? That seems a category mistake. (Zagwill, 1997, p.513)

Par ailleurs, Zagwill note que les réalistes moraux naturalistes, comme Nicholas Sturgeon et David Brink, qui ont établi l'analogie entre termes moraux et termes d'espèces naturelles emploient cette analogie concernant la relation de survenance. Ils ne défendent pas explicitement une théorie causale de la fixation du contenu des termes moraux.

Les réalistes peuvent, en outre, remettre en questions les intuitions sémantiques illustrées par l'expérience de la Terre Jumelle Morale. Eric Russert Kraemer remarque que les défenseurs du naturalisme sémantique causal peuvent invoquer le fait que les intuitions sémantiques prennent du temps à se mettre en place et que l'idée du naturalisme sémantique causale est nouvelle. Ils peuvent aussi arguer que nous n'en savons pas encore assez concernant quelles propriétés naturelles réalisent les propriétés mentales. Peut-être nos intuitions sémantiques changeraient-elles si nous avions plus de détails concernant ces propriétés naturelles. Kraemer propose une analogie pour mieux comprendre cet argument. Il se demande ce qu'il adviendrait de nos intuitions sémantiques si nous découvrions des formes de vie avancées sur une autre planète ne possédant pas d'ADN mais une autre combinaison chimique. Étant donné que nous identifions les gènes à des segments d'ADN, dirions-nous que ces créatures n'ont pas de gènes puisqu'elles n'ont pas d'ADN ou penserions-nous que nous avons identifier faussement les gènes avec des séquences d'ADN? Kraemer n'est pas certain de quelle serait la réponse. Ce qu'il veut montrer, c'est qu'il peut être difficile de prédire précisément comment nos intuitions sémantiques changeront face à des changements majeures dans notre corpus épistémique commun et dans notre compréhension de ce corpus. Il est certain que l'acceptation du naturalisme sémantique causal impliquerait des modifications de notre compréhension actuelle du savoir scientifique.

Une autre stratégie que les réalistes peuvent employer pour contrer l'argument de Horgan et Timmons consiste à dire que les intuitions à l'oeuvre dans l'expérience de la Terre Jumelle Morale appartiennent à la morale de sens commun qui est un compte-rendu désuet de la nature

de la moralité. Accepter le naturalisme sémantique causal impliquerait de mettre de côté certaines intuitions de l'homme de la rue et de se concentrer sur un programme théorique plus large.

Geoffrey Sayre-McCord propose, lui aussi, dans son article “‘Good’ on Twin Earth”, une manière de réconcilier le réalisme moral avec l'expérience de pensée de Horgan et Timmons. Il souligne le fait que nous ne nous tournons pas vers la science pour établir si la propriété régulant l'usage fait par une autre communauté d'un terme moral est la même propriété qui régule notre emploi de ce terme. Il introduit la notion d'espèce moral (*moral kind*) qu'il distingue de la notion d'espèce naturelle. Les espèces morales ne figurent pas dans les sciences naturelles, mais dans les théories morales.

Sayre-McCord pense qu'on peut trouver dans notre monde quelque chose de similaire à l'expérience de pensée de Horgan et Timmons. Il y a, en effet, des personnes qui ont des critères différents quant à l'application des termes éthiques. Par exemple, un partisan de la théorie de la justice rawlsienne ne considérera comme juste que les institutions conformes aux deux principes de la justice identifiés par Rawls, alors qu'un utilitariste pensera qu'une institution est juste si elle incarne des règles dont l'adoption promet le bonheur du plus grand nombre. Des égalitaristes seront d'avis qu'une institution juste favorise une distribution égalitaire des biens. Toutes ces personnes proposent des critères différents à la justice et leur terme “justice” est sensible à des traits naturels différents. Malgré cela, nous pensons que ces personnes peuvent dialoguer entre elles, donc qu'elles emploient des termes ayant une même signification, et qu'elles ont simplement des vues divergentes concernant la justice. Ces cas nous suggèrent, selon Sayre-McCord, que nous ne devrions pas considérer que le sens du terme “justice” est déterminé par les critères utilisés par les personnes employant ce terme. Horgan et Timmons en concluent que nous ne devrions pas penser que le terme “justice” réfère aux propriétés naturelles régulant causalement l'emploi du terme fait par une personne ou une communauté. Sayre-McCord est d'avis, néanmoins, que rien ne nous empêche de concevoir que le terme “justice” réfère à une espèce morale.

In fact, a crucial assumption behind our willingness to see Rawlsians, utilitarians, and egalitarians as all talking about justice is our sense that the various criteria they each rely on have emerged in response to (as we see it) samples of a single moral kind- justice. Often enough the criteria are, quite clearly, embraced on the grounds that they articulate the true nature of what they all agree we've been talking all along. When one or another criterion is embraced on these grounds, this will reflect the conviction that the best normative

theory will countenance seeing the criterion as specifying the true nature of justice (in a way that parallels our conviction that the best scientific theory will countenance seeing H₂O as specifying the true nature of water). (Sayre-McCord, 1997, p.288)

Afin de montrer la vraisemblance de cette conception, Sayre-McCord nous propose d'imaginer une communauté dans laquelle le terme "justice" ne serait pas régulé de manière appropriée par la justice. L'emploi de ce terme serait réduit aux choses respectant des règles sociales injustes. Le fait que leur emploi du terme soit insensible à la justice nous pousserait à considérer que les membres de cette communauté ne parle pas de justice, mais de ce qui est approuvé socialement. Ce n'est pas parce que nous pensons que notre emploi des termes moraux réfère à ce qui est valable, juste, etc que nous pensons que nos vues actuelles sur le sujet sont décisives. Sayre-McCord souligne que nous sommes prêts à remettre en cause nos vues, sans penser que nous ne parlons plus de justice. Nous sommes également capables de reconnaître que des des points de vue très éloignés sur la question portent néanmoins sur la justice.

As the moral kinds approach would have it, what a moral term refers to, if anything, is determined by whether, in light of the best moral theory, the use of that term can be seen as appropriately regulated by instances of a normatively significant kind. Our sincere deployment of the terms reflects, in turn, our conviction that we are using them to refer to what the best theory would reveal to be normatively significant kinds. But that conviction travels with a recognition that we ourselves may be mistaken in quite significant ways about the shape the best theory might take. (Sayre-McCord, 1997, p.291)

Il n'y a évidemment pas de consensus concernant la justice, le bien, etc et il est donc difficile de savoir quelle théorie normative saisit le mieux la nature de la justice, par exemple. Sayre-McCord demeure malgré tout confiant et pense que l'élaboration de théories normatives est une tentative pour clarifier ce qui compte comme normativement signifiant. Par ailleurs, il ne pense pas que l'absence de consensus soit une objection à la proposition sémantique selon laquelle nos termes moraux réfèrent aux espèces normativement signifiantes, quelles qu'elles soient, régulant de manière appropriée leur emploi. Un avantage de l'approche des espèces morales, souligne Sayre-McCord, c'est qu'elle permet de considérer que deux points de vues qui s'opposent sont en désaccord. Ces points de vues sont en désaccord à propos de quelque chose et ce quelque chose peut être saisi (même si cela n'est pas le cas actuellement) par une théorie normative établissant la nature de l'objet discuté.

L'approche en termes d'espèces morales permet de dire que les Terriens et les habitants de la Terre Jumelle emploient des termes ayant un même sens. L'emploi des termes moraux faits par les déontologistes et les conséquentialistes réfère à une espèce normativement signifiante. Il est peut-être difficile de déterminer laquelle des deux théories saisit au mieux la nature de ces termes, mais il y a néanmoins quelque chose auquel réfère les termes moraux.

David Brink, dans un article intitulé "Realism, Naturalism, and Moral Semantics", a également proposé une solution pour sauver le réalisme moral des conclusions de l'argument de Horgan et Timmons. Il suggère d'abandonner la théorie causale de la référence pour un compte-rendu de la sémantique morale en termes d'intentions référentielles. En effet, l'argument de Horgan et Timmons vise un réalisme moral basé sur le naturalisme sémantique causal.

Brink observe que ceux qui ont introduit le terme "eau" voulaient référer à la structure, quelle qu'elle soit, expliquant les traits perceptibles et fonctionnels de cette substance sans couleur et sans odeur, se trouvant dans les lacs et les rivières, servant à se désaltérer, se laver, etc. Il se trouve que c'est la structure chimique H_2O qui permet d'expliquer ces traits. Brink cherche alors quelles sont les intentions référentielles dans le cas moral. Cela implique de trouver une formulation du point de vue moral qui soit suffisamment abstraite pour permettre une grande variété de positions morales. Il faut distinguer, souligne Brink, concepts et conceptions. Par exemple, il faut distinguer le concept de justice distributive, d'une part, et l'utilitarisme, le libéralisme et l'égalitarisme, de l'autre. Brink pense que ce qui est distinctif du point de vue moral c'est que nous évaluons des personnes, des actions et des institutions en fonction de standards que d'autres peuvent et devraient accepter. Cette manière de comprendre le concept de moralité permet d'admettre des conceptions très diverses.

Even when appraisers from distinct communities use language in ways that are counterfactually regulated by different properties of their environments, we should interpret their language as moral language and their judgments that employ that language as moral judgments only if those judgments are based on standards that the appraisers endorse, if only implicitly, as interpersonally justifiable. [...] But our account of the shared referential intention to pick out people, actions and institutions that are interpersonally justifiable, in virtue of which the judgments of Earthlings and Moral Twin Earthlings are both moral judgments, identifies just such a common meaning or reference about the two communities have a disagreement in belief. Their disagreement is one about which features of people, actions, and institutions make them interpersonally justifiable, with Earthlings holding consequentialist views and Moral Twin Earthlings holding deontological views (Brink, 2001, pp.174-175)

Brink rend donc le réalisme moral et la théorie de la référence directe compatibles en abandonnant la thèse de la régulation causale. Il soutient que la référence est fixée par les intentions originelles d'adopter un point de vue moral, c'est-à-dire d'employer du langage moral pour saisir les propriétés qui sont l'objets d'une évaluation justifiable de manière inter-personnelle, et non par les propriétés régulant l'usage actuel du langage moral. Cela lui permet d'échapper à la conclusion que les termes moraux employés par des conséquentialistes et des déontologistes réfèrent à des propriétés différentes.

J'ai esquissé ici plusieurs réponses possibles à l'argument de Horgan et Timmons. Je ne m'étendrai pas plus longtemps sur cette question et ne me prononcerai pas quant à savoir laquelle est la meilleure. Ce qui est important, c'est le réalisme dispose de ressources pour répondre à cette objection.

3.3 Les explications morales

Le réalisme moral non-naturaliste et le réalisme naturaliste non réductionniste pensent tous deux que les propriétés morales ne sont pas réductibles aux propriétés naturelles mais que les premières sont constituées ou réalisées par ces dernières. Ces deux positions doivent alors nous dire comment des faits moraux peuvent expliquer des observations. En effet, il ne peuvent pas recourir à l'efficacité causale des propriétés morales en argumentant que celles-ci sont en fin de compte réductibles à des propriétés naturelles dont le pouvoir causal ne pose pas problème. Toutefois, il est largement admis que quelque chose est crédible au niveau ontologique si on a besoin de lui pour expliquer ce qui est le cas dans le monde.

À première vue, il semble que les faits moraux sont souvent invoqués dans des explications causales des événements du monde. Par exemple, on explique pourquoi un policier a accepté un pot-de-vin en invoquant sa corruption morale. Cependant, les choses ne sont pas aussi simples:

passing the counterfactual test of causation relied upon by some realists is insufficient to certify the realistic nature of the facts in question. The test claims that C is a cause of O if O would not have occurred as it did had C not occurred as it did. Yet non-realistic facts pass the test. For instance, if it weren't against the law to steal, then the burglar wouldn't have been prosecuted and jailed. The legal standards that make this claim

true are not realistic- they are not true independently of the attitudes anyone takes towards the propositions they express. (Schafer-Landau, 2003, p.100)

Les standards légaux, comme les règles d'un jeu, sont au mieux compris comme étant construits à partir des pratiques et attitudes des gens. La réalité sociale est construite ce qui signifie qu'il ne s'agit pas de faits bruts.

Par ailleurs, le critère causal est trop restrictif. Geoffrey Sayre-McCord note qu'outre les faits moraux, les énoncés mathématiques, les généralisations empiriques, les lois scientifiques échouent aussi au test du critère causal.³⁹ En effet, si les généralisations et les lois aident à expliquer pourquoi nous faisons les expériences que nous faisons, elles n'en sont pas la cause. Le fait que toutes les émeraudes sont vertes ne cause pas une émeraude particulière à être verte.

Le problème n'est alors pas tant de savoir si une certaine catégorie de faits peut figurer dans une explication causale, mais plutôt de savoir si elle est requise pour expliquer une observation. Or, selon l'argument classique contre l'objectivisme en éthique, nous n'avons pas besoin de faire appel aux faits moraux pour disposer d'explications complètes. Les explications en termes de croyances et de désirs sont suffisantes. Nous n'avons pas besoin de recourir à la malignité d'une action pour expliquer une réaction horrifiée face à celle-ci. Aussi ferions-nous mieux de comprendre les propriétés morales comme découlant de nos attitudes et nos accords.

Gilbert Harman soutient que les énoncés moraux ne peuvent expliquer un état de choses qu'à la condition d'être réductibles à des énoncés naturels, plus précisément à des énoncés psychologiques ou des sciences sociales. Il pense que les principes moraux ne peuvent pas être confirmés et testés de la même manière que les énoncés scientifiques. Nous testons généralement nos principes moraux en faisant appel à des expériences de pensée. Nous construisons des situations imaginaires pour voir si le principe moral examiné serait toujours applicable dans cette situation. Harman prend comme exemple le principe selon lequel il vaut mieux sacrifier une vie pour en sauver cinq plutôt que l'inverse. Ce principe semble tout à fait adéquat pour un médecin recevant six blessés, mais disposant de moyens limités. S'il doit choisir entre employer tous ses moyens pour sauver la personne la plus grièvement blessée ou sauver les cinq autres personnes et laisser mourir la sixième, il agira de manière juste en sauvant le maximum de personnes. Par contre, en imaginant qu'il y a cinq patients mourant dans un hôpital, chacune ayant besoin d'une

³⁹Sayre-McCord, 1988, p.266

transplantation d'organe, il nous paraîtrait horrible de sacrifier une personne bien portante pour distribuer ses organes aux autres personnes. En faisant cette expérience de pensée, nous voyons que le principe moral examiné ne recommande pas toujours la meilleure action. Toutefois, cela n'est pas la même chose que de confronter ce principe au monde.

This of course was a "thought experiment". We did not really compare a hypothesis with the world. We compare an explicit principle with our feelings about certain imagined examples. [...] But scientific hypotheses can also be tested in real experiments, out in the world. (Harman, 1988, p.120)

La question qu'il faut alors se poser, c'est de savoir si nous pouvons percevoir la bonté ou la malignité d'une action. Harman admet que lorsque nous sommes témoins d'une scène horrible, par exemple, nous n'avons pas besoin de conclure que les auteurs du crime font quelque chose de mal. Nous voyons que cela est mal. Cependant, notre réaction est-elle due à la malignité de l'action ou est-elle une réflexion de notre sens moral?

Harman admet qu'il n'y a pas d'observation pure. En science, comme en éthique, ce que nous percevons dépend en partie des théories auxquelles nous adhérons. L'observation implique la formation de croyances. Pour former une croyance nous avons besoin de comprendre les concepts pertinents et un concept a toujours un rôle dans une théorie ou un système de croyances. Dans l'exemple de Harman, nous voyons des enfants verser de l'essence et mettre le feu à un chat. Pour voir cela, il faut posséder au préalable un certain nombre de connaissances, comme par exemple des connaissances sur le développement humain, une connaissance de ce qu'est l'essence, etc. Lorsqu'une observation entre en conflit avec la théorie, en science, comme en éthique, nous avons le choix entre dire que nous avons commis une erreur d'observation ou remettre en question la théorie.

Harman pense que la différence entre la science et l'éthique, c'est que le scientifique doit supposer l'existence de faits physiques pour expliquer les occurrences de l'observation soutenant sa théorie. L'observation d'une trace de vapeur ne soutient sa théorie que parce qu'il est raisonnable de penser qu'il y a quelque chose dans le monde, un proton, qui explique son observation, allant au-delà de suppositions concernant la psychologie du scientifique. Par contre, Harman ne pense pas que nous aillions à supposer l'existence de faits moraux pour expliquer les cas d'observations morales. Tout ce qui est nécessaire dans le cas moral, ce sont des suppositions concernant la psychologie et la sensibilité morale de la personne faisant l'observation morale.

Indeed, an assumption about moral facts would seem to be totally irrelevant to the explanation of your making the judgment you make. It would seem that all we need assume is that you have certain more or less well articulated moral principles that are reflected in the judgments you make, based on your moral sensibility. It seems to be completely irrelevant to our explanation whether your intuitive immediate judgment is true or false. (Harman, 1988, p.122)

Cependant, selon Russ Schafer-Landau, quel que soit le contenu d'un jugement, il peut toujours être expliqué causalement par référence à un réseau de croyances plutôt qu'en faisant appel à l'effet produit par le monde extérieur sur le sujet. Ainsi on pourrait expliquer le jugement émis par un scientifique qu'il y a un proton qui a passé en faisant référence exclusivement aux croyances d'arrière plan du scientifique. En effet, sans ces croyances, ce jugement n'aurait pu être rendu. Pour Schafer-Landau, ce sont ces croyances additionnées d'une observation empirique qui suffisent à expliquer le jugement. Il pense qu'il en va de même pour l'éthique. Sans croyances morales d'arrière-plan, nous ne saurions émettre de jugements moraux. Ce sont ces croyances et une observation pertinente qui suffisent à expliquer un jugement moral.

Selon David Brink, des explications purement psychologiques sont inadéquates dans un cas comme dans l'autre. Nous aurions besoin de disposer d'une explication de pourquoi l'observateur, dont on invoque la psychologie, a les croyances qu'il a. Dans le cas du scientifique, nous voulons savoir pourquoi il croit qu'il y a une trace de valeur et pourquoi il soutient sa théorie microphysique. Dans le cas des enfants mettant le feu à un chat, nous aurions besoin d'une explication de pourquoi l'observateur croit que la cruauté délibérée et inutile est mauvaise. Pour expliquer ceci, Brink pense que nous sommes forcés de postuler des faits moraux. Il n'est pas suffisant d'invoquer sa psychologie et son éducation.

But we need to explain these influences, and unless we regard this education as mistaken or in some way biased, its explanation is likely to refer, eventually and in part, to moral facts, their recognition and transmission. (Brink, 1989, p.186)

Brink construit également un exemple où la croyance de l'observateur en certains faits moraux ne peut pas expliquer les événements observés, puisque l'observateur ne reconnaît pas lui-même ces faits. Il imagine le cas d'une personne qui considère sa société comme étant globalement juste

alors que ses lois et institutions sont en fait injustes. Cette personne fait partie de la classe défavorisée et, malgré le fait qu'elle continue à croire sa société juste, elle commence à ressentir du ressentiment. En se liant à d'autres membres de sa classe sociale et par réflexion, cette personne en arrive à la conclusion que sa société est injuste.

Here, social injustice seems to explain both the feelings [she] experienced before acquiring a belief in the injustice of [her] society and, of course, the development of this belief. So even if there is nothing wrong with citing [her] moral beliefs as part of the explanation of [her] subsequent protest activities, this belief cannot explain [her] earlier nonmoral psychological changes or the change in [her] moral consciousness. (Brink, 1989, p.189)

Toutefois, selon Harman, la différence entre une observation dans le cas scientifique ou dans le cas moral, c'est qu'une observation ne constitue pas de preuve en faveur ou en défaveur d'une théorie morale particulière puisque la vérité ou la fausseté de l'observation semble non pertinente à une explication de pourquoi l'observation a été faite. Le fait de faire une observation morale n'est pas une preuve d'un fait moral, mais témoigne de la sensibilité morale de l'observateur. Les faits moraux ne nous sont d'aucune aide pour expliquer pourquoi nous faisons l'observation. Par contre, dans le cas scientifique, l'observation constitue un indice en faveur de la théorie. La chaîne de l'explication allant du principe à l'observation semble rompue dans le cas moral. Dans le cas scientifique, l'observation constitue une preuve en faveur de la théorie physique car la théorie explique le proton qui explique la trace de vapeur. Dans le cas moral, l'observation ne constitue pas une preuve en faveur du principe moral car ce principe n'est d'aucune aide pour expliquer l'observation.

The moral principle may "explain" why it is wrong for the children to set the cat on fire. But the wrongness of that act does not appear to help explain the act, which you observe, itself. The explanatory chain appears to be broken in such a way that neither the moral principle nor the wrongness of the act can help explain why you observe what you observe. (Harman, 1988, p.123)

Nicholas Sturgeon a répondu à cet argument de Harman. Il pense qu'il y a des explications morales qui sont de bonnes explications ou des composantes de bonnes explications. Sturgeon fait partie des réalistes de Cornell et est donc un réaliste moral naturaliste non réductionniste. Il s'oppose au réductionnisme dans la mesure où il ne voit pas en quels termes le vocabulaire moral

devrait être défini. On ne saurait réduire le vocabulaire moral à celui de la physique car il y a plus de propriétés physiques que d'expressions physiques pour les représenter. En effet, il y a un continuum entre les états physiques du monde, alors que nous disposons d'un nombre fini de prédicats physiques. Même en admettant que le physicalisme soit vrai et que les propriétés biologiques, mentales soient en fin de compte physiques, nous ne disposerions pas d'autres terminologies que les terminologies biologique ou psychologique pour les représenter. Sturgeon ne voit pas non plus pourquoi il faudrait réduire le vocabulaire éthique au vocabulaire psychologique ou sociologique dans la mesure où l'on exige pas de ces disciplines-là qu'elles soient réductibles les unes aux autres.

Sturgeon est d'avis que le fait qu'Hitler était une personne moralement dépravée explique les actes qu'il a commis. En effet, une personne moralement admirable ne pourrait pas instiguer la mort de milliers de personnes. Nous citons souvent le caractère moral d'une personne pour expliquer ses actes. Sturgeon note que

[...] it would be difficult to find a serious work of biography, for example, in which actions are not explained by appeal to moral character: sometimes by appeal to a specific virtues and vices, but often enough also by appeal to a more general assessment. (Sturgeon, 1988, p.244)

Cependant, l'exemple proposé par Harman est destiné à montrer que nul n'est besoin de faire appel à la malignité de l'action pour expliquer la croyance de l'observateur que cette action est mauvaise. Sturgeon pense que les traits moraux d'une action ou d'une institution peuvent aussi jouer un rôle explicatif. Il donne comme exemple l'opposition massive à l'esclavage apparue aux XVIII^e et XIX^e siècles alors même que l'esclavage était une institution ancienne. Une explication possible de cette opposition consiste à dire que l'esclavage était devenu plus dur encore à cette époque. L'explication repose alors sur un trait moral de l'institution de l'esclavage, à savoir qu'elle était plus mauvaise encore aux XVIII^e et XIX^e siècles. On peut aussi expliquer l'instabilité politique d'un pays en invoquant l'injustice sociale qui y règne. Dans l'exemple de Harman, des enfants mettent le feu à un chat pour le plaisir. Harman soutient qu'il n'est pas pertinent de faire appel à la malignité de l'action pour expliquer la croyance que cette action est mauvaise. Cela implique que l'observateur aurait eu cette croyance même si l'acte n'était pas mauvais. Sturgeon répond qu'afin que l'action ne soit pas mauvaise, il aurait fallu qu'elle ne possède pas les traits de cruauté délibérée et inutile et que si elle n'avait pas posséder ces traits, l'observateur n'aurait

probablement pas formé la croyance que cette action était mauvaise. De plus, l'action doit être différente d'un point de vue naturel pour ne pas être mauvaise. En effet, il est impossible que de deux actions indiscernables du point de vue des propriétés naturelles, l'une soit mauvaise et l'autre non. Si cette action avait été différente d'un point de vue naturel, rien ne nous garanti que l'observateur aurait maintenu sa croyance que l'action était mauvaise.

Cependant Sturgeon pense qu'Harman répondrait que son exemple vise à montrer que si l'action des enfants mettant le feu au chat était un exemple de cruauté délibérée et inutile, mais que cela ne la rendait pas mauvaise, l'observateur conserverait sa croyance que l'action est mauvaise. Or Sturgeon pense, pour reprendre son exemple, qu'il n'y a pas de monde possible où Hitler a la même personnalité que celle qu'il y a eue dans le monde actuel, dans la même situation, et où Hitler n'est pas moralement dépravé. Si certains arrangements sociaux, politiques et légaux constituaient l'injustice du III^e Reich, il est impossible que ces faits naturels demeurent pareils mais que l'injustice n'existe pas. Sturgeon propose ainsi un test contrefactuel. Selon ce test, pour voir si *H* explique *E*, il faut voir si *E* aurait eu lieu en l'absence de *H*. Si *E* n'aurait pas eu lieu, son occurrence est en partie expliquée par *H*. Si *H* est le fait qu'Hitler était moralement dépravé et *E* le fait d'instiguer des milliers de morts, la question est de savoir s'il est possible d'instiguer des milliers de morts sans être moralement dépravé. Sturgeon répond par la négative à cette question et c'est pourquoi il pense que le fait qu'Hitler était moralement dépravé explique le fait qu'il ait instiguer des milliers de morts.

La réponse d'Harman implique qu'il est possible que nous nous trompions radicalement dans notre théorie morale. Puisqu'il est possible que notre théorie soit totalement fausse, les faits non moraux pourraient demeurer exactement pareils, nous continuerions à former les mêmes croyances, alors même que ces croyances seraient fausses. Toutefois, Sturgeon ne pense pas que nos théories morales soient à ce point erronées qu'elles ne sont en aucun cas fiables.

[...] if Hitler hadn't been morally depraved we would still have thought he was: for we may assume that our moral theory also is a least roughly correct, and it follows from the most central features of that theory that if Hitler hadn't been morally depraved, he wouldn't have done what he did. The *parallel* claim about the microphysical example is, instead, that if there hadn't been a proton there, but there *had* been a vapor trail, the physicist would still have concluded that a proton was present. (Sturgeon, 1988, p.252)

Sturgeon établit donc un parallèle entre le cas moral et le cas du scientifique. Si la théorie du physicien était erronée et si les faits microphysiques demeuraient les mêmes, le physicien arriverait à la même conclusion. Cependant, dans le cas scientifique, cela ne signifie pas qu'il n'y a pas de faits microphysiques. En effet, nous n'avons aucune raison de penser que la théorie du physicien est totalement fautive. Ainsi l'hypothèse que c'est un proton qui a causé la trace de vapeur est une explication tout à fait plausible du phénomène. Sturgeon est d'avis que dans le cas moral, nous n'avons pas, non plus, de raison de penser que notre théorie morale est totalement fautive. Les explications morales ne seront plausibles que dans la mesure où notre théorie morale est en gros correcte, puisque c'est notre théorie morale qui nous indique que, par exemple, seules les personnes moralement dépravées peuvent instiguer la mort de milliers de personnes. Ainsi, ce n'est que dans le cas où nous avons déjà des doutes au sujet des théories morales, que nous penserons que les faits moraux ne jouent aucun rôle explicatif. Si nous acceptons une théorie morale, les faits moraux auront des conséquences observables. David Brink donne l'exemple suivant pour illustrer ce phénomène:

[...] my principle that it is good people who keep their promises when doing so involves great personal sacrifice has no observational implications by itself. But when I conjoin it with my independent moral belief that Zenobia is a good person, I can obtain the observational prediction that Zenobia will keep her promise to Zelda even at great personal sacrifice. My independent belief that Barney is thoroughly rotten, together with my moral principle, leads me to expect that Barney will not suffer the personal hardship needed to keep his promise to Bonny. (Brink, 1989, p.184)

La thèse de Harman selon laquelle les faits moraux ne sont pas pertinents pour expliquer des observations présuppose donc le scepticisme moral. Si nous avons confiance dans les théories morales, les faits moraux peuvent jouer un rôle explicatif. Harman doit donc fournir une raison indépendante pour douter des théories morales, selon Sturgeon, afin que son argument soit valide.

Cependant, Harman n'accepte pas que les exemples d'explications morales fournis par Sturgeon soient de véritables explications. En effet, il ne s'agit pas d'explications permettant de tester les énoncés moraux en les confrontant au monde, alors que cela est possible dans le cas d'explications scientifiques. Admettons qu'Albert bat son chat et que Jane croit que cet acte est mauvais. Sturgeon accepte que la malignité de l'action explique la croyance de Jane dans la

mesure où si cette action n'avait pas été mauvaise, Jane ne serait pas parvenue à cette croyance. Harman ne partage pas cet avis:

The problem is that this sort of explanation does not make clear how the actual *wrongness* of Albert's act plays a role in generating Jane's belief. In the absence of a naturalistic reduction of wrongness, Jane's belief can be explained entirely by Jane's moral sensibility plus an observation of Albert's action, whether or not there is anything wrong with causing pain to animals. (Harman, 1986, p.62)

Tant que nous parviendrons pas à des réductions naturalistes des énoncés moraux, nous ne pourrons pas les tester empiriquement comme les énoncés scientifiques. Pour illustrer son point, Harman imagine une autre observatrice, Mary, qui ne pense pas qu'il est mal d'infliger de la souffrance aux animaux. Mary peut néanmoins expliquer la croyance de Jane et cela sans faire appel à la malignité de l'acte d'Albert.

Harman est d'avis que tant que nous ne disposerons pas de réductions des énoncés moraux à des énoncés psychologiques ou sociologiques, le problème de la testabilité des énoncés moraux demeurera d'actualité. Il ne voit pas comment la bonté ou la malignité d'une action peuvent se manifester dans le monde d'une manière qui puissent affecter les sens des gens. Harman suggère que les variétés de réductionnisme les plus plausibles sont les théories du spectateur impartial et les théories du raisonnement pratique. D'après les théories du spectateur impartial, dire qu'une action est mauvaise revient à dire qu'il s'agit du genre d'action qu'un spectateur impartial désapprouverait sous certaines conditions spécifiées par la théorie en question. On peut alors expliquer la croyance de Jane en disant que l'action d'Albert est le genre d'action qu'un spectateur impartial désapprouverait et que Jane fait partie des spectateurs dont l'opinion est pertinente. L'action d'Albert possède certaines propriétés qui font en sorte qu'un spectateur impartial la désapprouverait et ce sont ces propriétés qui amènent Jane à croire que l'action d'Albert est mauvaise. Dans le cas des théories du raisonnement pratique, un acte est mauvais si l'agent dispose de raisons suffisantes de ne pas l'accomplir. Cela revient à dire que si l'agent raisonnait correctement, il déciderait de ne pas faire cette action. L'action d'Albert est mauvaise car il dispose de raisons suffisantes de ne pas la commettre. Si Jane croit que son action est mauvaise parce qu'il n'a pas fait le raisonnement correcte, sa croyance est expliquée, en partie, par la malignité de l'acte d'Albert.

Pour revenir au cas de Mary, Sturgeon admet qu'il n'y a rien dans l'observation qui l'oblige à changer d'avis et à croire qu'il est mal d'infliger des souffrances inutiles aux animaux. Toutefois, pour que ceci constitue un problème pour la possibilité d'explications morales, il faudrait que ce genre de situation ne puisse pas se produire en dehors de l'éthique. Or, Sturgeon pense que ce genre de cas peut se produire en physique:

Harry sees a vapor trail in a cloud chamber and thinks "There goes a proton". Susan, who has been listening to too many neopositivist arguments in a philosophy of science course, is utterly agnostic about the existence of microphysical particles; she notes Harry's reaction, but explains it as due merely to his "physical sensibility" plus his observation of the vapor trail. (Sturgeon, 1986, p.71)

De plus, dans le domaine de la psychologie, les freudiens et leurs opposants tirent des conclusions différentes des mêmes observations et chaque camp est capable d'offrir une explication du point de vue de l'autre. Néanmoins, dans ces disciplines, des observations amènent souvent à des changements de croyances. La différence avec l'éthique serait alors que dans le domaine moral nous ne changerions jamais nos jugements étant confrontés à des observations. Cependant, Sturgeon pense qu'il est possible que l'observation d'un acte de cruauté et la constatation que Jane croit que cet acte est malveillant puissent amener Mary à réviser ses croyances. Cela se produirait si Mary considérait que Jane a un jugement fiable concernant les questions morales.

Because of her view about Jane's reliability as a judge, a view about moral explanations, this nonmoral observation will force a change somewhere in her moral view: either Jane is less admirable than she thought, or admirable people can be mistaken more easily than she thought, or else there is more wrong in Albert's action than she thought. If her evidence about Jane is quite good, [...] her suspicions may focus on Albert. [...] Observational evidence has change her mind. (Sturgeon, 1986, p.72)

Harman soutient que les faits moraux ne jouent aucun rôle explicatif pertinent et nous venons de voir que Sturgeon lui réplique que son argument présuppose le scepticisme moral. C'est pourquoi Geoffrey Sayre-McCord propose de changer le critère de pertinence au niveau de l'explication pour celui de pouvoir explicatif. Il pense que le critère de la pertinence permet trop facilement de défendre l'existence de propriétés morales. La question est alors de savoir si la reconnaissance de faits moraux nous permet de mieux expliquer nos expériences. Sayre-McCord

pense que nous pouvons très bien nous passer des faits moraux. Il relève que le fait que nous fassions appel à des faits moraux pour expliquer le comportement d'autrui n'est pas suffisant pour justifier notre croyance en des faits moraux. Sturgeon fait notamment appel à ceci pour montrer que des traits moraux peuvent faire partie d'explications. Sayre-McCord compare le cas moral avec une situation où la croyance aux sorcières deviendrait répandue. On pourrait apprendre que si le corps de certaines femmes flottent lorsqu'on les jette à l'eau et d'autre non, c'est que les premières sont des sorcières, etc. Avec un peu d'entraînement, nous serions sûrement capable de fournir des explications faisant intervenir des sorcières et nous pourrions alors conclure, en suivant les arguments de Sturgeon, que le fait d'être une sorcière est un trait pertinent afin d'expliquer nos observations. Toutefois, si ces observations pouvaient être expliquées sans faire appel à l'existence de sorcières, nous serions justifier de penser que les sorcières n'existent pas. La question que se pose Sayre-McCord est de savoir s'il est possible d'expliquer nos observations sans postuler de faits moraux.

Are our appeals to moral properties just intellectually sloppy concessions to effective socialization, or do we really strengthen our explanatory abilities by supposing that there are moral properties? This is a substantial challenge, and one not adequately answered by the observation that we often rely on moral properties to explain behavior. (Sayre-McCord, 1988, p.273)

Non seulement, les faits moraux doivent figurer dans nos meilleures explications, mais encore faut il que ces explications soient liées de manière adéquate à nos expériences du monde.

Sayre-McCord est un réaliste moral et il pense donc qu'il y a des propriétés morales. Sa défense des faits moraux repose plus sur la notion de justification que d'explication. Les hypothèses qui sont justifiées reposent sur des standards d'explication, puisque que c'est en figurant dans les meilleures explications disponibles qu'elles se trouvent justifiées. Si l'on pense que le critère explicatif mène au rejet de tous fait évaluatif, cela pose problème pour déterminer quelle est la meilleure explication. En cherchant à rendre légitime certains jugements, nous faisons appel à des standards d'explication et, ce faisant, nous faisons appel à des valeurs. Il y a au moins certains faits évaluatifs, c'est ceux qui concernent les meilleures explications. Sayre-McCord note que dans le cas d'évaluation d'actions et d'évaluation d'explication, nous cherchons à justifier notre jugement et non à expliquer nos expériences. Voici la défense qu'offre Sayre-McCord des valeurs morales:

[...] we might begin with evaluations of explanations, move up (in generality and abstraction) to principles justifying these evaluations, then move back down, along a different justificatory path, to recognizably moral evaluations of actions, characters, institutions, and so on. (Sayre-McCord, 1988, p.279)

Il pense que les jugements moraux particuliers ainsi que les principes moraux généraux peuvent être rendu légitimes au travers de leur connexion avec la théorie des valeurs qui sert à justifier nos jugements explicatifs.

Nous venons donc de voir trois réponse possibles au problème du rôle explicatif des faits moraux. Les trois auteurs s'entendent sur le fait que s'il est possible de fournir des explications de nos observations en nous passant de toute référence à des faits ou propriétés morales, il faut bannir ceux-ci de notre ontologie ou, au mieux, les considérer comme un épiphénomène n'ayant aucun pouvoir explicatif. Harman pense que les faits moraux ne pourraient jouer de rôle explicatif qu'à la condition d'être réductibles à des faits naturels mais que, en l'absence de telles réductions, ils ne sont absolument pas pertinents pour expliquer nos observations. Sturgeon refuse de réduire les faits moraux à des faits naturels et pense que les faits et traits moraux passent le test contrefactuel. Sayre-McCord pense que le critère proposé par Harman est trop tolérant. Néanmoins tout critère explicatif suppose de justifier un jugement s'il fait partie de la meilleure explication. Postuler une meilleure explication, c'est faire appel à un fait évaluatif. Sayre-McCord dérive ensuite l'existence de faits moraux de ces faits évaluatifs.

On pourrait reformuler le problème de l'explication morale en se demandant comment des propriétés survenantes peuvent jouer un rôle explicatif. On pourrait alors répondre que la réalisation de la propriété survenante assure la présence de propriétés de base qui elles possèdent une efficacité causale. Ainsi les propriétés morales pourraient être pertinente d'un point de vue explicatif même si elles ne sont pas efficaces au niveau de la causalité. Les propriétés morales assurent la présence de propriétés naturelles causalement efficaces. Néanmoins, nous ne pouvons pas remplacer les explications morales par des explications ne tenant en compte que des propriétés de base. Comme l'a relevé Sturgeon, le vocabulaire de la physique serait trop pauvre pour parvenir à la même richesse explicative que l'explication morale. Les explications morale seraient ainsi indispensables pour des raisons pratiques puisqu'il est extrêmement difficile d'identifier toutes les faits de bases physiques qui constituent un fait moral. Cependant, nous avons une bonne idée

de quels faits sociaux, politiques, économiques ou légaux constituent des faits moraux injustes, par exemple. Pouvons-nous maintenir la pertinence des explications morales dans le cas où nous pouvons identifier quels faits naturels constituent le fait moral? Brink relève que nous commençons souvent par employer une explication morale en disant, par exemple, que c'est l'injustice sociale qui est la cause des troubles politiques.⁴⁰ Nous nous appuyons sur une théorie morale concernant la nature et les causes de l'injustice, puis ensuite citons les structures particulières sociales, politiques et légales en tant que facteurs explicatifs. Si le fait moral, que telle situation est injuste, est constitué par des faits sociaux et politiques, alors même si je ne parle que de ces faits naturels, je serai en train de citer un fait moral. En effet, ces faits naturels constituent le fait moral. Brink suggère qu'il ne faut pas forcément voir les explications en termes des faits constituants et en termes de faits constitués comme étant en rivalité. Ils sont plutôt complémentaires. Par exemple, il relève que je peux expliquer pourquoi ma crème glacée coule sur mon pantalon en faisant appel à la température de la pièce. Cette explication n'est néanmoins pas supplantée par une explication en termes d'énergies cinétiques moléculaires moyennes. Elle est plutôt un complément à la première explication. De la même façon, les faits moraux peuvent être explicatifs, même si une explication en termes des faits naturels qui les constituent est un ajout à l'explication morale. Brink pense aussi qu'il y a plus de généralisations intéressantes concernant les faits survenants que concernant les faits qui réalisent ces derniers. Par exemple, il y a plus de généralisations intéressantes concernant les états mentaux des gens et leurs actions que de généralisations intéressantes concernant les états neuronaux. En effet, les états mentaux peuvent être réalisés de multiples façons et les arrangements des états physiques constituant un état mental ne représentent pas forcément une espèce naturelle. Puisqu'il est plus aisé d'obtenir des généralisations concernant les faits survenants, ces faits possèdent un rôle explicatif.

The particular arrangement of natural facts that could realize explanatory moral facts (types or tokens) may in many cases represent no natural kind from the point of view of economics, social theory, or nutritional theory, and certainly not from the point of view of microphysics. If so, moral explanans will generalize better than would explanans in terms of the lower-order facts that constitute these moral facts. (Brink, 1989, p.195)

⁴⁰Brink, 1989, p.192

Il vaut mieux citer l'injustice sociale comme cause de l'instabilité d'un pays que des faits sociaux, politiques et économiques particuliers parce qu'il aurait encore pu y avoir de l'injustice même avec des faits sociaux, politiques et économiques différents. Toutes les bases sociales, économiques et politiques possibles de l'injustice ne forment pas une catégorie unifiée. C'est pourquoi, les explications morales ont un rôle explicatif privilégié.

3.4 Conclusion

Dans ce chapitre, nous avons essayé de voir s'il était possible qu'il y ait des faits moraux. La plupart des réalistes font appel à la théorie de la survenance pour rendre les faits moraux compatibles avec une ontologie naturaliste. Les faits moraux sont réalisés par des faits naturels et les propriétés morales sont réalisées par des propriétés naturelles et ils ne sont en rien des entités mystérieuses. Les faits moraux seraient, en ce sens, similaires aux faits mentaux ou macroscopiques, etc qui eux, également, surviennent sur des propriétés de base. Il me semble que le réalisme non réductionniste est plus convaincant puisqu'il permet la réalisation multiple des propriétés morales. Différentes combinaisons peuvent réaliser la même propriété morale et c'est pourquoi l'injustice n'est pas identifiée à une seule configuration de facteurs économiques et sociaux, mais peut être constituée d'autres facteurs. Les propriétés morales ne sont donc rien de plus que des configurations de propriétés naturelles. D'ailleurs, David Brink relève que la notion de constitution ou de réalisation est assez familière et ne pose pas problème dans un grand nombre de cas.⁴¹ Les objets sont constitués d'arrangements de particules microphysiques, les processus biologiques sont constitués d'événements physico-chimiques, les états psychologiques sont constitués d'états du cerveau, etc. La question de savoir quels propriétés et faits naturels constituent telle propriété ou tel fait moral doit être répondue avec l'aide des théories morales normatives.

Il y aurait ainsi plusieurs niveaux d'explication. On peut expliquer les choses au niveau physique, mais aussi au niveau moral ou mental. Les explications en termes des faits survenant et en termes des faits qui les réalisent ne sont pas en concurrence. Les faits moraux sont constitués de faits naturels et un fait moral est réalisable de façon multiple. Comme il n'y a pas de catégorie unifiée correspondant à toutes les configurations de faits naturels pouvant réaliser tel fait moral,

⁴¹Brink, 1989, p.177

une explication morale aura plus de pertinence. Il est, en effet, plus facile d'obtenir des généralisations en faisant appel aux faits moraux qu'aux faits naturels les constituant.

Si on accepte le réalisme moral, il n'y a plus de raison de conserver une fracture ontologique entre les faits et les valeurs, puisque le réalisme suppose qu'il existe des faits axiologiques. Dans le monde, nous avons affaire à une multitude de faits différents, dont des faits moraux, des faits mentaux, des faits biologiques, des faits institutionnels, etc. Les faits axiologiques surviennent sur des faits de base naturels et on pourrait alors formuler différemment la dichotomie fait-valeur en tant que dichotomie entre propriétés de base et propriétés survenantes. Néanmoins, cette dernière distinction ne permet pas d'exclure l'éthique ou l'esthétique du monde naturel. La théorie de la survenance permet de comprendre le monde en tant que constitué de couches différentes. À la base, nous trouvons les propriétés physiques sur lesquelles surviennent de nombreuses autres propriétés. Chaque niveau de survenance a sa légitimité explicative. Nous perdrons de l'information à ne vouloir donner que des explications en termes physiques.

Conclusion

Dans ce travail, j'ai passé en revue quelques arguments donnés par les non-cognitivistes en faveur d'une dichotomie fait- valeur. Premièrement, j'ai examiné l'argument selon lequel les énoncés moraux ne sauraient avoir de valeur de vérité puisqu'ils ne réfèrent à rien dans le monde et qu'il est impossible de les vérifier empiriquement. Nous avons vu les problèmes inhérents au vérificationnisme. En effet, même en science, les principes ne peuvent pas être testés de manière indépendante et de nombreux principes théoriques de la science n'ont aucune implication concernant les observations lorsqu'ils sont pris individuellement. Des prédictions vérifiables par l'observation peuvent être dérivées de principes scientifiques uniquement lorsque'elles sont combinées à un arrière-plan approprié. C'est la théorie dans son ensemble qui peut être vérifiée et non des énoncés individuels. Le fait que les énoncés moraux ne sont pas vérifiables individuellement, n'est alors pas suffisant à montrer qu'ils ne réfèrent à rien. Par ailleurs, en refusant des conditions de vérités aux énoncés moraux, le non-cognitivisme met à mal toute possibilité de raisonnements moraux. En effet, l'expressivisme, même dans ses versions les plus élaborées, a de la difficulté à rendre compte du fonctionnement des énoncés évaluatifs dans des contextes non assertés, dont les conditionnels. Il lui est alors difficile d'expliquer la validité des *modus ponens* moraux. Le problème des solutions expressivistes, c'est qu'elles emploient une notion de cohérence logique qui est trop large et qui permet d'y inclure ce qui n'est que de l'incohérence pragmatique.

J'ai ensuite examiné les problèmes liés à la nature motivationnelle des jugements moraux. Les non-cognitivistes soutiennent deux thèse, la thèse internaliste et la théorie humienne de la motivation, pour prouver que les jugements moraux ne peuvent pas être des croyances, mais seulement des projections d'états affectifs. La thèse internaliste stipule qu'il y a un lien nécessaire entre émettre un jugement moral et être motivé en fonction de celui-ci. J'ai argumenté en faveur de la thèse externaliste qui soutient que ce lien n'est que contingent. Les cas de faiblesse de la volonté, de dépression, etc, sont des contre-exemples à la thèse internaliste forte, alors que la possibilité de l'amoraliste va à l'encontre de la thèse internaliste faible. Par contre, j'admet que la théorie humienne me paraît plausible. En effet, pour toute explication en termes de croyances évaluatives, il me semble qu'on peut trouver un désir plus fondamental à l'arrière-plan. Toutefois la théorie humienne soutient le non-cognitivisme qu'à la condition que l'internalisme soit vrai, ce que je réfute.

Pour terminer, j'ai regardé les questions liées au réalisme moral. Il me semble que si le réalisme moral est vrai et que, par conséquent, il y a des faits moraux, il n'y a plus de raison de postuler une fracture ontologique profonde entre faits et valeurs. Il y aurait juste des niveaux de réalité différentes, les propriétés évaluatives survenant sur les propriétés naturelles. J'ai examiné plusieurs défis posés au réalisme. J'ai d'abord examiné l'argument de la question ouverte qui met à mal le réalisme moral naturaliste réductionniste. D'après cet argument, on ne saurait établir de relation d'identité entre propriétés morales et naturelles. Nous avons vu que cet argument s'attaquait aux relations d'identité analytiques, mais qu'il existe de relations d'identité synthétique qui sont établies par investigation. Un réaliste naturaliste réductionniste peut alors concevoir que par une réflexion sur les théories morales normatives il sera un jour possible de parvenir à la meilleure théorie qui établira de telles identités entre propriétés morales et naturelles. Par contre, l'argument de la question ouverte n'atteint pas le réalisme moral non naturaliste, ni le réalisme moral naturaliste non réductionniste. Ces théories affirment, en effet, que les propriétés morales surviennent sur les propriétés naturelles, mais non pas qu'elles sont identiques. Il y a une relation asymétrique entre les deux. Si tout changement au niveau naturel implique une modification au niveau des propriétés survenantes, ces dernières peuvent être réalisées de manière multiple. Des ensembles de propriétés naturelles différents peuvent donc être la base d'une même propriété survenante. Néanmoins, Blackburn, ainsi que Horgan et Timmons pensent que le réalisme ne permet pas de rendre compte de la survenance. Nous avons examiné ces arguments et vu que le réalisme dispose de moyens suffisants pour répondre à ces difficultés.

Je n'ai, cependant, pas eu la place pour examiner tous les arguments mettant en doute la pertinence de la dichotomie fait-valeur et je voudrais les évoquer brièvement ici. Il existe notamment un argument qui est souvent mis à contribution lorsqu'il s'agit de défendre la dichotomie fait-valeur: il s'agit du problème de l'inférence logique entre énoncés factuels et axiologiques ou entre énoncés factuels et normatifs. En effet, la dichotomie fait-valeur a souvent été mise en parallèle avec la loi de Hume qui établit qu'on ne peut pas inférer un devoir-être à partir d'un être. Selon cet argument, on ne saurait tirer de conclusion évaluative à partir de prémisses purement factuelles. Pour parvenir à une conclusion évaluative, il faudrait toujours qu'au moins une des prémisses soit évaluative. John Searle, cependant, estime être parvenu à un exemple d'inférence d'un "doit" à partir d'un "est". L'inférence qu'il propose est la suivante:

- (1) Jones prononce les mots « Je te promets de te donner cinq dollars, à toi, Smith. »
- (2) Jones a promis à Smith de lui donner cinq dollars.
- (3) Jones s'est placé lui-même sous l'obligation de donner cinq dollars à Smith.
- (4) Jones est sous l'obligation de donner cinq dollars à Smith
- (5) Jones doit donner cinq dollars à Smith.

L'argument est formellement valide, la première prémisse est factuelle et la conclusion (5) est normative. Nous avons vu au premier chapitre que, selon Searle, les faits institutionnels, comme la promesse, existent au sein de systèmes de règles constitutives. Ces systèmes de règles constitutives impliquent des obligations, des engagements, des responsabilités. C'est au sein de tels systèmes que l'on peut, selon lui, inférer un "doit" d'un "est".

Toutefois, même en refusant l'argument de Searle, Ruwen Ogien relève que la version logique de la dichotomie fait-valeur est assez inoffensive. Tout ce qu'elle dit, c'est qu'un terme ne peut pas apparaître dans les conclusions s'il ne figure pas dans les prémisses du raisonnement. Cette limite s'applique aussi aux tentatives de dériver des propositions portant sur les épinarads de propositions portant sur les philosophes grecs.⁴² De plus, même si on ne peut pas dériver logiquement un "doit" d'un "est", cela ne suffit pas à conclure qu'il n'existe aucun autre lien entre "devoir être" et "être".

Par ailleurs, aucune proposition d'une discipline donnée ne peut être déduite de propositions appartenant à une autre discipline sans ajouter une prémisse faisant le lien nomologique entre les deux disciplines. David Brink donne comme exemple l'impossibilité de déduire les lois de la thermodynamique de lois de la mécanique statistique sans l'aide d'une prémisse additionnelle comme la loi de Boyle-Charles.⁴³ Ainsi avec l'aide d'une prémisse faisant le lien, il est possible de dériver une conclusion morale de prémisses factuelles. L'éthique se retrouve aussi sur un pied d'égalité avec les autres disciplines puisque'on ne peut pas non plus dériver de conclusion biologique de prémisses venant de la physique fondamentale sans ajouter

⁴²Ogien, 2004, p.297

⁴³Brink, 1989, p.169

une prémisses liant nomologiquement certaines propriétés physiques à certaines propriétés biologiques.

Il y a, en outre, des raisons de penser que certains concepts illustrent un enchevêtrement entre des éléments factuels et évaluatifs. C'est le cas, par exemple, des concepts moraux épais. Ceux-ci me semblent être des concepts mixtes contenant une composante descriptive et une composante évaluative. La question est de savoir si on peut facilement séparer ces deux composantes, comme l'affirme le prescriptivisme, ou si elles sont inextricablement mêlées, comme le soutient Bernard Williams.

Bernard Williams dans *Ethics and the Limits of Philosophy* établit une distinction entre concepts moraux minces et épais. Les concepts moraux épais, tels que "courageux", "généreux", "malhonnête", etc. sont plus spécifiques, plus riches que les concepts minces tels que "bien", "mal", "meilleur", "pire", "doit" qui sont généraux et abstraits. Les jugements impliquant des concepts épais sont à la fois déterminés par le monde (*world-guided*) et guides pour l'action (*action-guiding*), ils donnent souvent une raison d'agir. L'emploi de tels concepts dépend des états de choses dans le monde, mais il détermine aussi ce que nous devrions faire. Par contre, les jugements comprenant un concept mince ne font que guider l'action. Ils n'ont pas de contenu descriptif et se contentent de nous indiquer quoi faire.

L'intuition qui guide une analyse à deux composants des concepts moraux épais, c'est que lorsqu'une action est courageuse, par exemple, elle correspond à une prise de risque, à une maîtrise de soi face au danger mais que, par ailleurs, elle est aussi une bonne chose. Ainsi les concepts minces, comme "bon", seraient plus généraux que les concepts épais. En effet, comme on ne peut pas leur attribuer de partie descriptive, ils sont moins spécifiques. Afin d'illustrer cette intuition, on peut donner un exemple. L'éditorialiste qui après les attentats du 11 septembre écrit "Les terroristes étaient courageux", signifiant par là qu'ils ont surmonté leur peur face à une situation de danger, a suscité une réaction d'indignation. Le terme "courageux" ne pouvait pas convenir à la qualification de terroristes. Or, cela semble indiquer que ce terme a bien une connotation positive. De même, lorsqu'on est prêt à qualifier une action de cruelle, il semble qu'on la désapprouve fortement.

Williams veut montrer que les concepts épais expriment une union entre faits et valeurs. En effet, la façon dont nous employons ces termes est déterminée par comment est le monde, mais cet emploi implique également une évaluation de l'état de choses considéré.

What they have found are a lot of those “thicker” or more specific ethical notions I have already referred to, such as *treachery* and *promise* and *brutality* and *courage*, which seem to express a union of fact and value. The way these notions are applied is determined by what the world is like (for instance, by how someone has behaved), and yet, at the same time, their application usually involves a certain valuation of the situation, of persons or actions. Moreover, they usually (though not necessarily directly) provide reasons for action. Terms of this kind certainly do not lay bare the fact-value distinction. (Williams, 1985, pp. 129-130)

Hilary Putnam s'est également penché sur la question des concepts moraux épais.⁴⁴ Il note que le terme “cruel” a un emploi éthique, plus précisément il possède une connotation négative. Par exemple si je dis “Mon oncle est très cruel”, je n'ai pas besoin d'ajouter “Ce n'est pas une bonne personne”. On ne comprendrait pas un énoncé comme “C'est un homme cruel et une bonne personne”. Cela tendrait à montrer que “ne pas être une bonne personne” fait partie de la signification de “cruel”. Toutefois, “cruel” a aussi un emploi descriptif. Par exemple, lorsqu'un historien écrit que tel empereur était très cruel et que sa cruauté a provoqué des révoltes, il décrit un état de choses. Selon Putnam, les concepts moraux épais ignorent la dichotomie fait-valeur et ont parfois un emploi axiologique, parfois un emploi descriptif.

Les non-cognitivistes peuvent offrir les réponses suivantes. Ils peuvent, en premier lieu, dire que les concepts moraux épais sont de simples concepts factuels et non pas des concepts éthiques ou normatifs. Ils peuvent aussi dire que les concepts épais sont séparables en une composante purement descriptive et une composante attitudinale. La composante descriptive rend compte d'un état de choses auquel correspond le prédicat et la composante attitudinale exprime une attitude, émotion ou volition.

Le problème de l'analyse à deux composantes, c'est qu'il est impossible de dire ce que la “signification descriptive” de “cruel” dit sans employer le mot “cruel” ou un synonyme. Il n'est pas le cas que l'extension de “cruel” est simplement “causer d'intenses souffrances”. Par exemple, avant l'anesthésie, les chirurgiens causaient de grandes souffrances sans généralement être cruels. Par ailleurs, des comportements qui ne causent pas de souffrances évidentes peuvent néanmoins

⁴⁴Putnam, 2002, pp.34-43

être cruels. En effet, on pourrait dire d'une personne entraînant un jeune dans la débauche pour l'empêcher de développer de grands talents qu'elle est cruelle. Il est difficile de déterminer à quels traits du monde référerait ces concepts épais une fois pelés de l'aspect évaluatif.

Putnam pense que l'enchevêtrement entre faits et valeurs n'est pas seulement à l'oeuvre dans les concepts moraux épais.⁴⁵ Cet enchevêtrement a une portée beaucoup plus large. Lorsque nous effectuons des évaluations éthiques, les descriptions dont nous avons besoin, qu'elles concernent des motifs ou des caractères d'êtres humains, sont généralement des descriptions énoncées dans le langage des romanciers et non dans celui du jargon scientifique. Lorsqu'une situation ou une personne est adéquatement décrite, la décision quant à savoir si celle-ci est bonne ou correcte s'ensuit fréquemment automatiquement. Le monde dans lequel nous vivons ne se divise pas de manière claire et précise entre faits et valeurs, nous vivons dans un monde humain complexe dans lequel voir la réalité avec toutes ces nuances, c'est-à-dire à la manière des grands romanciers tels que Flaubert et Dostoïevski, et porter des jugements de valeur appropriés ne sont pas des aptitudes séparées. Des textes littéraires peuvent, par ailleurs, nous aider à former une appréciation d'une situation donnée en organisant et forgeant notre attitude au monde et une rencontre avec un texte peut modifier notre perspective morale. En outre, je pense que certains écrits et discours peuvent être l'expression d'une pensée éthique, bien qu'ils ne contiennent pas ou peu de termes spécifiquement moraux. Imaginons un texte témoignant de l'atrocité des camps de concentration nazis. Un tel texte pourrait être essentiellement descriptif et néanmoins exprimer une position morale. La portée éthique serait ainsi comme l'humour ou l'ironie, que l'on peut attribuer seulement en fonction d'une appréciation globale.

Selon Putnam l'enchevêtrement des faits et des valeurs s'applique non seulement à la sphère privée décrite par les romanciers, mais également à la sphère publique. Il ne pense pas qu'on puisse construire des cas où tous s'accordent sur les faits, mais divergent sur l'évaluation de ces derniers. Il est d'avis que les descriptions employées par chacun seront différentes. Pour lui, concernant une question de justice sociale, par exemple, un communiste et un social-démocrate ne s'accordent pas sur les faits. Ils emploient, pense-t-il, chacun une description différente des faits. Ceci tendrait à montrer que les faits et les valeurs sont entremêlés dans notre expérience du monde. Nous ne ferions pas de distinction nette et tranchée entre énoncés factuels

⁴⁵Putnam, *Reason, Truth and History*, 1981, pp. 165-66

et jugements évaluatifs, mais nous aurions plutôt affaire à un continuum entre des énoncés purement descriptifs et des jugements clairement moraux. Entre les deux, pourraient se trouver des énoncés de forme factuelle mais ayant une connotation axiologique induite par le contexte.

Si la question de la dichotomie fait-valeur m'est apparue importante, c'est que si les valeurs sont totalement séparées du monde factuel et qu'on leur refuse des conditions de vérité, tous nos débats moraux semblent être une agitation inutile et vaine. Les discussions sur des questions morales seraient du même ordre que des disputes sur des questions de goûts. Or il me semble que les débats moraux ont une importance et qu'il y a bien quelque chose à propos de quoi porte le débat. Dans ce travail, j'ai argumenté en faveur du réalisme moral car je pense que nos énoncés moraux sont de véritables croyances et qu'ils réfèrent donc à quelque chose dans le monde. Je suis d'avis que les propriétés morales surviennent sur des propriétés de base naturelles. Si l'on pense que les propriétés morales ont un statut ontologique suspect, pourquoi n'en irait-il pas de même pour les autres propriétés survenantes, comme les propriétés biologiques ou mentales par exemple, pourtant généralement acceptées en tant que propriétés réelles? Le monde est constitué de couches différentes et les faits et les valeurs ne se situent pas dans les mêmes strates, puisque les faits moraux sont constitués par des faits naturels. La distinction entre propriétés de base et propriétés survenantes me semble suffisante. À partir du moment où nous comprenons que les propriétés mentales sont réalisées par des états neuronaux, nous sommes amenés à rejeter le dualisme entre physique et mental. Si les faits moraux surviennent sur des faits naturels, je ne vois pas pourquoi nous devrions, alors, conserver une fracture ontologique profonde entre les deux.

Bibliographie

- Ayer A.J., *Langage, Vérité et Logique*, trad. J. Chana, Paris, Flammarion, 1956
- Blackburn S., *Spreading the Word*, Oxford, Clarendon Press, 1984
- Blackburn S., *Essays in Quasi-Realism*, Oxford, Oxford University Press, 1993
- Blackburn S., "Gibbard on Normative Logic", *Philosophical Issues*, Vol.4, Naturalism and Normativity, (1993b), pp.60-66
- Blackburn S., *Ruling Passions*, Oxford, Clarendon Press, 1998
- Brink D., "Externalist Moral Realism", *The Southern Journal of Philosophy*, Vol.24, Supplement, (1986), pp.23-41
- Brink D., *Moral Realism and the Foundations of Ethics*, Cambridge University Press, 1989
- Brink D., "Moral Motivation", *Ethics*, Vol.108, N°1 (Oct. 1997), pp.4-32
- Brink D., "Realism, Naturalism, and Moral Semantics", *Social Philosophy and Policy*, Vol.18, N°2, (Summer 2001), pp.154- 176
- Damasio A., *L'erreur de Descartes. La raison des émotions*, trad. M. Blanc, Paris, Éditions Odile Jacob, 1994, 1995
- Dreier J., "The Supervenience Argument Against Moral Realism", *The Southern Journal of Philosophy*, Vol.30, N°3, (1992), pp.13-38
- Dreier J., "Dispositions and Fetishes: Externalist Models of Moral Motivation", *Philosophy and Phenomenological Research*, Vol.61, N°3, (Nov. 2000) pp. 619-638
- Geach P.T., "Assertion", *The Philosophical Review*, Vol. 74, N°4 (Oct 1965), pp.449-465
- Gibbard A., *Wise Choices, Apt Feelings. A Theory of Normative Judgment*, Cambridge, Mass., Harvard University Press, 1990
- Hare R.M. *The Language of Morals*, Oxford, Oxford University Press, 1952
- Harman G., "Moral Explanations of Natural Facts- Can Moral Claims Be Tested Against Moral Reality?", *The Southern Journal of Philosophy*, Vol. 24, Supplement, (1986), pp.57-67
- Harman G., "Ethics and Observation" in G. Sayre.McCord (ed.), *Essays in Moral Realism*, Ithaca and London, Cornell University Press, 1988
- Harman G. & Thomson J.J., *Moral Relativism and Moral Objectivity*, Cambridge, Mass., Oxford, Blackwell Publishers, 1996
- Horgan T. & Timmons M., "Troubles on Moral Twin Earth: Moral Queerness Revived", *Synthese*, Vol.92, N°2, (Aug 1992), pp.221-260
- Kim J., *Philosophy of Mind*, Boulder, Oxford, Westview Press, 1998

- Klagge J., "An Alleged Difficulty Concerning Moral Properties", *Mind*, New Series, Vol.93, N°371 (Jul. 1984), pp.370-380
- Kraemer E.R., "On the Moral Twin Earth Challenge to New-Wave Moral Realism", *Journal of Philosophical Research*, Vol.16, (Oct 1991), pp.466-472
- Mackie J.L., *Ethics. Inventing Right and Wrong*, London, Penguin Books, 1977, 1990
- Miller A., "Emotivism and the Verification Principle", *Proceedings of the Aristotelian Society*, New Series, Vol.98, (1998), pp.103-124
- Miller A., *An Introduction to Contemporary Metaethics*, Cambridge, Polity Press, 2003
- Moore G.E., *Principia Ethica*, Cambridge, Cambridge University Press, 1966
- Nagel T., *The Possibility of Altruism*, Princeton, Princeton University Press, 1970
- Ogien R., *Le réalisme moral*, Paris, Presses Universitaires de France, 1999
- Ogien R., *La panique morale*, Paris, Grasset, 2004
- Pettit P., "Humeans, Anti-Humeans, and Motivation" *Mind*, New Series, Vol.96, N°384, (Oct.1987), pp.530-33
- Putnam, H. *Mind, Language and Reality. Philosophical Papers Volume 2*, Cambridge, Cambridge University Press, 1975
- Putnam H., *Reason, Truth and History*, Cambridge, Cambridge University Press, 1981
- Putnam H., *Realism with a Human Face*, Cambridge Mass., London, Harvard University Press, 1990
- Putnam H., *The Collapse of the Fact/ Value Dichotomy and other essays*, Cambridge, Mass., London, Harvard University Press, 2002
- Quine W.v.O., "Main Trends in Recent Philosophy: Two Dogmas of Empiricism", *The Philosophical Review*, Vol. 60, N°1 (Jan. 1951), pp. 20-43
- Sayre.McCord G., "'Good' On Twin Earth", *Philosophical Issues*, Vol.8, Truth, (1997), pp.267-292
- Sayre.McCord G., "Moral Theory and Explanatory Impotence" in G. Sayre.McCord (ed.), *Essays in Moral Realism*, Ithaca and London, Cornell University Press, 1988
- Schueler G.F., "Modus Ponens and Moral Realism", *Ethics*, Vol. 98, N°3, (Apr. 1988), pp.492-500
- Shafer-Landau R., *Moral Realism. A Defence*, Oxford, Clarendon Press, 2003

- Searle J., "Meaning and Speech Acts", *The Philosophical Review*, Vol.71, N°4 (Oct. 1962), pp. 423-432
- Searle J., "How to Derive "Ought" From "Is", *The Philosophical Review*, Vol. 73, N°1 (Jan. 1964), pp.43-58
- Smith M., "The Humean Theory of Motivation", *Mind*, New Series, Vol.96, N°381, (Jan. 1987), pp.36-61
- Smith M., *The Moral Problem*, Cambridge, Mass., London, Blackwell,, 1994
- Sturgeon N., "Harman on Moral Explanations of Natural Facts", *The Southern Journal of Philosophy*, Vol. 24, Supplement, (1986), pp.69-78
- Sturgeon N., "Moral Explanations" in G. Sayre.McCord (ed.), *Essays in Moral Realism*, Ithaca and London, Cornell University Press, 1988
- Svavarsdottir S., "Moral Cognitivism and Motivation", *The Philosophical Review*, Vol.108, N° 2 (Apr. 1999), pp. 161-219
- Timmons Mark, *Morality without Foundations. A Defense of Ethical Contextualism*, New-York, Oxford, Oxford University Press, 1999
- Van Roojen M., "Expressivism and Irrationality", *The Philosophical Review*, Vol. 105, N°3 (Jul. 1996)
- Williams B., *Ethics and the Limits of Philosophy*, London, Fontana, 1985
- Wittgenstein L., *Tractatus logico-philosophicus* suivi de *Investigations philosophiques*, trad. P. Klossowski, Paris, Gallimard, 1961
- Zagwill N., "Explaining Supervenience: Moral and Mental", *Journal of Philosophical Research*, Vol.22, (Apr 1997), pp.509-518

