

Une défense du moralisme

Contre l'égoïsme méthodologique et l'adaptation
des normes morales aux motivations humaines

Thèse présentée à la Faculté des Lettres et Sciences Humaines

Institut de philosophie

Université de Neuchâtel

Pour l'obtention du grade de docteur en sciences humaines et sociales
par

Judith Würgler

Acceptée par le jury de thèse :

Prof. Richard Glauser, Université de Neuchâtel, directeur de thèse

Prof. Christine Clavien, Université de Genève

Prof. Marc Fleurbaey, Université de Princeton, USA

Prof. Laurent Jaffro, Université Paris I Panthéon-Sorbonne, France

Prof. Mark Nelson, Westmont College, Santa Barbara, USA

Soutenue le 11 juin 2020

IMPRIMATUR

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de M. Richard Glauser, directeur de thèse, professeur émérite, Université de Neuchâtel ; Mme Christine Clavien, MER, Université de Genève ; M. Marc Fleurbaey, professeur, Université de Princeton ; M. Laurent Jaffro, professeur, Université Paris-I Panthéon-Sorbonne ; M. Marc Nelson, professeur, Westmont College, Santa Barbara autorise l'impression de la thèse présentée par Mme Judith Würigler en laissant à l'autrice la responsabilité des opinions énoncées.

Neuchâtel, le 11 juin 2020



Le doyen
Pierre Alain Mariaux

Résumé

Ce travail soutient une variante *extrême* de moralisme éthique et politique. Cette variante est *extrême* car elle soutient simultanément les trois composantes constitutives du moralisme. Une variante plus modérée pourrait ne soutenir que l'une ou l'autre de ces composantes. Ces trois composantes sont les suivantes :

1. Il est raisonnable d'attendre de citoyens informés qu'ils adaptent leur comportement en fonction des exigences morales qui portent sur eux, même lorsque cela requiert pour eux de sacrifier une partie de leur intérêt personnel.
2. Il n'existe pas de *limites*, fondées sur des considérations au sujet des motivations humaines, à ce qui peut être exigé des individus sur le plan moral et politique.
3. La recherche en philosophie morale est *indépendante* et *prioritaire* relativement à la recherche empirique au sujet des motivations humaines. Seule la recherche normative – et non la recherche empirique – peut nous renseigner sur la nature et le contenu de la motivation morale.

L'attitude moraliste est décriée en philosophie par des critiques de différents bords. Le but premier des pages qui suivent consiste à défendre le moralisme contre ces attaques.

La première critique, fondée sur l'accusation de « sophisme moraliste », est formulée par les partisans du *réalisme politique*, auquel appartiennent de nombreux économistes. Ils accusent les théories politiques moralistes de reposer sur une conception (trop) optimiste et naïve de la nature humaine. Selon cette accusation, au lieu de s'attendre à ce que les êtres humains renoncent spontanément à leurs intérêts privés au nom du bien public ou de la morale, il faut plutôt « faire l'économie de la vertu » dans l'organisation de la société ; le but est d'obtenir un comportement extérieurement vertueux sans faire appel à des motivations vertueuses – notamment par la mise en place d'incitations monétaires.

La seconde critique, fondée sur l'accusation d'« erreur moraliste », est formulée par les partisans de l'« humanisme moral ». Ces derniers accusent les théories éthiques moralistes de ne pas tenir compte de la nature humaine et de ses défaillances motivationnelles dans la *formulation* des principes moraux eux-mêmes. Selon cette accusation, une théorie morale qui est incompatible avec les motivations humaines, par exemple en exigeant des sacrifices individuels trop importants, ne respecte pas le principe « devoir-implique-pouvoir » et doit donc être rejetée.

La thèse consiste avant tout à réfuter à la fois l'accusation de « sophisme moraliste » et celle d'« erreur moraliste ». La dernière partie du travail consiste à défendre la troisième composante de la thèse moraliste, i.e. la priorité de la recherche morale normative par rapport à la recherche empirique. Il s'agit, dans cette partie, de réhabiliter le raisonnement dit « moraliste », souvent condamné comme fallacieux, consistant à dériver une conclusion *descriptive* sur la base d'une prémisse exclusivement *normative*. Par la réhabilitation de ce raisonnement, nous serons en mesure de tirer la conclusion *descriptive* suivante: il existe, chez l'être humain, une motivation morale suffisamment forte pour surmonter tous les obstacles motivationnels qui pourraient s'opposer à la réalisation de l'action morale.

Mots-clés

Altruisme / Egoïsme (méthodologique) / Motivation morale / Amoralisme / Internalisme / Science économique / Éthique des défaillances des marchés / Métaéthique / « Devoir-implique-pouvoir » / Sophisme et erreur naturalistes / Sophisme et erreur moralistes / Humanisme / Logique déontique

Abstract

The thesis defends an *extreme* version of ethical and political moralism. It is *extreme* because it simultaneously upholds all the three constituent claims of moralism. A milder version would support only one or two of them. The three constituent claims of moralism are the following:

1. It is reasonable to expect informed citizens to adapt their behavior according to the moral requirements to which they are subject, even when that implies sacrificing part of their personal interest.
2. There are no *limits*, grounded on considerations about human motivations, to what can be morally or politically required of individuals.
3. Research in moral philosophy is *independent* and has priority over empirical researches on human motivations. Normative inquiries alone – and not empirical investigations – can inform us about the nature and content of moral motivation.

This moralistic attitude has been criticized in philosophy from different perspectives. The primary aim of the present work is to argue against these criticisms.

The first criticism is formulated by advocates of “political realism”, among whom figure many economists. It claims that moralistic theories commit a “moral fallacy”. According to this line of criticism, moralistic political theories rely on an (overly) optimistic and naïve conception of human nature. According to this charge, rather than expecting human beings to spontaneously give up personal advantages for the sake of the public good or for the sake of morality, one should rather “economize on virtue” in the organization of social institutions; the aim is to make citizens act in an outwardly virtuous manner without relying on virtuous motivations – most notably, by organizing monetary incentives appropriately.

The second criticism is voiced by supporters of “moral humanism”. It claims that moralistic ethical theories fail to take account of human nature and its motivational deficiencies in the formulation of moral principles. By failing to do so, they commit what can be called a “moral error”. According to this line of criticism, a moral theory that is incompatible with human motivations, by requiring, for example, too great personal sacrifices, fails to conform to the principle that “ought-implies-can” and must, therefore, be rejected.

The present PhD thesis consists primarily in refuting the two charges of “moral fallacy” and “moral error”. The last part of the work is a defense of the third component of the moralistic thesis, i.e. the priority of normative research over empirical research. The aim of this part is to rehabilitate a form of “moralistic reasoning”, often condemned as fallacious. This kind of reasoning derives a *descriptive conclusion* on the basis of a *normative* premise. Through the rehabilitation of this “moralistic” way of thinking, we will be able to derive the following conclusion: human beings are endowed with a moral motivation that is sufficiently strong to enable them to overcome any motivational impediment in the realization of moral action.

Keywords

Altruism / (Methodological) egoism / Moral motivation/ Amoralism / Internalism / Economic science/ Market failure approach to business ethics / Meta-ethics / “Ought-Implies-Can” / Naturalistic fallacy and mistake / Moralistic fallacy and mistake / Humanism / Deontic Logic

Remerciements

J'adresse mes remerciements tout d'abord à mon directeur de thèse, Richard Glauser, désormais professeur émérite de philosophie générale à l'Université de Neuchâtel. Je lui suis reconnaissante de m'avoir accordé sa confiance pour entreprendre un projet de recherche en dehors de ses thèmes de prédilection et d'avoir relu mes ébauches avec tout le soin et le haut niveau d'exigence qui le caractérisent.

Je tiens également à remercier les membres de mon jury de thèse pour le temps et l'énergie consacrés à lire et commenter ce long travail. Il va sans dire que la seconde version de la thèse a largement profité de leurs conseils et suggestions. Ce fut ainsi un honneur de pouvoir bénéficier de l'expertise de Christine Clavien (MER, Université de Genève), Marc Fleurbaey (professeur, Université de Princeton), Laurent Jaffro (professeur, Université Paris-I Panthéon-Sorbonne) et Mark Nelson (professeur, Westmont College, Santa Barbara).

J'adresse également ma gratitude au Fonds National Suisse de la recherche pour son soutien financier dans le cadre d'un projet de recherche à l'Université de Francfort-le-Main. Ce séjour a permis de rediriger et recentrer ma recherche, tout en travaillant dans un contexte intellectuellement très riche et stimulant.

Mes collègues Antoine, Antoine, Antoine, Claudio, Lauréline, Léa, Maria, Vincent, ont formé une équipe imbattable pour élaborer des critiques, des objections et des recommandations lors de nos discussions de groupe et durant les fameux colloques du mardi. Olivier Massin, actuel professeur de philosophie générale à l'Université de Neuchâtel, a aussi largement contribué à la conclusion de ce travail, en acceptant de me prendre comme assistante et en ne désespérant pas (trop) devant mon envie de défendre les causes perdues – notamment les rats-taupes.

Pour finir, je tiens à remercier mes fidèles relecteurs et relectrices, non seulement pour avoir accepté de passer en revue des textes difficiles dans un domaine qui leur est inconnu, mais aussi pour avoir prêté leurs noms aux personnages qui animent le propos de la thèse. Merci ainsi aux membres de ma famille, Thomas, Lena, Gabrielle et Felix, à mon rayonnant conjoint Gilles et à mes éternel-le-s ami-e-s Ludivine, Nathalie et Gilles. Le plaisir de leur compagnie a seul permis de surmonter les phases difficiles de ce marathon doctoral.

TABLE DES MATIÈRES

PREMIÈRE PARTIE: ÉLÉMENTS FONDAMENTAUX

INTRODUCTION	17
1. LA QUESTION DE RECHERCHE	17
2. STRUCTURE DU TRAVAIL	19
3. QUELQUES DÉFINITIONS	22
3.1. ÉNONCÉS DESCRIPTIFS ET NORMATIFS	22
3.2. LA VARIÉTÉ DES ÉNONCÉS NORMATIFS	25
3.3. LE MAL MORAL	26
LES MOTIVATIONS HUMAINES	29
1. EGOÏSME, ALTRUISME ET MOTIVATION MORALE	29
1.1. L'ÉGOÏSME	29
1.1.1. Une conception générale de l'égoïsme	29
1.1.2. L'égoïsme économique	32
1.2. L'ALTRUISME	35
1.2.1. Une conception générale de l'altruisme	35
1.2.2. L'altruisme économique	37
1.3. LA MOTIVATION MORALE	37
2. LA MOTIVATION MORALE	40
2.1. COMPATIBILITÉ AVEC L'ENSEMBLE DES DOCTRINES MÉTAÉTHIQUES ET NORMATIVES	40
2.1.1. Compatibilité avec le réalisme moral fort non naturaliste	40
2.1.2. Compatibilité avec le réalisme moral naturaliste	43
2.1.3. Compatibilité avec le réalisme moral <i>faible</i>	45
2.1.4. Compatibilité avec les antiréalismes	46
2.1.5. Compatibilité avec l'ensemble des théories morales normatives	47
2.2. LA DIFFICULTÉ DU CHOIX DU FORMALISME	48
2.2.1. Un « goût » pour la morale ?	49
2.2.2. Une motivation morale égalitariste, kantienne ou utilitariste?	50
2.3. DES MOTIVATIONS PROCHES MAIS DIFFÉRENTES	52
2.3.1. L'effet du « sentiment chaleureux »	52

2.3.2. Le désir d'avoir ou de susciter une image positive de soi	55
3. LE DÉSIR DE RÉPUTATION MORALE	56
3.1. UN DÉSIR IDENTIQUE À LA MOTIVATION MORALE	56
3.2. UN DÉSIR DIFFÉRENT DE LA MOTIVATION MORALE	57
3.2.1. La variété des désirs liés à l'image morale de soi	59
3.2.2. La particularité des désirs liés à l'image morale de soi	60
3.2.3. Le désir d'être une personne vertueuse et le désir d'avoir une image positive de soi	63
3.2.4. Note sur la conception kantienne de la motivation morale	64
4. CONCLUSION	68
LA MOTIVATION MORALE EN MÉTAÉTHIQUE	71
1. LE PROBLÈME DE LA MOTIVATION MORALE DANS LA MÉTAÉTHIQUE CONTEMPORAINE	71
1.1. LE CONSTAT DE FAIT	71
1.2. L'EXPLICATION INTERNALISTE	73
1.2.1. Les conséquences métaéthiques de l'explication internaliste	74
1.2.2. Les contre-exemples	77
1.2.3. L'objection externaliste	81
1.3. UN CONFLIT ENLISÉ	83
1.4. L'ORIGINE PROBABLE DE L'ENLISEMENT	86
1.5. L'OBJET DE LA CONTROVERSE COMPARÉ À L'OBJET DU PROBLÈME QUI NOUS INTÉRESSE	88

DEUXIÈME PARTIE: LE SOPHISME MORALISTE

L'ACCUSATION DE SOPHISME MORALISTE EN ÉCONOMIE	93
1. L'HYPOTHÈSE ÉGOÏSTE EN ÉCONOMIE	93
1.1. L'ÉGOÏSME MÉTHODOLOGIQUE	95
1.2. LE SOPHISME MORALISTE SELON LES ÉCONOMISTES	99
1.3. L'ÉCONOMIE COMME SCIENCE DE « MÉCANISATION DE LA VERTU »	102
2. LES CONSÉQUENCES SOCIALEMENT DANGEREUSES DE L'HYPOTHÈSE DE L'ÉGOÏSME	106
2.1. L'EFFET CONTREPRODUCTIF DES INCITATIFS ÉCONOMIQUES	106
2.1.1. L'achat d'un droit de faire des actions immorales	107
2.1.2. La diminution des actions vertueuses : le cas du Don de sang	109
2.2. LA DISPARITION DES COMPÉTENCES DE DISCERNEMENT MORAL	112

2.2. DEUX FORMES D'ÉGOÏSME	186
DÉDRAMATISER LE CONFLIT ENTRE MORALE ET INTÉRÊT PERSONNEL	189
1. DEUX ARGUMENTS	189
2. L'ARGUMENT DE « DÉDRAMATISATION DE L'ÉGOÏSME »	191
2.1. LES COÛTS DE LA POSTURE DE L'AMORALISTE	194
2.1.1. Les coûts psychologiques	194
2.1.2. L'opposition à nos dispositions affectives	196
2.1.3. L'impossibilité de maîtriser les effets de notre action	197
2.1.4. Le risque de désintégration sociale	198
2.1.5. Notre disposition naturelle au bien	198
2.2. CRITIQUE DE L'ARGUMENT DES COÛTS DE L'AMORALISME	199
2.3. LE PASSAGE EN DOUCEUR DE L'AMORALISME À LA MORALE	201
2.4. L'ENTRÉE RAPIDE DANS LA SPHÈRE DE LA MORALE	204
2.4.1. L'impossibilité de la justification morale de la stratégie égoïste	204
2.4.2. L'impossibilité de la révolte morale	205
2.4.3. L'impossible supériorité morale de l'amoraliste	206
3. CONCLUSION	207
L'ERREUR MORALISTE	211
1. L'ERREUR MORALISTE	211
1.1. LA CONTRAINTE HUMANISTE	212
1.2. LES THÉORIES MORALES QUI ÉCHOUENT À RESPECTER LA CONTRAINTE HUMANISTE	214
1.3. LE PROBLÈME DU SOPHISME NATURALISTE	220
1.3.1. Le sophisme naturaliste sur le plan logique	221
1.3.2. Les théories affectées par l'accusation de « sophisme naturaliste »	223
1.3.3. La contestation logique du sophisme naturaliste	225
1.3.4. Le raisonnement naturaliste au niveau sémantique	227
1.3.5. Différents « Principes Sémantiques de Transition » (PST)	228
1.4. LA CONTRAINTE HUMANISTE COMME APPLICATION DU PRINCIPE « DEVOIR-IMPLIQUE-POUVOIR »	233
1.5. LA DÉFENSE TRADITIONNELLE DE L'INTERNALISME MORAL	238
1.5.1. L'internalisme des raisons d'agir	238
1.5.2. L'argument de la motivation	240
1.5.3. Critiques de l'argument de la motivation	242

1.5.4. « Devoir-implique-pouvoir » en faveur de l'internalisme des raisons d'agir	246
1.6. LE PRINCIPE DIP-MOTIVATION	248
LE PRINCIPE « DEVOIR-IMPLIQUE-POUVOIR »	251
1. LES VARIÉTÉS DE POSSIBILITÉS	252
2. L'ARGUMENT SÉMANTIQUE EN FAVEUR DU PRINCIPE DIP	256
2.1. LA SÉMANTIQUE MINIMALE DE LA NOTION DE « DEVOIR »	257
2.2. LES DEUX SÉMANTIQUES DU DEVOIR MORAL	259
2.2.1. La première sémantique de l'obligation morale	260
2.2.2. La seconde sémantique de l'obligation morale	264
3. LES OBJECTIONS CONTRE L'ANALYSE CONCEPTUELLE DE DIP	267
3.1. LES CONTRE-EXEMPLES	268
3.2. L'IMPOSSIBILITÉ DES DILEMMES MORAUX ET DE LA FAIBLESSE DE LA VOLONTÉ	271
3.2.1. Les dilemmes moraux	272
3.2.2. La faiblesse de la volonté	274
3.3. LA VIOLATION DE LA LOI DE HUME	275
3.4. L'ABSENCE DE NEUTRALITÉ NORMATIVE DE LA LOGIQUE DÉONTIQUE	276
4. QUELLE INTERPRÉTATION DE DIP ?	281
4.1. LA PRÉSUPPOSITION	281
4.2. L'INEFFICACITÉ DU DISCOURS MORAL	283
4.3. L'IMMORALITÉ DES PRINCIPES MORAUX QU'IL EST IMPOSSIBLE DE RESPECTER	287
4.3.1. Le respect de la loi de Hume	289
4.3.2. Le débat initial : péché originel, prédestination et la cruauté de Dieu	290
4.4. CRITIQUE DE L'INTERPRÉTATION MORALE	293
4.4.1. Le dilemme moral du réalisme métaéthique	294
4.4.2. Les jugements moraux qui ne sont pas avantageux pour la reproduction des gènes	298
5. LA VACUITÉ DU PRINCIPE « DEVOIR-IMPLIQUE-POUVOIR »	302
5.1. UNE RELATION ÉTROITE ENTRE LE BIEN NON MORAL ET LA MOTIVATION	304
5.2. DEUX FORMES D'INTERNALISME AU SUJET DU BIEN PERSONNEL	307
5.2.1. Deux conceptions du « bien pour »	307
5.2.2. Deux variétés d'internalisme du « bien pour »	309
5.3. LE BIEN PERSONNEL COMME OBJET POSSIBLE DE SA MOTIVATION OU DE SON SOUCI	312
5.3.1. Le bien personnel mis en relation avec la motivation par le biais de DIP	312

5.3.2. La faiblesse de l'internalisme du bien personnel	315
5.3.3. Le bien des chats et des pierres	316
5.4. PREMIÈRE CRITIQUE : COMMENT DÉTERMINER L'INCAPACITÉ ?	319
5.5. SECONDE CRITIQUE : L'IMPOSSIBILITÉ DE VOULOIR RENONCER À SON PROPRE BIEN	322
5.5.1. Deuxième version de la critique : désirer ce que l'on croit être son propre bien	324
5.5.2. Interprétation conceptuelle de DIP-motivation	326
5.5.3. Interprétation pragmatique de DIP-motivation	328
5.5.4. Interprétation morale de DIP-motivation	331
5.6. TROISIÈME CRITIQUE : LES LIMITES BIOLOGIQUES AUX EXIGENCES SOCIALES	332
5.6.1. Des individus non humains	337
5.6.2. Ce que les êtres humains devraient désirer, mais ne désirent pas nécessairement	340
5.6.3. La valeur morale des limites motivationnelles pour la justification	342
5.6.4. La légitimité du blâme	346
6. CONCLUSION	352

QUATRIÈME PARTIE: LE RAISONNEMENT MORALISTE RÉHABILITÉ

UNE DÉFENSE DU MORALISME	355
1. LE SOPHISME MORALISTE	355
1.1. QUEL LIEN ENTRE LES ÉNONCÉS DESCRIPTIFS AU SUJET DES MOTIVATIONS HUMAINES ET LES ÉNONCÉS MORAUX ?	356
1.2. QU'EST-CE QUE LE SOPHISME MORALISTE ?	360
1.2.1. Quelques exemples	361
1.2.2. Le sophisme moraliste motivationnel	367
1.2.3. Le caractère suspect de la motivation morale	370
1.2.4. Réhabiliter le raisonnement moraliste	375
1.3. LA VERSION MORALISTE DE « DEVOIR-IMPLIQUE-POUVOIR »	380
1.3.1. Le problème des dilemmes moraux	382
1.3.2. Le problème de la faiblesse de la volonté	392
1.3.3. Que faire de Luther ?	395
2. ET SI LES DEVOIRS MORAUX CATÉGORIQUES N'EXISTAIENT PAS ?	397
3. LE PROBLÈME FINAL : L'IMPOSSIBILITÉ DE MOTIVATION SANS MOTIVATION	402
3.1. L'OBJECTION	402
3.2. LE DILEMME	404

4. PEUT-ON COMPLÈTEMENT IGNORER LES DIFFICULTÉS MOTIVATIONNELLES DES ÊTRES HUMAINS ?	406
5. CONCLUSION	412
CONCLUSION	413
ANNEXES	419
1. LA MÉTAÉTHIQUE	419
1.1. LES DIFFÉRENTES POSITIONS DU DÉBAT MÉTAÉTHIQUE CONTEMPORAIN	420
1.1.1. Cognitivisme vs non-cognitivisme	420
1.1.2. Réalisme vs théorie de l'erreur	425
1.1.3. Objectivisme vs non-objectivisme	427
1.1.4. Réalisme moral naturaliste vs non naturalisme	429
1.1.5. Naturalisme moral analytique vs synthétique	433
1.1.6. Naturalisme moral synthétique <i>réductionniste</i> vs non réductionniste	434
1.1.7. Tableau des doctrines métaéthiques	436
2. LES NÉGATIONS DE NORMES SONT-ELLES DES NORMES ?	438
2.1. L'INTER-DÉFINISSABILITÉ DES NOTIONS NORMATIVES ET LES NÉGATIONS DE NÉGATION	438
2.2. LA COMPATIBILITÉ AVEC LE NIHILISME MORAL	441
2.3. CONCLUSION	443
3. LEXIQUE	444
BIBLIOGRAPHIE	447

PREMIÈRE PARTIE

Éléments fondamentaux

INTRODUCTION

1. LA QUESTION DE RECHERCHE

Dans ce travail, nous allons explorer différentes manières de représenter les êtres humains dans les théories politiques. Nous allons, entre autres, nous poser les questions suivantes : devons-nous représenter les êtres humains comme étant des êtres purement égoïstes ? Devons-nous, au contraire, concevoir une société où l'égoïsme se trouve complètement éliminé ? Par ailleurs, une société composée d'individus altruistes est-elle vraiment meilleure qu'une société composée d'individus égoïstes ? L'égoïsme constitue-t-il l'origine de tous les maux sociaux ? L'égoïsme est-il la ressource motivationnelle fondamentale des sociétés capitalistes et, de ce fait, condamne-t-il ces sociétés à un échec certain ? Peut-on raisonnablement attribuer aux citoyens d'une société une motivation morale désintéressée ou serait-ce faire preuve de naïveté ?

Nous allons tenter plus précisément de déterminer s'il est réaliste de faire reposer le système politique sur une conception des membres de la société comme étant capables de *vertu*, c'est-à-dire de se soucier de la valeur morale de leurs actions et, éventuellement, de sacrifier leurs intérêts personnels à cette valeur. La question que nous allons nous poser a été parfaitement formulée par Christopher Hamel :

[Une] politique de la vertu [est-elle] *réaliste*, eu égard aux présupposés psychologiques et anthropologiques censés la rendre possible ? Prétendre qu'elle est réaliste implique-t-il d'assumer l'idée d'une « vertu non motivée » (Pettit 2012, p. 247), c'est-à-dire désintéressée ? Est-il raisonnable d'attribuer aux individus des sociétés contemporaines la disposition à agir pour le bien commun et à sacrifier leurs intérêts personnels ? Par ailleurs, fût-elle réaliste, serait-elle *faisable* ? Peut-on la mettre en œuvre sans conférer à l'État un pouvoir excessif de modeler les dispositions des individus ? (HAMEL, 2015, p. 103)

Nous allons dans ce travail défendre la plausibilité d'une telle « politique de la vertu ». Nous soutiendrons, plus précisément, la (triple) thèse suivante :

1. Il est raisonnable d'attendre de citoyens informés qu'ils adaptent leur comportement en fonction des exigences morales qui portent sur eux, même lorsque cela requiert pour eux de sacrifier une partie de leur intérêt personnel.
2. Il n'existe pas de *limites*, fondées sur des considérations au sujet des motivations humaines, à ce qui peut être exigé des individus sur le plan moral et politique.
3. La recherche en philosophie morale est *indépendante* et *prioritaire* relativement à la recherche empirique au sujet des motivations humaines. Sur la base de cette indépendance et priorité, nous défendrons la thèse de l'existence, chez l'être humain, d'une motivation morale catégorique, universelle et toute-puissante, susceptible de remplacer l'intérêt personnel comme ressource motivationnelle fondamentale dans nos théories politiques.

Ces trois thèses réunies constituent ensemble une doctrine dite « moraliste ». Pour défendre cette doctrine, nous devons répondre à deux objections majeures : l'accusation de « sophisme moraliste » développée par les économistes et l'accusation d'« erreur moraliste » développée par les métaéthiciens.

L'accusation de « sophisme moraliste » concerne le premier point de notre thèse. Selon cette accusation, envisager que les membres d'une société disposent d'une motivation morale désintéressée, et baser une théorie politique sur cette idée, cela revient à considérer les hommes tels qu'ils devraient être – vertueux, coopératifs, altruistes, loyaux, honnêtes – plutôt que tels qu'ils sont réellement – philistins, égoïstes, manipulateurs, déloyaux, malhonnêtes, etc.

Selon les économistes, le sophisme moraliste est un raisonnement fallacieux récurrent parmi les planificateurs sociaux¹, par lequel ces derniers sont amenés à naïvement croire que les citoyens et acteurs économiques se conformeront spontanément aux exigences morales, même lorsque cela est contraire à leurs propres intérêts. Ce raisonnement fallacieux consiste à dériver un énoncé *descriptif* au sujet des motivations humaines – i.e. « les êtres humains sont vertueux, altruistes, coopératifs » – sur la base d'un énoncé *normatif* au sujet de la manière dont les êtres humains *devraient* se comporter moralement – i.e. « les êtres humains *devraient* être coopératifs, altruistes et vertueux ».

¹ Nous entendons par planificateur social toute personne chargée de l'organisation de la vie en société – que ce soit dans la fonction de législateur, de gouvernant ou de conseiller du gouvernement.

L'accusation d'« erreur moraliste », développée par les métaéthiciens, concerne le second point de notre thèse. Selon cette accusation, une théorie morale est déficiente, voire fausse, si elle ne tient pas compte de la réalité de la nature humaine – potentiellement défaillante – directement dans la *formulation* des principes moraux régissant la vie en société. Une erreur moraliste consiste, en ce sens, à formuler des principes qui ne sont pas « adaptés » à l'être humain et à ses motivations fondamentales.

L'objectif principal du présent travail consistera à répondre aux deux objections susmentionnées. A la fin du travail, nous esquisserons également un argument *positif* en faveur de l'attribution, aux êtres humains, d'une motivation morale désintéressée. Par cet argument, nous serons amenés, simultanément, à soutenir la thèse fondamentale suivante : la philosophie morale est indépendante, voire prioritaire, par rapport à la recherche empirique au sujet des motivations humaines, notamment en ce qui concerne la motivation morale.

2. STRUCTURE DU TRAVAIL

Le travail se partage en quatre parties principales. La première sert à introduire de nombreux éléments. Nous commencerons par poser quelques définitions de notions normatives qui nous seront utiles par la suite. Dans une partie assez longue, nous définirons ensuite quelle est la nature de la motivation morale et comment cette motivation se distingue d'autres types de motivation. Pour finir, nous exposerons la manière dont est traitée la « question de la motivation morale » dans la métaéthique contemporaine et justifierons notre choix de ne pas entrer dans ce débat, malgré les affinités évidentes avec notre question de recherche.

La seconde partie de la thèse expose et répond à l'objection de « sophisme moraliste » telle qu'elle est formulée par les économistes. Selon cette objection, une doctrine politique qui suppose que les individus sont capables de renoncer à une partie de leurs intérêts au nom de la morale est une doctrine qui manque de « réalisme » au sujet des motivations humaines. A l'inverse, une doctrine politique réaliste – non moraliste – ne suppose pas que les individus disposent d'une quelconque motivation morale. Elle fait plutôt « l'économie de la vertu » en supposant que les êtres humains sont motivés à agir exclusivement par intérêt personnel.

Pour répondre à l'accusation de « sophisme moraliste », notre attitude consistera principalement à contre-attaquer. En effet, nous montrerons que les théories politiques qui « font l'économie de la vertu » rencontrent des difficultés que ne rencontrent pas les théories politiques « moralistes » – i.e. les théories fondées sur les dispositions morales des citoyens. Ces difficultés se partagent en deux catégories.

En premier lieu, fonder une société sur l'intérêt personnel peut avoir des conséquences sociales néfastes. Nous distinguerons quatre de ces conséquences : (1) la disparition de la valeur morale des actions, (2) une baisse des compétences morales, (3) des coûts importants pour la collectivité et (4) un transfert de la responsabilité morale des individus vers l'État.

En second lieu, nous soutiendrons que faire l'hypothèse que les individus sont motivés à agir exclusivement par intérêt personnel ne permet pas, contrairement à ce que pensent les partisans de cette hypothèse, de répondre à l'exigence de « réalisme » à l'égard des motivations humaines. Nous montrerons ainsi que l'hypothèse de l'intérêt personnel est elle-même encore trop « utopiste » ou « angélique » à l'égard des motivations de l'être humain. A nouveau, cette partie se divisera en quatre critiques distinctes : (1) les motivations humaines destructrices, (2) la morale des politiciens, (3) l'égoïsme de l'être humain et, finalement, (4) la norme implicite du marché. Nous parviendrons à la conclusion que le projet de « faire l'économie de la vertu » dépend d'une conception erronée du réalisme politique.

La troisième partie de notre travail concerne la seconde objection, celle formulée par les métaéthiciens. Selon ces derniers, toute doctrine morale qui n'est pas adaptée aux motivations des êtres humains est une doctrine invalide ou déficiente. Elle commet une « erreur moraliste ». A l'aide du principe « devoir-implique-pouvoir », les métaéthiciens soutiennent ainsi qu'une doctrine morale ne peut contenir ni des normes que les êtres humains ne sont pas en mesure d'accomplir – sur le plan motivationnel, ni des normes qui requièrent des sacrifices trop importants.

Notre tâche, dans cette troisième partie, consistera à répondre à l'accusation d'« erreur moraliste ». Nous procéderons comme suit : d'abord, nous présenterons la « contrainte humaniste » censée justifier l'exclusion des doctrines morales trop difficiles à respecter pour les êtres humains. Nous donnerons quelques exemples des théories qui

seraient exclues si la contrainte humaniste était appliquée. Dans un second temps, nous discuterons du principe « devoir-implique-pouvoir » sur lequel repose la « contrainte humaniste ». Nous montrerons que ce principe peut recevoir plusieurs interprétations, qui toutes posent des difficultés. Nous montrerons ensuite que le principe rencontre une difficulté spécifique lorsqu'il porte sur les possibilités motivationnelles. Cette difficulté consiste dans le fait que le principe est *vide*, car il n'existe rien que l'être humain est essentiellement incapable de désirer. Nous répondrons à plusieurs objections qui peuvent être soulevées contre notre argument et conclurons qu'il n'existe pas de limites motivationnelles à ce qui peut être demandé d'un être humain sur le plan moral et politique.

La dernière partie du travail esquissera une défense de la thèse positive selon laquelle l'être humain possède une motivation morale indépendante de ses désirs non moraux. Pour ce faire, nous utiliserons un raisonnement qui présente, en apparence, les caractéristiques d'un « sophisme moraliste ». Un sophisme moraliste consiste à dériver fallacieusement une proposition descriptive sur la base d'une prémisse normative. Notre tâche consistera à montrer qu'il existe des circonstances dans lesquelles ce raisonnement moraliste n'est pas fallacieux, notamment lorsqu'il s'appuie sur le principe « devoir-implique-pouvoir », interprété comme un principe analytique. Nous montrerons que de nombreuses doctrines s'appuient en fait sur ce raisonnement moraliste. La principale conclusion de cette dernière partie consistera à soutenir que la philosophie morale est indépendante et *prioritaire* par rapport à la recherche empirique au sujet des motivations humaines. Nous aurons donc, finalement, poussé le moralisme à son extrême.²

² En effet, les philosophes sont décriés comme étant « moralistes » soit (a) parce qu'ils s'opposent au réalisme politique, soit (b) parce qu'ils s'opposent à l'humanisme moral, ou (c) parce qu'ils soutiennent l'indépendance de la philosophie morale par rapport à la recherche empirique. Cette troisième acceptation du « moralisme » est visible dans la citation d'Edward C. Moore suivante : « J'appellerai *sophisme moraliste* l'affirmation que les jugements moraux appartiennent à un ordre différent des jugements factuels. » [“What I would like to call the Moralistic Fallacy is the assertion that moral judgments are of a different order from factual judgments” (MOORE E. C., 1957, p. 29)]. De notre côté, nous soutiendrons les *trois* formes de « moralisme ».

3. QUELQUES DÉFINITIONS

Avant toute chose, il nous faut rapidement définir un certain nombre de concepts centraux de la métaéthique qui nous seront utiles pour la suite. Le premier est celui d'*énoncé moral*, qui doit être distingué des énoncés descriptifs et, surtout, des énoncés descriptifs *naturels*.

3.1. ÉNONCÉS DESCRIPTIFS ET NORMATIFS

La distinction la plus générale est celle qui oppose les énoncés descriptifs et les énoncés normatifs – dont font partie les énoncés moraux. Pour notre propos, il suffit de caractériser la distinction dans les termes très généraux dans lesquels la distinction est généralement formulée : les énoncés descriptifs concernent *ce qui est*, alors que les énoncés normatifs – et moraux – concernent *ce qui devrait être*. Les énoncés descriptifs ont pour fonction de représenter le monde alors que les énoncés normatifs ont pour fonction de l'évaluer. Ainsi, les énoncés « Jean se marie avec Jeanne », « Le Père-Noël a une barbe blanche », « Dieu a élu le peuple Juif » sont des énoncés descriptifs, alors que les énoncés « Jean ne devrait pas se marier avec Jeanne », « Le Père-Noël devrait se couper la barbe » et « Dieu ne devrait pas faire de choix arbitraires » sont des énoncés normatifs.

Intéressons-nous maintenant aux énoncés normatifs proprement moraux. Comment identifie-t-on un énoncé moral ? Un énoncé moral est tout d'abord un énoncé qui contient une notion morale – i.e. bien/mal, juste/injuste, vicieux/vertueux. Cela ne suffit cependant pas pour séparer les énoncés moraux et d'autres énoncés descriptifs contenant, eux-aussi, des notions morales. Par exemple, l'énoncé « Les Suisses croient qu'il est irrespectueux d'arriver en retard à un rendez-vous » est un énoncé descriptif tout à fait commun, bien qu'il contienne le terme « irrespectueux ». Pour ne pas devoir considérer cet énoncé comme étant un énoncé moral, nous devons préciser que, dans les énoncés moraux, le concept moral apparaît *dans un contexte non subordonné*. A l'aide de cette précision, il devient possible de ne pas considérer la description de croyances et d'attitudes morales comme étant elle-même un énoncé moral. Ainsi, nous suivons Jaquet et Naar pour dire que

[...] dans les énoncés « Emmanuel Kant était convaincu que mentir est immoral », « Peter Singer soutient que manger de la viande est injuste » et « La plupart des gens considèrent qu'il est mal de se réjouir de la souffrance d'autrui », les termes

« immoral », « injuste » et « mal » apparaissent dans des phrases subordonnées ». Ces énoncés ne sont donc pas des énoncés moraux selon notre définition. (JAQUET & NAAR, 2019, pp. 193-194)

Un autre critère pertinent, que nous empruntons à Arthur Prior, pour identifier un énoncé *moral*, est que les notions morales qui y figurent doivent influencer la valeur de vérité de l'énoncé. Autrement dit, à la suite de Prior, nous maintenons que les notions morales doivent apparaître dans l'énoncé de manière *essentielle* :

Il est [...] nécessaire [...] que l'une au moins des expressions qui est présente apparaisse de manière « essentielle », i.e. elle ne doit pas être remplaçable par n'importe quelle autre expression (du type grammatical approprié) sans que cela ne modifie la valeur de vérité de l'énoncé. Par exemple, « soit il est le cas, soit il n'est pas le cas que je dois me battre pour mon pays » n'est pas un énoncé éthique, car la valeur de vérité de l'énoncé dans son ensemble ne serait pas modifiée si la phrase éthique « je dois me battre » était remplacée par « je me bats fréquemment » ou « mes voisins croient que je me bats », ou si la clause totale « je dois me battre pour mon pays » était remplacée par « deux et deux font 5 » ou « le pain coûte six francs la livre ». ³

Comme le dit Prior, l'énoncé « soit il est le cas, soit il n'est pas le cas que je dois me battre pour mon pays » n'est pas un énoncé moral, car les notions morales peuvent être remplacées par n'importe quelle notion non morale sans influencer la valeur de vérité de l'énoncé complet. Par exemple, si nous remplaçons l'expression « je dois me battre » par « le pain coûte six francs la livre », alors nous obtenons la proposition complète « soit il est le cas, soit il n'est pas le cas que le pain coûte six francs la livre ». Or, cette proposition a exactement la même valeur de vérité que la proposition « soit il est le cas, soit il n'est pas le cas que je dois me battre ». Cela signifie que le contenu proprement *moral* de l'énoncé n'affecte pas la valeur de vérité de la proposition complète. Nous concluons donc que les énoncés qui contiennent des notions morales dans des contextes non subordonnés, mais

³ "It is also necessary, for example, that at least one of the ethical expressions which are present should occur "essentially", i.e., should not be just replaceable by any expression whatsoever (of the appropriate grammatical type) without change of truth-value. For example, "It either is or is not the case that I ought to fight for my country" is not an ethical statement, since the truth-value of the whole would be unaltered if the ethical phrase "ought to" were replaced by "frequently do" or "am believed by my neighbours to", or if the whole clause "I ought to fight for my country" were replaced by "two and two are five" or "bread is sixpence a loaf"." (PRIOR, 1960, p. 200)

dans lesquels ces notions morales ne jouent aucun rôle pour déterminer leur valeur de vérité, ne sont pas des énoncés moraux.

Nous ajouterons encore que l'énoncé moral ne doit pas contenir *uniquement* des termes moraux et logiques. Pour qu'un énoncé soit proprement normatif ou moral, il doit non seulement contenir des notions morales et des notions logiques, mais également des notions *descriptives non morales*. Sur ce point, nous suivons encore Arthur Prior : les énoncés « il est obligatoire de faire ce qu'il est obligatoire de faire » ou « ce qui est obligatoire est aussi permis » ou « il est interdit de faire quoi que ce soit de contraire au devoir » ne sont pas de véritable énoncés moraux, car ils ne font que définir les notions morales entre elles, sans faire porter ces notions *sur des objets du monde* (PRIOR, 1960, p. 201). Pour qu'un énoncé soit moral, il doit mettre en relation des notions morales et des notions descriptives non morales, comme c'est le cas dans les énoncés « Georges est bienveillant », « les distinctions sociales basées exclusivement sur l'appartenance à une race sont injustes » ou encore « les parents doivent prendre soin de leurs enfants ».

En résumé, l'énoncé « il est immoral de commettre un adultère » est un énoncé moral car (i) il contient une notion morale dans un contexte non subordonné, (ii) la notion morale apparaît de manière essentielle dans l'énoncé et (iii) l'énoncé ne contient pas exclusivement des notions morales et logiques, mais met en relation des concepts moraux (« immoral ») et non moraux (« adultère »).

Intéressons-nous pour finir aux énoncés naturels. Nous avons vu que les énoncés descriptifs, contrairement aux énoncés normatifs – dont les énoncés moraux constituent une sous-catégorie – portent sur *ce qui est*. Parmi ces énoncés *descriptifs*, il convient ensuite de distinguer plus spécifiquement les *énoncés naturels*. Quelle est la caractéristique spécifique des énoncés *naturels* parmi les énoncés descriptifs ?

Notons en premier lieu que les énoncés naturels sont des énoncés qui portent sur des *faits naturels*. Il n'existe toutefois pas de consensus au sujet de la nature des *faits naturels*, mais deux critères sont en général retenus : un fait naturel est (a) un fait qui est susceptible d'être étudié par les sciences empiriques « classiques » – par exemple la biologie, la psychologie, l'économie, la physique, la chimie, etc. et/ou (b) un fait qui joue un rôle explicatif non redondant pour rendre compte de phénomènes observés (JAQUET & NAAR, 2019, pp. 106-107). Dans ce contexte, les énoncés descriptifs « Dieu est omniscient » ou

« les sorcières ont des pouvoirs magiques » ne sont pas des énoncés *naturels*, car ils portent sur des faits qui ne sont pas susceptibles d'être étudiés par les sciences empiriques classiques et (nous supposons) ils ne jouent pas de rôle explicatif autonome pour rendre compte de phénomènes observés.

3.2. LA VARIÉTÉ DES ÉNONCÉS NORMATIFS

Il est possible de distinguer trois catégories principales de concepts normatifs (a) Les concepts déontiques, (b) les concepts évaluatifs *minces* et (c) les concepts évaluatifs *épais*.

(a) Les concepts déontiques

Les concepts « obligatoire », « interdit », « permis » et « facultatif », constituent les principales notions dites *déontiques*, c'est-à-dire relatives au *devoir*, auxquelles nous pouvons ajouter celles de « droit », de « juste » et de « raison d'agir ». La principale caractéristique des énoncés déontiques est qu'ils n'admettent pas de *degré* ou *gradation* : une action ne peut pas être, par exemple, plus ou moins obligatoire. De la même manière, je ne peux pas avoir plus ou moins droit à la vie.

Une autre caractéristique spécifique des concepts déontiques est qu'ils portent exclusivement sur des *actions*, contrairement aux concepts évaluatifs qui peuvent caractériser aussi bien des personnes, des actions, des choses ou des états de choses (TAPPOLET, 2011, p. 166).

Voici quelques exemples d'énoncés *déontiques* :

- « Il ne faut pas humilier autrui. »
- « Les pays riches ont le devoir moral de soulager la faim dans le monde. »
- « Tout le monde a le droit de s'exprimer librement. »
- « Il est injuste que les plus riches aient accès à la santé alors que les plus pauvres n'y ont pas accès. »
- « Le fait que la guerre cause beaucoup de souffrance est une raison de ne pas faire la guerre »

(b) Les concepts évaluatifs *minces*

Les notions de « bon » ou « mauvais », « bien » ou « mal » sont les principales notions évaluatives *minces*. Ces notions se distinguent des notions *déontiques* principalement par le

fait qu'elles admettent la *gradation* : il est possible de considérer un état du monde comme *plus ou moins bon ou mauvais*. En outre, ces notions ont pour caractéristiques – contrairement aux concepts évaluatifs *épais*, que nous présenterons au point (c) – de ne pas avoir de contenu *descriptif*. Elles constituent des notions *purement évaluatives*.

Parmi les exemples d'énoncés évaluatifs *minces* figurent par exemple les énoncés suivants :

- « La douleur est a priori un *mal* pour la personne qui en souffre. »
- « Il est *bon* que la guerre cesse dans le monde. »
- « L'introduction de cultures OGM en Suisse est une mauvaise chose. »
- « Le plaisir est meilleur que la douleur »

(c) Les concepts évaluatifs « épais »

Les notions servant à désigner des *vertus* – par exemple la générosité, la bienveillance, ... – et celles servant à désigner des *vices* – par exemple, l'orgueil, l'envie, ... – représentent des concepts évaluatifs *épais* par excellence. Leur particularité est que, à l'instar des concepts évaluatifs minces, ces notions admettent la *gradation*. Une personne peut être plus ou moins honnête, plus ou moins vertueuse. A l'inverse des notions évaluatives minces, par contre, les notions évaluatives *épaisses* présentent un contenu *descriptif* en plus de leur contenu évaluatif. Autrement dit, les notions évaluatives *épaisses* ne sont pas purement évaluatives.

Comme exemple d'énoncés évaluatifs « épais », nous trouvons :

- « Thomas est généreux et honnête. »
- « Le président est un lâche. »
- « La présidente est orgueilleuse. »
- « Felix est cruel avec les chats. »

3.3. LE MAL MORAL

Lorsque nous parlons du bien moral ou du mal moral, nous les distinguons d'autres types de *biens* et de *maux*. Il convient de brièvement caractériser ces différents types de

biens et de maux, afin de clarifier, dans les grandes lignes, la notion de « bien » et de « mal » que nous utiliserons par la suite.

Nous pouvons distinguer trois principales acceptations de la notion de « mal » (SCHULTHESS, 1986, p. 171) :

a) Le mal physique

Le mal physique qualifie la souffrance qui affecte le corps ou l'esprit. Sont des maux en ce sens la douleur, l'infirmité, la mort, la maladie, la folie. Les expressions « tu me fais mal » ou « malaise » font référence au mal dans son acceptation physique (SCHULTHESS, 1986, p. 171).

b) Le mal *instrumental*

Le mal instrumental désigne une déficience des *moyens* dans la réalisation d'une fin. Un moyen peut être *mauvais*, en ce sens qu'il est inadapté pour atteindre un objectif. Les expressions « un mauvais calcul », « je vois mal » ou « maladroit » font référence au mal dans son sens instrumental (SCHULTHESS, 1986, p. 172).

c) Le mal moral

La troisième signification du terme « mal » fait cette fois référence aux qualités morales des personnes et de leurs actions. Les expressions de « malveillance », « maltraitance » ou « malfaisance » désignent des défaillances morales des personnes (SCHULTHESS, 1986, p. 172). La norme en fonction de laquelle est évalué le *mal moral* n'est ni celle du bien-être (contrairement au mal physique), ni celle de l'adéquation des moyens aux fins (contrairement au mal instrumental), mais celui du respect de certaines *règles qui régissent* (ou pourraient régir) *les rapports entre les personnes*. Ces règles sont celles qui rendent possibles la vie commune. Le mal, dans son acceptation morale, est donc causé par la violation et la négligence des règles qui permettent (ou permettraient) la vie commune, mais aussi par l'indifférence à leur égard (SCHULTHESS, 1986, p. 177).⁴

⁴ Cette conception du mal moral, comme atteinte ou indifférence à des règles permettant la vie commune, peut sans doute paraître trop *engageante* du point de vue métaéthique et normatif. Il nous semble néanmoins que cette conception – ou des conceptions s'en rapprochant – sont suffisamment partagées pour offrir un point de départ à la réflexion, point de départ révisable s'il le faut. L'idée que la morale concerne des règles de vie commune et les possibilités d'accord ou d'entente entre les personnes, se retrouve aussi bien chez les constructivistes comme Jürgen Habermas (1999) et Karl-Otto Apel (1996-1998), John Rawls (2009) ou encore Gauthier (1986), chez les réalistes moraux tels que Larmore (2008) ou Scanlon (1998), et même chez

des expressivistes tels que Gibbard (1990). En outre, même les conceptions qui n'attribuent pas de place centrale à la possibilité de l'accord ou de l'entente dans leur théorie morale considèrent néanmoins, il nous semble, que le fait de vivre une vie morale fait partie des conditions permettant la vie en communauté. En effet, dans l'ensemble des théories morales, les règles morales sont, au minimum, des règles qui régissent les rapports entre les personnes.

LES MOTIVATIONS HUMAINES

1. EGOÏSME, ALTRUISME ET MOTIVATION MORALE

Commençons par introduire les différentes motivations dont il sera principalement question par la suite. Pour le moment, nous nous contenterons de trois motivations : (I) l'égoïsme, (II) l'altruisme et (III) la motivation morale. Nous proposerons une définition informelle et générale de chaque motivation, puis nous exposerons la manière dont les économistes définissent et formalisent ces motivations à l'intérieur de leur discipline.

1.1. L'ÉGOÏSME

1.1.1. UNE CONCEPTION GÉNÉRALE DE L'ÉGOÏSME

Un égoïste est un individu dont les actions sont motivées par la satisfaction de ses intérêts personnels. Il cherche à réaliser son propre bien. Un égoïste ne se soucie du bien d'autrui ou de morale que lorsque cela lui permet de mieux satisfaire ses propres intérêts. En cas de *conflit* entre, d'un côté, la satisfaction de ses propres intérêts et, de l'autre, la satisfaction des intérêts d'autrui ou la réalisation de valeurs morales, un égoïste sacrifiera les intérêts d'autrui et les valeurs morales à son propre profit. Nous pouvons donc définir l'égoïsme de la manière suivante :

Définition de l'égoïsme	L'acteur A est égoïste seulement si, par son action, il vise à satisfaire ses propres intérêts, à réaliser son propre bien.

Cette caractérisation de l'égoïsme comporte deux éléments qu'il importe de souligner. En premier lieu, cette définition de l'égoïsme est incompatible avec la conception *biologique* de l'égoïsme, définie en termes exclusivement comportementaux, indépendamment de l'attitude subjective. Elle est également incompatible avec la conception de l'égoïsme comme étant un processus causal uniquement inconscient. En effet, selon notre définition de l'égoïsme, il est nécessaire que *subjectivement* un individu poursuive son propre bien pour que cet individu puisse être adéquatement caractérisé d'égoïste.

Le second élément à noter au sujet de la définition proposée est qu'elle ne formule qu'une condition *nécessaire* pour qu'un individu soit égoïste, et non une condition *nécessaire et suffisante*. Ainsi, si nous imaginons un individu qui vise (subjectivement) à satisfaire ses propres intérêts, mais qui (objectivement) est mû par le souci du bien d'autrui, alors cet individu ne serait pas forcément considéré comme égoïste selon notre définition. Par exemple, un mécène peut faire don de sa fortune à une œuvre de charité dans le but (subjectif) de marquer l'histoire de son empreinte. Cependant, il s'avère que son véritable motif (inconscient) est de venir en aide à ses concitoyens. Ainsi, bien que sa motivation subjective soit égoïste, il est causalement poussé à agir par un motif altruiste, dont il n'a pas conscience. Notre définition n'affirme pas qu'un tel individu soit égoïste. La question reste ouverte.⁵

Il existe différentes manières de caractériser l'égoïsme. Une caractérisation assez générale de l'égoïsme est celle proposée par Bernard Williams. Selon le philosophe, nous pouvons caractériser une personne égoïste par le fait que ses désirs, portant sur des états de choses, sont déterminés exclusivement par la manière dont sa propre situation est affectée dans ces états de choses.

⁵ Il nous a été objecté que, en définissant l'égoïsme de la sorte, nous rompons avec la discussion historique de l'égoïsme, telle que formulée notamment par Bernard Mandeville. Nous ne pensons toutefois pas que notre conception de l'égoïsme nous écarte de la discussion historique. En effet, si Mandeville soutient bien que les comportements *aux apparences* altruistes ou vertueuses sont *en fait* le résultat de motifs égoïstes, il n'affirme pas que ces motivations égoïstes doivent toujours être inconscientes. Il existe certes, selon lui, un fossé entre les motivations *déclarées* (*publiques*) et les véritables motivations (*privées*), mais ce fossé ne se situe pas entre les motivations *conscientes* (*subjectives*) et les motivations *inconscientes* (*objectives*). Le caractère conscient de la poursuite de l'intérêt personnel chez Mandeville est évident, par exemple, dans la manière dont l'auteur explique l'origine de l'échange, fondé la simulation de l'amitié pour obtenir un prix à son avantage (MANDEVILLE, 1998 (1714), pp. 57-58), et l'origine de la politesse, fondée sur l'expérience du caractère avantageux de la dissimulation des pensées et motivations privées (MANDEVILLE, 1991 (1729)).

En outre, les spécialistes de Mandeville semblent accepter également que l'élément important concerne la distinction entre les motifs *publics* (apparents) et les motifs *dissimulés* (réels) plutôt qu'entre les motifs *conscients* et les motifs *inconscients*. En effet, dans le passage suivant, Rudi Verburg décrit l'homme mandevillien comme soucieux de constamment dissimuler ses véritables motifs (*conscients*) à autrui : « Derrière la version polie de l'homme et de la société, se trouve un monde de simulations, de fraudes, de ruses et de tromperies, dans lequel les hommes tentent de satisfaire leurs instincts naturels, tout en dissimulant leurs intentions pour ne pas péjorer leurs chances de succès. Au final, cette mascarade a pour fonction de satisfaire les désirs égoïstes de l'homme. Le respect des règles de la société est finalement motivé par le calcul des bénéfices, ce qui rend la coopération et la cohésion fondamentalement instables. » [« Underneath the polished version of man and society, there is a world of pretensions, fraud, cunning, and deception as people try to satisfy their natural impulses while concealing their intentions as not to hurt their chances of success. At the end of the day, the whole masquerade should serve the individual purpose of gratifying man's selfish desires. Compliance with the rules of society is ultimately motivated by a calculation of benefits, whereby cooperation and cohesion are inherently unstable. » (VERBURG, 2015, pp. 663-664).]

Le critère pour distinguer les motivations égoïstes est alors le suivant : les désirs d'une personne égoïste contiennent nécessairement une référence à cette personne. Dit autrement, si le contenu du désir ne peut pas être spécifié sans une référence à l'individu lui-même ou à sa situation, alors il s'agit d'un désir égoïste. Les désirs de l'individu égoïstes sont, pour reprendre l'expression de Bernard Williams, des « désirs pour soi » (*I-desire*), c'est-à-dire des désirs portant sur la manière dont le « moi » est affecté dans un certain état du monde.

[Toutes les choses que] l'égoïste désire sont des choses *pour lui* [...]. Toutes les choses que l'égoïste veut prennent la forme de « je veux que... », où cela est suivi par une description de ce que l'égoïste obtient, par la description d'un état désirable dans lequel l'égoïste se trouve, ou quelque chose de ce genre ; soit, sinon, si ses désirs n'ont pas cette forme, ils dépendent d'un autre désir qui, lui, a cette forme, dans le sens que c'est seulement parce que l'égoïste a ce second désir qu'il a aussi les premiers, et si le second désir disparaissait ou était satisfait autrement, alors les premiers désirs disparaîtraient avec lui.⁶

Tout désir dont les conditions de satisfaction sont déterminées par ce qui se produit *pour la* personne elle-même est considéré comme un désir égoïste. Cela inclut également les désirs qui ne contiennent pas *eux-mêmes* de référence à la personne, mais qui *dérivent* de désirs qui, eux, contiennent une telle référence. Il se peut, par exemple, que je désire que « Ludivine ne soit pas en retard ». Le contenu du désir peut être spécifié sans faire usage des concepts de « moi » ou de « je ». Cependant, si ce désir dérive d'un autre désir, égoïste quant à lui, que « *je* ne perde pas de temps à attendre Ludivine » alors le désir que « Ludivine ne soit pas en retard » est un désir égoïste en vertu du fait qu'il dérive d'un désir égoïste.

A l'inverse, tout désir dont les conditions de satisfaction ne sont pas déterminées par ce qui se produit *pour la personne elle-même* est un désir qui n'est pas égoïste ; on dira qu'il s'agit d'un désir désintéressé. A nouveau, il est possible qu'un désir qui contient une référence à soi dérive d'un désir plus général qui ne contient pas une référence à soi. Dans

⁶ "[A]ll [of the egoist's] desires are for things for him. [...] All the egoist's wants are either of the form 'I want that I . . .', where this is followed by something which specifies his getting something, entering into a desirable state, or whatever; or else, if they are not of that form, they depend on another desire which is of that form, in the sense that it is only because he has that latter desire that he has the former, and that if the latter went away or were satisfied in some other way, the former would go away." (WILLIAMS, 1973, p. 260).

ce cas, le désir dérivé du désir désintéressé est un désir lui-même désintéressé, bien qu'il contienne le terme « moi » ou « je ». Par exemple, si mon désir que « je vienne au secours d'une personne en danger » dérive du désir désintéressé que « la personne en danger soit aidée par *quelqu'un* », alors le désir dérivé est lui-même désintéressé – malgré le fait qu'il contient le terme « je ».

Cette caractérisation de l'égoïsme proposée par Williams rend compte du fait que, dans la pratique ordinaire, les motivations contenant une référence à soi sont qualifiées d'« égoïstes ». Par exemple, le désir que « j'éprouve du plaisir », que « je devienne célèbre », que « je devienne riche », que « je réussisse une carrière académique », que « je sois meilleur que les autres », etc. sont des désirs considérés de manière paradigmatique comme des désirs *égoïstes*. En suivant Williams, le caractère égoïste de ces désirs provient du fait que la référence à soi est inéliminable dans la spécification de leur contenu.

1.1.2. L'ÉGOÏSME ÉCONOMIQUE

Le désir de richesse, de gloire, de réussite personnelle, représentent des cas emblématiques de désirs égoïstes dans l'approche proposée par Williams – conception qui correspond, à notre avis, à la conception ordinaire de l'égoïsme. Il existe toutefois des conceptions plus spécifiques de l'égoïsme. La théorie économique néo-classique adopte notamment une conception plus restreinte. Etant donné que nous nous intéresserons plus loin à cette conception « économique » de l'égoïsme, il convient de la spécifier dès maintenant.

L'égoïsme dans la théorie économique néo-classique se caractérise par une indifférence *totale* à la situation d'autrui. L'égoïste ne se préoccupe pas *du tout* de la situation des autres. Cela exclut les désirs de type *comparatifs* que nous trouvons dans la caractérisation de l'égoïsme offerte par Bernard Williams, par exemple le désir d'être (soi-même) meilleur que les autres. Ainsi, dans cette conception, il ne suffit pas que le contenu du désir contienne une référence à la situation de la personne elle-même pour qu'il puisse

être considéré comme égoïste. Il doit, en outre, ne contenir *aucune* référence à la situation d'autrui (PHAM, 2008).⁷

Ensuite, et il s'agit de l'élément le plus important, le bien ou intérêt de la personne est déterminé en fonction de sa situation *matérielle* ou *économique*.⁸ Autrement dit, les variables constitutives de la *fonction d'utilité* des personnes sont des biens matériels ou économiques. Cela exclut, cette fois, la possibilité de considérer, par exemple, le désir de gloire (sans richesse), de réussite académique (possiblement désastreuse matériellement), le désir d'avoir une image morale positive de soi (qui peut exiger le sacrifice matériel), comme étant des cas de désirs intéressés.⁹

En termes mathématiques, l'homme égoïste de la théorie économique, ou *Homo Economicus*, peut être décrit par la fonction d'utilité suivante (CLAVIEN & CHAPUISAT, 2016) :

Homo Economicus

$$U_i(x_i, x_j) = w_i(x_i, x_j)$$

⁷ Il est intéressant que la motivation la plus plausible pour expliquer ce refus, en économie, des désirs « comparatifs » consiste dans le fait que de tels désirs, dont fait partie notamment *l'envie*, sont considérés comme moralement problématiques.

⁸ La raison pour laquelle l'égoïsme est défini de manière restreinte en économie, comme étant la poursuite de l'intérêt strictement matériel, est que la richesse, ou le bien-être matériel, est considéré comme la principale incarnation de cet égoïsme. Autrement dit, « [l']expression [...] principale de [l']amour exclusif de soi résid[e] dans le désir d'enrichissement matériel [...] » (CAILLÉ, LAZZERI, & SENEL, 2001, p. 22)

⁹ Cette interprétation restreinte de l'intérêt personnel ne poserait pas tellement de problème si elle était bel et bien toujours considérée par les économistes comme une forme *restreinte* de l'intérêt personnel, spécifique à la sphère économique ou matérielle de l'existence. Le problème vient du fait que cette conception restreinte de l'intérêt est parfois considérée comme définissant l'intérêt personnel dans n'importe quelle sphère de l'existence, ou, plus précisément, dans n'importe quel contexte institutionnel. Cela est problématique, par exemple, lorsque les économistes appliquent cette conception restreinte de l'intérêt personnel à la sphère de la recherche académique. Dans ce contexte, les économistes sont amenés à définir l'intérêt personnel des chercheurs comme consistant à vouloir s'enrichir – plutôt que, disons, comme consistant à vouloir être publié et reconnu par ses pairs.

Cette définition de l'intérêt personnel a pour conséquence les deux éléments suivants : (a) la publication des recherches par les chercheurs, sans attente de gains monétaires, est perçue comme un cas de « don », d'altruisme désintéressé. Comme le remarque John O'Neill : « Etant donnée l'introduction de relations de propriété de marché dans la science, l'acte de publier sans obtenir un gain monétaire compte comme un acte surérogatoire de don. » [“Given the introduction of market property relations into science, the act of publishing without some form of payment will count as a superogatory act of gift-giving” (O'NEILL J. , 1990, p. 615).] Or, il ne semble pas que, lorsque les académiciens cherchent à être publiés, ils agissent *contre* leurs intérêts. Au contraire, dans la science, il est dans l'intérêt du chercheur que ses résultats soient publiés.

L'introduction de relations de marché amène, en outre, (b) les universités à redéfinir la propriété intellectuelle, de manière à ce qu'un chercheur ne publie ses recherches que s'il peut en obtenir un gain monétaire. Les universités *interdisent même* parfois de publier en l'absence d'opportunité de gains monétaires. Le problème est que cette démarche va à l'encontre du but de la recherche, qui progresse par le partage des connaissances plutôt que par le secret de la propriété intellectuelle. Pour ces problèmes, voir (O'NEILL J. , 1990).

L'utilité U pour l'individu i est déterminée par le bien-être objectif w_i que i obtient dans la situation matérielle (x_i, x_j) .¹⁰ Pour rendre compte plus explicitement du caractère *égoïste* de cette fonction d'utilité, nous pouvons la résumer par la formule suivante :

$$\text{Homo Economicus} \quad U_i = w_i$$

Cette formule montre plus explicitement que le seul composant ayant une influence sur la fonction d'utilité U de i , c'est-à-dire la valeur que i accorde au résultat, est le bien-être (w) de i lui-même. Ce bien-être est déterminé par les préférences des individus sur des quantités objectives mesurables – par exemple le travail, le repos, la richesse économique, la consommation de biens et services, le profit.

En résumé, l'*Homo Economicus*, ou « égoïste de la théorie économique », est un individu qui cherche à améliorer sa propre situation *matérielle*, sans accorder d'importance à la situation matérielle d'autrui.¹¹ Lorsque nous discuterons de l'objection du « sophisme moraliste » telle qu'elle est formulée par les économistes, nous traiterons de l'égoïsme dans l'acceptation purement économiste du terme, c'est-à-dire comme désir d'enrichissement personnel. Par contre, lorsque nous élargirons la réflexion en dehors de la sphère du marché, pour nous intéresser notamment à l'objection d'« erreur moraliste » telle qu'elle est formulée par les métaéthiciens, nous emploierons la définition plus large de l'égoïsme, selon laquelle l'égoïsme consiste à poursuivre ses propres intérêts – entendus en un sens plus large – ou, autrement dit, à avoir des désirs pour soi.

¹⁰ Une manière plus explicite de rendre compte de l'indépendance de la fonction d'utilité d'un individu, par rapport à la situation d'autrui, consisterait à adopter la formalisation de l'*homo economicus* proposée par Pham, à savoir $U_i = U(y_i)$ (PHAM, 2008, p. 542), dans laquelle la situation matérielle du second individu j n'apparaît pas dans la fonction d'utilité de i . Nous avons retenu la formalisation plus générale, proposée par Clavien, car la caractéristique déterminante de la figure de l'*homo economicus* est l'indifférence *au bien-être d'autrui*, c'est-à-dire à la satisfaction des préférences d'autrui. L'important est donc l'indépendance des *fonctions d'utilité* – i.e. le fait que la fonction d'utilité de i ne contient pas de référence à la fonction d'utilité de j .

¹¹ Cela est conforme à la définition de l'égoïsme qu'en donnent les économistes : « Le seul fait qu'une personne tire avantage de la consommation d'un bien n'affecte pas l'utilité des autres personnes ». [“[T]he fact that one person is enjoying the consumption of a good does not in itself increase or decrease anyone's utility” (SUGDEN, 1981, p. 33)].

1.2. L'ALTRUISME

Dans la section précédente, nous avons proposé une conception générale de l'égoïsme et nous avons exposé la manière, plus spécifique, dont la théorie économique définit l'égoïsme. Dans la présente section, nous allons procéder de façon similaire pour définir l'altruisme ; nous commencerons par proposer une définition générale de l'altruisme puis exposerons les manières plus spécifiques dont l'altruisme est défini en économie.

1.2.1. UNE CONCEPTION GÉNÉRALE DE L'ALTRUISME

Un individu altruiste est un individu qui se préoccupe du bien d'autrui. Un individu altruiste est prêt à sacrifier (au moins une partie de) ses propres intérêts en faveur du bien d'autrui. A l'extrême, un altruiste *pur* ne se soucie de son propre bien que dans la mesure où préserver ce bien lui permet de mieux satisfaire les intérêts d'autrui. Nous pouvons caractériser l'altruisme, dans des termes généraux, de la manière suivante :

Définition de l'altruisme L'acteur A est altruiste seulement si, par son action, il vise à satisfaire les intérêts de l'acteur B, à réaliser le bien de l'acteur B, lorsque $A \neq B$.¹²

Comme nous l'avons fait avec l'égoïsme, il convient de souligner deux éléments importants relativement à cette caractérisation de l'altruisme. En premier lieu, notre définition est incompatible avec la conception *biologique* de l'altruisme. Elle serait également incompatible avec les doctrines qui – si elles existaient – considéreraient l'altruisme comme une motivation *inconsciente*, comme le moteur causal inconscient de l'action. En second lieu, cette définition ne propose qu'une condition *nécessaire* de l'altruisme. Elle laisse ouverte la question de savoir si une personne motivée subjectivement à faire le bien, mais poussée objectivement par un motif égoïste, doit également être considérée comme altruiste – bien que nous pencherions plutôt vers une réponse négative à cette question.

¹² Cette définition de l'altruisme est compatible avec toutes les conceptions de l'intérêt ou du bien des personnes. Notamment, elle est compatible avec la conception du bien personnel, répandue en économie, comme étant identique à la *satisfaction des préférences*. Dans ce cas, un altruiste A qui vise à faire le bien de B est un individu qui, par son action, vise à satisfaire les préférences de B. Notre définition de l'altruisme est, bien entendu, également compatible avec des conceptions objectives du bien personnel ou de l'intérêt. À noter que cette définition de l'altruisme est également compatible, voire implique, l'absence d'égard pour le bien moral ou la justice.

Dans la classification proposée par Bernard Williams, l'altruiste n'est pas simplement un individu dont les désirs ne contiennent pas de référence à « soi ». Il existe en effet de nombreux désirs qui ne contiennent pas de référence à soi et qui ne contiennent pas non plus de référence aux intérêts de quelqu'un d'autre. De tels désirs ne sont ainsi ni égoïstes, ni altruistes. Par exemple, le désir esthétique que « la collégiale de Neuchâtel soit rénovée », le désir que « la connaissance de l'univers soit élargie » ou que « la nature soit préservée » sont des désirs désintéressés qui ne contiennent pas de référence au bien d'autrui.

En outre, comme le fait remarquer Williams (1973, p. 263), il peut exister des désirs non égoïstes qui portent sur le *malheur* d'autrui : le désir que quelqu'un échoue, qu'il lui arrive un malheur, etc. On parlera dans ce cas de *désintérêt* malveillant¹³ ou de désirs malveillants désintéressés. Nous reviendrons plus loin sur cette idée. De tels désirs ne sont pas égoïstes, ils ne concernent pas le « moi », mais ils ne sont pas non plus des désirs altruistes, puisqu'ils ne portent pas sur le *bien* d'autrui. Autrement dit, ils sont *désintéressés* mais ils ne sont pas altruistes. Les désirs *altruistes* ne constituent donc qu'un sous-ensemble des désirs non égoïstes ou désintéressés, c'est-à-dire des désirs ne portant pas sur le bien personnel.

Une complication supplémentaire, pour partager les désirs égoïstes et altruistes, provient des désirs dits *partiaux*.¹⁴ Par exemple, Cécile peut désirer que *son cousin* Michel se porte bien. En un sens, le désir de Cécile est un désir altruiste, puisqu'il porte sur le bien d'une autre personne. Toutefois, Cécile ne se préoccupe du bien de Michel que dans la mesure où Michel est bien *son* cousin à elle. Le désir de Cécile ne peut donc pas être décrit sans faire usage d'un indexical référant à sa propre personne. En ce sens, il s'agit d'un désir égoïste. Il semblerait donc que nous soyons en présence d'un désir à la fois altruiste et égoïste. Pour notre propos, nous considérerons que de tels désirs appartiennent plutôt à la catégorie des désirs *égoïstes*, étant donnée l'indépassable référence à soi dans la caractérisation du contenu. Mais nous reconnaissons qu'il s'agit d'un cas difficile à catégoriser à l'aide des définitions que nous proposons.

¹³ Nous reprenons ici l'expression utilisée par Alain Wolfelsperger (2010) dans l'un de ces articles.

¹⁴ Cette complication a été relevée par le professeur Laurent Jaffro

1.2.2. L'ALTRUISME ÉCONOMIQUE

La théorie économique définit l'altruisme dans le même formalisme que celui utilisé pour définir l'égoïsme. Nous l'avons vu, la fonction d'utilité de l'égoïste est déterminée exclusivement par la manière dont son propre bien-être est affecté. La fonction d'utilité de l'altruiste contient quant à elle nécessairement une référence au bien-être *d'autrui*. Par exemple, dans sa forme égalitariste, l'altruiste accorde une importance égale à son propre bien-être et au bien-être d'autrui (CLAVIEN & CHAPUISAT, 2016, p. 25):

Altruisme égalitariste

$$U_i = w_i + w_j \quad \text{où } w_i \text{ fait référence au bien-être de l'individu } i \text{ lui-même et } w_j \text{ fait référence au bien-être d'un autre individu.}$$

Il existe encore d'autres variantes de l'altruisme, qui varient en fonction de l'importance relative que l'individu accorde au bien-être d'autrui par rapport à son propre bien-être. A l'extrême, un altruiste pur est un individu qui n'accorde pas de poids à son propre bien-être. Sa fonction d'utilité se traduirait de la manière suivante :

Altruisme total

$$U_i = w_j \quad \text{où la seule variable déterminante de la fonction d'utilité est le bien-être d'autrui } j.$$

Quelle que soit la conception spécifique de l'altruisme qui est retenue, toutes les conceptions partagent l'idée que le bien-être d'autrui constitue une variable déterminante de la fonction d'utilité d'une personne – i.e. de l'ordre de préférences déterminant ses choix.

1.3. LA MOTIVATION MORALE

Dans les deux sections précédentes, nous avons présenté l'égoïsme et l'altruisme d'abord en termes généraux, puis dans les termes plus spécifiques de l'économie. Dans la présente section, nous nous contenterons de définir la motivation morale dans des termes généraux. Nous consacrerons un chapitre complet à l'analyse de cette motivation.

Alors que l'égoïsme caractérise le fait de se soucier de son propre bien et l'altruisme le fait de se soucier du bien d'autrui, la motivation morale consiste dans le fait de chercher à réaliser une valeur morale, notamment le *bien* (moral), le *juste*, ou la *vertu*. La motivation morale consiste ainsi dans le fait, pour un individu, d'être motivé à faire une action

simplement parce que cette action possède une certaine qualité ou propriété morale (intrinsèque ou extrinsèque). Ainsi, un individu a une motivation morale si et seulement si il est motivé à faire une action pour la seule raison que cette action est intrinsèquement et/ou instrumentalement bonne/correcte/vertueuse du point de vue moral.

Une telle motivation, à l'instar de l'altruisme, peut être décrite comme *désintéressée* : elle ne porte pas sur les intérêts de la personne elle-même. Son contenu ne contient pas de référence indépassable à « soi ». Toutefois, cette motivation se distingue de l'altruisme en tant qu'elle ne vise pas à satisfaire les intérêts d'autrui, mais qu'elle vise à réaliser une valeur morale – ce qui peut, le cas échéant, aller à l'encontre de la satisfaction des intérêts personnels mais aussi à l'encontre de la satisfaction des intérêts d'autrui. Par exemple, le souci d'équité ou d'impartialité sont des motivations morales, mais elles ne sont pas des motivations altruistes – puisqu'elles ne visent pas nécessairement à améliorer le bien-être d'autrui ; elles visent à réaliser une valeur morale. Une personne cherchant, par exemple, à réaliser la justice distributive pourra être amenée à prendre une partie des richesses d'une personne pour les donner à autrui, ou pour elle-même. Dans un sens, elle agit donc *contre* l'intérêt matériel d'autrui, bien qu'elle agisse de manière désintéressée – au nom de la justice ou de l'équité.¹⁵

Nous pouvons donc définir la motivation morale de la manière suivante :

Définition de la motivation morale L'acteur A dispose d'une motivation proprement morale seulement si, par son action, il vise à réaliser le bien, à faire ce qui est juste ou à être une personne vertueuse.¹⁶

¹⁵ Bien entendu, de la même manière qu'un désir portant sur le bien d'autrui peut, en fait, être un désir égoïste lorsqu'il dérive d'un désir égoïste, il peut arriver qu'un désir portant sur le bien d'autrui ou sur son propre bien dérive d'une motivation morale plus générale. Dans ce cas, le désir portant sur le bien d'autrui ou sur son propre bien est lui-même un désir proprement moral. Par exemple, si la motivation de Lena pour rendre visite à son ami Vladimir à l'hôpital dérive d'un désir de faire ce qu'il est juste de faire à l'égard d'un ami, alors la motivation de Lena pour rendre visite à Vladimir est une motivation proprement morale. Réciproquement, nous pourrions avoir des désirs portant sur des valeurs morales qui soient dérivés de motivations altruistes, dans quel cas les désirs portant sur les valeurs morales seraient en fait eux-mêmes des désirs altruistes. Par exemple, il se peut que Vladimir se soucie du bien-être de son amie Lena. Or, Lena accorde beaucoup d'importance aux questions d'injustice. En conséquence, pour faire plaisir à Lena, Vladimir choisit de faire des actions visant à supprimer les injustices. Dans ce cas, sa motivation de « supprimer les injustices dans le monde », bien qu'ayant un contenu proprement moral, est en fait un désir altruiste.

¹⁶ Le désir « d'être une personne vertueuse » semble, de prime abord, correspondre à ce que nous qualifions de « désir égoïste ». En effet, il semble contenir une « référence indépassable » à soi. Néanmoins, il

Comme pour l'égoïsme et l'altruisme, cette description de la motivation morale doit être comprise en termes intensionnels : l'acteur A, par son action, vise à réaliser ce qu'il *juge* être le bien, à faire ce qu'il *juge* être juste et à être une personne qu'il *juge* vertueuse. A nouveau aussi, cette définition ne porte que sur une condition *nécessaire* et non suffisante de la motivation morale. Ainsi, si une personne cherche subjectivement à faire le bien, à faire ce qui est juste ou à être une personne vertueuse, mais qu'elle est mue inconsciemment par le désir de promouvoir ses propres intérêts, notre définition ne se prononce pas sur le caractère moral ou non de sa motivation globale, considérée à la fois subjectivement et objectivement. Certes, lorsque nous tenterons, plus loin, de distinguer la motivation morale de motivation proches mais différentes, nous tendrons à considérer que l'attribution d'une motivation proprement morale à un individu dépend à la fois des motifs subjectifs et objectifs. Mais rien dans la définition ci-dessus ne nous y contraint.

Notons également que nous interprétons le concept de *jugement moral* de manière neutre du point de vue métaéthique, c'est-à-dire comme ne présupposant ni que le jugement soit une *croyance*, ni que ce jugement soit un état mental non cognitif. Notre usage de la notion de « jugement moral » laisse ouverte la question de la nature cognitive ou non cognitive de cet état mental. Nous utilisons donc, et procéderons ainsi dans l'ensemble de ce travail, la notion de « jugement moral » comme étant *neutre* entre le cognitivisme et le non-cognitivisme métaéthiques; un jugement moral consiste, simplement, à prendre position, à avoir une conviction morale. Cet usage neutre est conforme à l'usage accepté dans la littérature métaéthique.¹⁷

s'agit, selon nous, d'un désir dérivé d'un désir lui-même désintéressé. Or, nous l'avons dit, un désir qui est dérivé d'un désir désintéressé est lui-même un désir désintéressé, même si ce désir dérivé contient une référence à soi. Le désir « d'être une personne vertueuse » est un désir qui est dérivé du désir plus général de réaliser la vertu, chez soi ou chez autrui. Si un individu désire être *lui-même* vertueux, ce n'est que parce qu'il accorde de la valeur, d'abord, à la vertu en général. En cela, le désir d'être une personne vertueuse diffère, par exemple, du désir d'être une personne riche. La personne qui désire être riche n'accorde pas de la valeur à la richesse *en général* et à la richesse *pour soi* de manière dérivée. Si tel était le cas, alors nous pourrions dire qu'elle a effectivement un désir désintéressé de richesse. Mais cela semble peu plausible ; si elle désire la richesse, ce n'est qu'en tant que cette richesse est *la sienne*. A l'inverse, la personne qui désire être une personne vertueuse accorde de la valeur à la vertu en elle-même, et désire devenir elle-même une personne vertueuse de manière dérivée. Le désir d'être une personne vertueuse n'est donc pas un désir égoïste dans le sens défini précédemment ; il s'agit d'un désir dérivé du désir désintéressé de « réaliser la vertu ». Je remercie mon directeur de thèse, Richard Glauser, pour m'avoir fait remarquer cette tension dans la définition proposée du désir « d'être une personne vertueuse ».

¹⁷ Comme le prouve l'usage du terme dans le livre d'introduction à la métaéthique de François Jaquet et Hichem Naar (2019).

Cette définition de la motivation morale offre certains avantages, mais rencontre aussi certaines difficultés, notamment pour la traduire dans le langage formel de l'économie. Dans la section suivante, nous présenterons les avantages et les inconvénients de la définition générale de la motivation morale que nous proposons.

2. LA MOTIVATION MORALE

Dans cette section, nous allons (2.1) voir que la conception proposée de la motivation morale est compatible avec l'ensemble des doctrines métaéthiques et des doctrines morales normatives existantes, (2.2) présenter quelques difficultés concernant la formalisation de la motivation morale avec les outils de l'économie et (2.3) distinguer la motivation morale d'autres motivations proches mais néanmoins différentes.

2.1. COMPATIBILITÉ AVEC L'ENSEMBLE DES DOCTRINES MÉTAÉTHIQUES ET NORMATIVES

L'avantage de la définition générale de la motivation morale que nous avons proposée est qu'elle est compatible avec l'ensemble des conceptions métaéthiques existantes¹⁸ et, surtout, avec l'ensemble des doctrines morales ou politiques *normatives* existantes.

2.1.1. COMPATIBILITÉ AVEC LE RÉALISME MORAL FORT NON NATURALISTE

Concernant la compatibilité avec les doctrines métaéthiques, la définition proposée est, par exemple, compatible avec les différentes variétés de réalismes métaéthiques *forts*¹⁹, et notamment avec les différentes variétés de réalismes métaéthiques *non naturalistes*. En effet, elle est compatible avec le fait que la valeur morale est une propriété non naturelle des *états de choses*, des *actions* ou des *personnes* et avec le fait qu'elle est une propriété *intrinsèque* ou *extrinsèque* des actions.

¹⁸ Une description, ainsi qu'un tableau résumé, des différentes positions métaéthiques contemporaines, est disponible dans l'annexe 1.

¹⁹ Le réalisme moral fort est la position métaéthique selon laquelle (i) les énoncés moraux sont vrais ou faux – et expriment des croyances, (ii) certains énoncés moraux sont vrais, (iii) les énoncés moraux sont rendus vrais par des faits indépendants des attitudes des agents. On distingue, au sein du réalisme moral fort, la version non naturaliste, selon laquelle les propriétés morales sont des propriétés *sui generis* irréductibles à des propriétés naturelles, et la version naturaliste, selon laquelle les propriétés morales sont des propriétés naturelles.

Elle s'accommode facilement, par exemple, de la conception métaéthique de G. E. Moore selon laquelle la *bonté* est une caractéristique intrinsèque et inanalysable des *états de choses* et une action est *moralement correcte* si elle permet de réaliser l'état de chose contenant le maximum de « bonté » (MOORE G. E., 1976). La motivation morale consisterait, dans cette perspective, à désirer réaliser un état de chose intrinsèquement bon et à vouloir faire une action parce que cette action est « juste », i.e. parce qu'elle maximise la quantité totale de « bonté » dans le monde.

Notre définition de la motivation morale est aussi compatible avec la conception kantienne de la valeur morale. Selon cette conception, une action doit satisfaire deux conditions pour avoir une valeur morale positive : en premier lieu, elle doit être *conforme* à un principe moral fondamental, à savoir l'Impératif Catégorique.²⁰ Une action conforme à l'Impératif Catégorique est considérée comme moralement *droite* ou *correcte*, mais elle ne possède pas encore de *valeur* morale, elle n'est pas *bonne*. Pour avoir une telle valeur morale positive, une deuxième condition doit être satisfaite : il faut que la personne qui accomplit cette action soit motivée à agir précisément par le fait que cette action est conforme au devoir.

Par exemple, si un entrepreneur renonce à dissimuler de l'information sur les produits qu'il vend dans le but d'éviter des sanctions légales, alors son action est *conforme* à son devoir moral, mais son action n'a pas de *valeur* morale ; l'entrepreneur n'ayant pas agi par conscience morale, mais uniquement pour éviter des ennuis, son action est dépourvue de valeur morale. Si, par contre, l'entrepreneur renonce à dissimuler des informations simplement parce qu'il sait qu'il est immoral de les dissimuler, alors son action est non seulement *conforme* au devoir moral, mais elle possède également une valeur morale positive. La valeur morale d'une action est déterminée par le fait que la considération du caractère moral ou immoral de l'action suffit à motiver ou démotiver l'acteur de réaliser l'action. Dans la conception kantienne, la motivation morale consiste donc plus spécifiquement à vouloir faire une action en vertu du fait que cette action est conforme à

²⁰ Il existe différentes variantes de ce principe. Les plus connues sont les deux suivantes : (1) « Agis uniquement d'après la maxime grâce à laquelle tu peux vouloir en même temps qu'elle devienne une loi universelle. » (KANT, 2007 (1785), p. 54), (2) « Agis de telle sorte que tu traites l'humanité, aussi bien dans ta propre personne que dans la personne de tout autre, toujours en même temps comme une fin, jamais seulement comme un moyen. » (KANT, 2007 (1785), p. 65)

l'Impératif Catégorique ou à vouloir renoncer à faire une action du fait que cette action n'est pas conforme l'Impératif Catégorique.

La conception de la motivation morale que nous proposons est encore compatible avec le réalisme normatif développé par Thomas Scanlon (SCANLON, 1998). Selon ce dernier, toutes les notions normatives (bon, juste, vertueux) peuvent être définies par la notion de *raisons d'agir*. Les raisons d'agir sont des entités normatives *primitives* ; elles sont des « considérations qui comptent en faveur d'une action » (SCANLON, 1998, p. 17). La notion normative de *bonté* est définie par celle de *raison d'agir* de la manière suivante : un état de chose est *bon* s'il possède des propriétés naturelles qui offrent des *raisons* d'adopter une attitude positive à son égard.

Par exemple, ce qui rend *bonne* l'activité sportive est le fait que l'activité sportive possède une propriété naturelle (i.e. favoriser la santé) qui offre une *raison* de désirer pratiquer une telle activité. Autrement dit, la propriété naturelle de favoriser la santé rend l'activité sportive *désirable* et, *de ce fait*, bonne. De la même manière, ce qui peut rendre *mauvaise* une activité économique est que cette activité possède une propriété naturelle (i.e. causer des dégâts environnementaux) qui offre une *raison* de ne pas continuer cette activité économique.

Ainsi, la valeur ou la bonté d'un état de chose consiste dans le fait de posséder des propriétés naturelles qui offrent ou constituent des raisons d'adopter une attitude positive à l'égard de cet état de chose:

Etre bon, ou avoir de la valeur, n'est pas une propriété qui elle-même offre une raison d'adopter une certaine attitude. Plutôt, être bon ou avoir de la valeur consiste dans le fait d'avoir une autre propriété qui constitue de telles raisons. [...] Cette approche [...] considère la bonté et la valeur comme étant des propriétés non naturelles, à savoir la propriété purement formelle, de second ordre, d'avoir des propriétés de premier ordre qui donnent des raisons du type pertinent.²¹

²¹ "Being good, or valuable, is not a property that itself provides a reason to respond to a thing in certain ways. Rather, to be good or valuable is to have other properties that constitute such reasons. [...] [T]his account [...] takes goodness and value to be non-natural properties, namely the purely formal, higher-order properties of having some lower-order properties that provide reasons of the relevant kind". (SCANLON, 1998, p. 97)

Vouloir un objet parce que cet objet est bon consiste donc à vouloir cet objet en vertu de ses propriétés naturelles qui le rendent digne d'être désiré – i.e. qui donnent des *raisons* de le désirer.

Jusqu'à présent, nous nous sommes contentés de présenter la théorie de Scanlon concernant les propriétés *normatives* en général. Or, nous sommes plus spécifiquement intéressés par sa théorie des propriétés *morales*. La qualité *morale* des actions est constituée, dans la conception *contractualiste* de Scanlon, par le fait que l'action est conforme à des principes auxquels aucune personne informée et libre n'a de *raison* de s'opposer.

Le contractualisme [...] soutient qu'un acte est moralement incorrect si sa réalisation dans des circonstances données serait interdite par n'importe quel ensemble de principes, destinés à la régulation générale du comportement, auxquels personne n'a de raison de s'opposer en tant que base pour un accord informé et non contraint.²²

A nouveau, la notion normative de *rectitude morale* d'une action, à l'instar de la notion de *bonté*, se définit par le biais de la notion normative primitive de *raison d'agir*. En résumé, une action est *correcte* si elle est conforme à un ensemble de principe auquel personne n'a de raison de s'opposer. La motivation morale correspondant à cette théorie de la rectitude morale consiste donc à désirer faire une action du fait que cette action peut être *justifiée* à autrui, puisqu'elle est conforme à des principes que personne n'a de raison de rejeter. Autrement dit, la motivation *morale* chez Scanlon consiste dans le fait de vouloir vivre avec les autres personnes selon des règles qu'elles peuvent également accepter.²³ C'est en cela que consiste, dans sa doctrine, la motivation (morale) à faire ce qui est juste.

2.1.2. COMPATIBILITÉ AVEC LE RÉALISME MORAL NATURALISTE

Pour Scanlon (2014) – et, avec lui, Thomas Nagel (1970), Derek Parfit (2011) et Charles Larmore (2008), notamment – les propriétés normatives sont des propriétés *non naturelles*.

²² "Contractualism [...] holds that an act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a basis for informed, unforced general agreement." (SCANLON, 1998, p. 153)

²³ Les individus qui sont motivés à agir par des raisons morales « veulent faire des actions qui peuvent être justifiées aux autres » "[...] want to act in ways that could be justified to others" (SCANLON, 1998, p. 154).]

Il est néanmoins aussi possible de soutenir que la qualité morale des actions est une propriété *naturelle* découverte empiriquement. Par exemple, pour Peter Railton, un comportement est moralement approprié s'il correspond à ce qu'il est rationnel de faire d'un point de vue social. Et une situation rationnelle du point de vue social se caractérise par le fait qu'elle résulte d'une décision parfaitement informée dans laquelle les intérêts objectifs de toutes les personnes affectées ont été pris en compte (RAILTON, 2016, p. 191). Or, étant donné que cette propriété de « tenir compte des intérêts de tous » joue un rôle causal dans l'explication de certains phénomènes sociaux²⁴, la propriété d'une action d'être rationnelle du point de vue social – et donc d'être moralement bonne – est une propriété naturelle. C'est pourquoi le réalisme moral de Railton est un réalisme moral *naturaliste*.

La caractérisation de la motivation morale, dans la théorie de Railton, est alors la suivante : lorsque nous cherchons à faire ce qui est juste, nous cherchons à faire des actions qui peuvent être justifiées aux autres personnes de la société du point de vue de la rationalité collective. Etre motivé à agir moralement revient, dans sa théorie, à être motivé à agir d'une manière que nous pouvons justifier à autrui, comme le montre le passage suivant :

Etant donné que dans le cadre du discours public et de la réflexion privée, nous sommes souvent concernés par la question de savoir si notre conduite est justifiable d'un point de vue général plutôt que d'un point de vue seulement personnel, [cela explique] le fait que nous attachions autant d'importance à la moralité à titre de standard et de critique [...] (RAILTON, 2016, p. 200).

A chaque fois que nous nous soucions de pouvoir justifier nos actions à nos propres yeux et aux yeux d'autrui, nous nous préoccupons de morale. Dans la théorie de Railton, cette préoccupation implique, substantiellement, de se soucier de compter les intérêts de chacun de manière égale. Cette conception de la motivation morale est, à nouveau,

²⁴ « [U]ne organisation sociale [*social arrangement*] – par exemple, une forme de production, une hiérarchie sociale ou politique, etc. – qui dévie par rapport à la rationalité sociale en négligeant considérablement les intérêts d'un groupe spécifique sera potentiellement sujette à l'insatisfaction et à l'agitation. » (RAILTON, 2016, p. 192) et plus loin dans le texte : « Outre le fait qu'elles sont caractérisées par un certain type d'agitation sociale, les sociétés qui ne parviennent pas à se rapprocher de la rationalité sociale peuvent également partager des caractéristiques supplémentaires : elles peuvent manifester une tendance vers certaines doctrines religieuses ou idéologiques, ou vers certains types d'appareils de répression ; elles peuvent être moins productives à certains égards (par exemple, ne pas être en mesure de développer certaines ressources humaines) et plus productives dans d'autres contextes (par exemple, dans l'extraction à moindre coût d'une plus grande force de travail à partir de certains groupes) [...] » (RAILTON, 2016, pp. 192-193).

compatible avec la conception de la motivation morale que nous avons proposée au début du chapitre.

2.1.3. COMPATIBILITÉ AVEC LE RÉALISME MORAL FAIBLE

Nous avons vu jusqu'à présent que notre définition de la motivation morale est compatible avec les variantes du réalisme moral *fort*, aussi bien dans sa version non naturaliste (2.2.1) que naturaliste (2.2.2). Nous allons voir dans la présente sous-section qu'elle est, en outre, compatible avec les conceptions métaéthiques *faiblement* réalistes,²⁵ selon lesquelles les faits moraux sont des faits *au sujet* des attitudes des agents. Par exemple, elle est compatible avec la conception métaéthique relativiste de Gilbert Harman (2000) selon laquelle la vérité d'un énoncé au sujet de ce qu'il est bon ou mal pour un agent S de faire est déterminée par le contexte et, notamment, par ce qui est accepté ou non par le groupe de personnes auquel S appartient.²⁶ La motivation morale, i.e. le désir de faire le bien ou ce qui est juste, consiste donc, dans cette théorie, à être motivé à agir par le fait que l'action est conforme à un principe accepté par le groupe dont on est membre.

Notre conception de la motivation morale est compatible avec d'autres théories faiblement réalistes, notamment les théories dites *constructivistes*.²⁷ Selon les partisan-ne-s

²⁵ Une doctrine métaéthique est faiblement réaliste lorsqu'elle admet les thèses suivantes : (i) les énoncés moraux sont vrais ou faux – et expriment des croyances, (ii) certains énoncés moraux sont vrais et (iii) les faits qui rendent vrais les énoncés moraux sont des faits au sujet des attitudes des êtres humains – et ne sont donc pas des faits indépendants des attitudes des acteurs moraux (vs réalisme moral fort).

²⁶ « Les raisons morales d'un agent dérivent typiquement de principes que l'agent accepte en tant que membre d'un groupe ». [“The agent's moral reasons typically derive from principles the agent accepts as a member of a group.” (HARMAN, 2000, p. 197)]. Cette conception de la nature de la valeur morale nous semble être celle adoptée par certains économistes, notamment Tirole et Bénabou, que nous retrouverons régulièrement tout au long de ce travail. « Ce qui rend un comportement donné socialement ou moralement inacceptable est souvent le seul fait que « cela ne se fait pas », ce qui implique que seules des personnalités extrêmes, des marginaux sociaux, ne seraient pas dissuadés par la honte intense rattachée à ce comportement. Dans d'autres lieux à d'autres temps, étaient acceptés des normes et des codes d'honneur différents et le fait que « tout le monde le fait » permet que le même comportement soit délivré de toute stigmatisation. Des exemples concernent le choix de la reddition plutôt que de la mort, ne pas se rendre à l'église, ne pas voter, divorcer, la faillite bancaire, le chômage, la dépendance économique, les évasions fiscales mineures, et des modes de consommation ostentatoires. ». [“What makes a given behavior socially or morally unacceptable is often the very fact that “it is just not done”, meaning that only people whose extreme types make them social outliers would not be dissuaded by the intense shame attached to it. In other places or times different norms or codes of honor prevail, and the fact that “everyone does it” allows the very same behavior to be free of all stigma. Examples include choosing surrender over death, not going to church, not voting, divorce, bankruptcy, unemployment, welfare dependency, minor tax evasion, and conspicuous modes of consumption”. (TIROLE & BENABOU, 2006, p. 1666)]

²⁷ Une doctrine métaéthique est constructiviste lorsque (i) elle admet le réalisme moral faible et (ii) spécifie que les attitudes qui constituent les conditions de vérité des énoncés moraux sont les attitudes que l'agent aurait dans une situation hypothétique ou idéale.

du constructivisme métaéthique, ce qu'il est bon, correct ou vertueux pour S de faire est déterminé par ce que S serait motivé à faire dans certaines circonstances idéales ou hypothétiques – par exemple, si S était pleinement rationnel, c'est-à-dire s'il ne possédait pas de fausses croyances et connaissait l'ensemble des faits non moraux pertinents, s'il avait des désirs cohérents, s'il respectait les normes de raisonnement dans la prise de décision, etc.

Par exemple, Michael Smith soutient qu'une personne S a le devoir moral de faire X dans les circonstances C seulement si l'ensemble des êtres humains, s'ils étaient parfaitement rationnels, désireraient que tout le monde fasse X dans les circonstances C (SMITH, 1994). Pour illustrer son propos, le philosophe donne l'exemple du devoir moral de rendre à son propriétaire tout bien égaré ou perdu que nous trouverions dans la rue.

Il est moralement correct que je rapporte la valise que j'ai trouvée dans la rue dans telles ou telles circonstances [...] seulement si, dans des conditions de rationalité complète, nous voudrions tous que, si nous trouvions une valise dans la rue dans telles ou telles circonstances, alors nous la rapporterions.²⁸

Dans le constructivisme de Michael Smith, comme dans le constructivisme moral de Christine Korsgaard (1996) ou John Rawls²⁹, la motivation morale consiste donc à être motivé à faire une action parce que cette action serait celle que nous serions tous motivés à faire dans certaines circonstances idéales ou hypothétiques.

2.1.4. COMPATIBILITÉ AVEC LES ANTIRÉALISMES

Notre conception de la motivation morale est compatible également avec l'antiréalisme, i.e. la doctrine métaéthique selon laquelle il n'existe pas de faits moraux.

²⁸ "My handing back a wallet I found in the street in such and such circumstances is right, for example, only if, under conditions of full rationality, we would all want that if we find a wallet in the street in such and such circumstances, then we hand it back." (SMITH, 1994, p. 187).

²⁹ « [L]e sentiment de la justice n'est pas un désir qui diffère de celui d'agir d'après des principes que des individus rationnels accepteraient dans une situation initiale donnant à chacun une représentation égale en tant que personne morale [...]. » (RAWLS, 2009, p. 518).

L'antiréalisme peut prendre deux formes principales ; (A) le non-cognitivism³⁰ et (B) la théorie de l'erreur.³¹

Pour le non-cognitivism (A), les jugements moraux sont des états mentaux conatifs, notamment des désirs, ou des états affectifs, tels que des émotions. La motivation morale consiste ainsi, dans le non-cognitivism, à être motivé à agir par une attitude morale non cognitive à l'égard d'une action.

Pour la théorie de l'erreur morale (B), les propriétés morales auxquelles nous faisons référence dans notre discours moral n'existent pas en réalité. Dans ce contexte, la motivation morale consiste à être motivé à agir parce que l'on *croit* (faussement) que l'action permet de réaliser le bien, de faire ce qui est juste ou d'être une personne vertueuse.

2.1.5. COMPATIBILITÉ AVEC L'ENSEMBLE DES THÉORIES MORALES NORMATIVES

Il nous semble avoir montré que la conception de la motivation morale que nous proposons ne nous engage aucunement à adopter l'une plutôt que l'autre des théories métaéthiques existantes. Il nous semble, en second lieu, qu'elle ne nous engage pas non plus du point de vue *normatif* ; notre conception de la motivation morale est compatible avec l'ensemble des doctrines normatives substantielles au sujet de *ce qui, dans le monde, est bon, juste ou vertueux*. Par exemple, dans une doctrine conséquentialiste utilitariste, l'individu disposant d'une motivation morale cherche à réaliser, par son action, l'état du monde qui maximise le bonheur du plus grand nombre de patients moraux. Dans une conception normative déontologique, l'individu moralement motivé désire faire les actions moralement correctes et éviter les actions moralement incorrectes, par exemple donner de l'argent à une œuvre de charité, ne pas mentir, ne pas tuer autrui, etc. Dans l'approche normative de l'éthique de la vertu, les individus motivés à agir moralement désirent développer les traits de caractères des personnes vertueuses – par exemple, être honnêtes,

³⁰ Une théorie métaéthique non cognitiviste se caractérise de la manière suivante : (i) il n'existe pas de faits moraux, (ii) les énoncés moraux ne sont ni vrais ni faux et (iii) ils n'expriment pas des croyances (ou tout autre état mental cognitif).

³¹ La théorie de l'erreur est la doctrine métaéthique selon laquelle : (i) les énoncés moraux sont vrais ou faux – et expriment des croyances, (ii) les énoncés moraux *prétendent* décrire des faits moraux objectifs, mais (iii) il n'existe pas de faits moraux objectifs et, donc, (iv) tous les énoncés moraux sont faux.

loyaux, tempérants, justes et généreux, mais aussi ne pas être irascibles, cupides et envieux, etc.³²

Avant de continuer, il nous faut néanmoins apporter la précision suivante : si nous avons montré que l'ensemble des théories normatives sont *compatibles* avec notre conception de la motivation morale, elles sont toutes également compatibles avec le fait que cette motivation morale *n'existe pas*. Par exemple, certains partisans de l'utilitarisme considèrent qu'il n'existe pas, chez l'être humain, une motivation morale correspondant au principe utilitariste ; il n'existe pas chez eux de motivation à faire ce qui maximise le bonheur du plus grand nombre. Pour l'utilitariste David Brink (1997), notamment, les êtres humains sont motivés à agir pour réaliser soit leur propre bien (égoïsme), soit le bien d'autrui (altruisme), mais il n'existe pas de motivation morale *supplémentaire* consistant à vouloir faire ce qui est *juste* du point de vue impartial, à savoir ce qui maximise le bonheur du plus grand nombre.³³ Selon lui, les êtres humains ne choisiront de faire les actions qui ont la propriété de maximiser le bonheur du plus grand nombre que lorsque ces actions auront la propriété *différente* de réaliser le bien de la personne elle-même (égoïsme) ou de réaliser le bien des personnes qui ont de l'importance pour elle (altruisme).

Pour l'instant, il nous importe seulement de montrer que notre définition de la motivation morale *n'exclut* aucune théorie morale normative. Rien ne nous autorise encore ni à affirmer qu'une telle motivation morale *existe* dans les faits, ni même qu'elle constituerait une ressource motivationnelle fiable pour la vie en société, si elle existait.

2.2. LA DIFFICULTÉ DU CHOIX DU FORMALISME

La représentation formelle de la motivation morale en théorie économique pose un plus grand nombre de difficultés que la représentation formelle de l'égoïsme et de l'altruisme. En premier lieu (1), le formalisme n'est pas le même si l'on considère la

³² Notre conception générale de la motivation morale est neutre concernant la valeur morale des différentes motivations. Il ne nous est pas nécessaire de considérer qu'une personne qui vise à faire ce qui est juste ou à être une personne vertueuse est meilleure qu'une personne qui cherche directement à faire le bien d'autrui (altruisme) ou qu'une action ne possède de valeur morale que si elle est le résultat d'une motivation morale. Nous cherchons, pour le moment, uniquement à distinguer conceptuellement la motivation morale des autres motivations, mais nous ne prenons pas position concernant la valeur morale de ces motivations elles-mêmes.

³³ Il nous semble qu'un utilitariste peut refuser l'existence d'une motivation proprement morale uniquement au niveau des citoyens ou membres d'une société. Il nous semble qu'il est contraint d'attribuer par contre aux législateurs ou aux individus occupant une charge publique un véritable souci utilitariste de maximiser le bonheur du plus grand nombre.

motivation morale, par exemple, comme un simple « goût » pour la morale ou plutôt comme une préférence de second ordre « d'être une personne qui a des préférences morales ». En second lieu (2), la représentation de la motivation morale varie en fonction du *contenu substantiel* attribué à la morale.

2.2.1. UN « GOÛT » POUR LA MORALE ?

Une première question qui se pose est la suivante : faut-il représenter la motivation morale comme un « goût moral » homogène aux autres types de goûts – par exemple au goût pour le chocolat ? Certains économistes ont favorisé cette représentation, en considérant la motivation morale comme un « goût de donner » ou comme une « préférence pour l'équité » ou encore un « goût pour l'égalité ». Dans de telles perspectives, la fonction d'utilité des individus varie en fonction de l'importance que cet individu accorde à la valeur morale de son action. Nous pourrions ainsi imaginer un individu, désireux de manger une glace, qui se dirige en direction du marchand de glace sur une plage. Sur son chemin, notre individu aperçoit une personne en train de se noyer. Il est le seul à pouvoir la sauver, mais s'il décide de la sauver, il ne pourra plus manger de glace car le marchand sera parti. Dans ce contexte, si le « goût » de notre individu pour les glaces est plus fort que son « goût » pour la morale, il risque bien d'ignorer les appels au secours de la malheureuse afin de pouvoir manger la glace tant désirée.

Le problème est que cette caractérisation ne semble pas représenter adéquatement le type de conflit qui a lieu lorsque nous *renonçons à satisfaire une envie* pour accomplir un devoir moral. Dans la caractérisation proposée, en effet, nous ne renonçons tout simplement pas à satisfaire une envie : nous satisfaisons l'envie la plus forte, que celle-ci soit une envie morale ou une envie de glace. Il n'y a donc pas de conflit entre satisfaction des désirs et accomplissement du devoir moral. Or, lorsque nous sauvons une personne en danger, il semble bien que nous *renoncions* à la satisfaction de nos envies afin d'accomplir notre devoir d'aider la personne en détresse.

D'autres économistes ont préféré représenter la motivation morale d'une autre manière. Au lieu que la motivation morale soit traitée comme une préférence parmi d'autres, portant directement sur l'évaluation des états de choses, elle est considérée plutôt comme une *préférence de préférences*, ou *préférence de second ordre* (SEN, 1977). La

motivation morale n'est donc pas une préférence pour un *état de chose* ayant une certaine valeur morale, mais plutôt une préférence pour des préférences ayant une certaine valeur morale.

Par exemple, si une personne préfère être *une personne équitable*, alors elle choisira d'agir selon ses préférences de premier ordre qui sont compatibles avec ce désir d'équité – de second ordre. Cette personne préfère être le genre de personne qui a une fonction d'utilité de tel ou tel type (moral), mais les fonctions d'utilité sont elles-mêmes définies indépendamment de toute considération morale.

Pour illustrer cette idée, admettons par exemple qu'un individu ait le choix entre trois ordres de préférences différents (A, B, C) sur trois actions possibles (x, y, z):

$$A = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad B = \begin{pmatrix} y \\ x \\ z \end{pmatrix} \quad C = \begin{pmatrix} z \\ y \\ x \end{pmatrix}$$

Si notre individu a la préférence de second ordre d'être une personne équitable, alors il aura une préférence de second ordre pour les préférences de premier ordre qui respectent le critère d'équité. Admettons que, du point de vue de l'équité, le classement B est meilleur que A et que A est meilleur que C. Notre individu aura alors une préférence (de second ordre) pour des préférences du type :

$$\text{Préférence pour l'équité : } \begin{pmatrix} B \\ A \\ C \end{pmatrix}$$

Ce classement de classements de préférence correspond à sa motivation morale. Il s'agit d'une préférence pour les préférences qui respectent les exigences morales.

2.2.2. UNE MOTIVATION MORALE ÉGALITARISTE, KANTIENNE OU UTILITARISTE?

Un second problème, pour la formalisation de la motivation morale, concerne les diversités d'interprétation de cette motivation, en fonction du contenu substantiel de la morale. Par exemple, la manière de représenter la motivation morale varie selon qu'elle est considérée soit comme un souci d'*équité* ou d'égalité, soit comme une *aversion pour l'injustice*, soit comme un désir utilitariste de maximiser le bonheur du plus grand nombre ou encore comme la volonté kantienne de faire l'action qui serait la meilleure si tout le monde faisait la même action.

Prenons, par exemple, le cas où la motivation morale est un « goût pour l'égalité » ; elle est alors formalisée comme suit :

Désir d'égalité $U_i(x_i, x_j) = ax_i - \beta|x_i - x_j|$

L'utilité de l'individu i sur la distribution des biens (x_i, x_j) entre i et j est définie par la valeur variable a que l'individu accorde à sa propre situation matérielle, valeur de laquelle on soustrait la valeur β que l'individu accorde à la *différence* entre sa propre situation et celle de l'individu j . Plus β est élevée, plus le « goût pour l'égalité » de l'individu est fort.

D'autres fonctions d'utilité représentent d'autres valeurs morales. Certains économistes ont, par exemple, proposé de retranscrire le souci « kantien » de faire des actions qui, si elles étaient faites par tout le monde, donneraient lieu au meilleur état de choses. Cette « propension à faire l'action qui produirait le meilleur résultat pour soi-même, *si tout le monde* (soi-même compris) choisissait de faire cette action » est traduite par un facteur k dans la formule d'utilité suivante (CLAVIEN & CHAPUISAT, 2016, p. 26) :

L'homme moral $U_i(x_i, x_j) = (1 - k)w_i(x_i, x_j) + kw_i(x_i, x_i)$

Si le facteur k est faible, l'utilité de i augmente principalement en fonction de la partie gauche de l'addition, à savoir en fonction du bien-être de i dans une situation où les individus ne sont pas identiques $w_i(x_i, x_j)$. Par contre, plus la valeur de k est élevée, plus l'utilité de i dépend de la partie *droite* de l'addition, dans laquelle la situation des deux individus est identique $w_i(x_i, x_i)$.

Plus k est élevé, plus l'individu choisit les actions qui lui seraient bénéfiques *si* tout le monde agissait comme lui. Plus k est faible, plus la fonction d'utilité devient « égoïste ». Lorsque $k = 0$, on retrouve la fonction d'utilité de l'égoïste. Lorsque par contre $k = 1$, alors l'individu est un véritable kantien : il est motivé exclusivement à faire les actions qui donneraient le meilleur résultat pour lui, si tout le monde agissait de la même manière. Sa fonction d'utilité peut être formalisée de la manière suivante :

L'homme Kantien $U_i(x_i, x_j) = w_i(x_i, x_i)$

D'autres fonctions d'utilité sont encore possibles, en fonction de la théorie morale adoptée. L'individu pourrait, par exemple, avoir une préférence morale utilitariste ou parétienne.³⁴

Dans cette sous-section, nous avons vu que l'une des difficultés pour représenter la motivation morale provenait de la variété d'interprétations de la valeur morale. Si l'individu adopte une doctrine morale égalitariste, kantienne ou utilitariste, sa fonction d'utilité ne sera pas la même. Dans la section suivante, nous présenterons une autre difficulté pour la formalisation de la motivation morale : il existe des motivations proches, mais différentes, qu'il est difficile de distinguer dans le formalisme économique de la motivation morale.

2.3. DES MOTIVATIONS PROCHES MAIS DIFFÉRENTES

Une difficulté supplémentaire s'ajoute au problème précédent concernant le choix du formalisme ; la « motivation morale » se rapproche mais n'est pas identique à d'autres motivations que l'on trouve en économie. Ces motivations sont (i) le « sentiment chaleureux » ressenti lorsqu'on fait une bonne action et (ii) le désir « d'avoir et/ou de susciter une image morale positive de soi ». Dans la suite de cette section, nous présenterons ces deux désirs et tenterons de les distinguer, à la fois conceptuellement et empiriquement – i.e. par leurs effets observables sur le comportement – de la motivation morale.

2.3.1. L'EFFET DU « SENTIMENT CHALEUREUX »

Certains économistes, notamment (ANDREONI, 1990), ont distingué deux manières différentes dont les individus sont motivés à agir pour le bien public.³⁵ D'un côté, ils sont motivés à agir par pur souci du bien public. Dans ce cas, les économistes parlent d'« altruisme pur ». Pour un altruiste pur, la manière dont l'objectif moral est réalisé n'a pas

³⁴ Une situation est « efficiente » ou « optimale » dans le sens de Pareto lorsqu'il n'est pas possible d'améliorer la situation d'une personne sans détériorer la situation d'une autre personne. Au niveau social, une situation x est considérée comme meilleure qu'une situation y si (i) la situation d'au moins une personne est meilleure dans x que dans y et (ii) la situation d'aucune personne n'est moins bonne dans x que dans y . Le bien-être des personnes dans chaque situation se mesure par la satisfaction de leurs préférences (HAUSMAN & MCPHERSON, 2007, pp. 241-243). Un optimum de Pareto est donc « un état de l'économie x tel que les agents ne puissent s'accorder à lui préférer un autre état y » (FLEURBAEY & MONGIN, 1996, p. 296).

³⁵ La doctrine que nous présentons dans ces lignes concerne la provision de biens publics, mais nous pouvons l'étendre à l'ensemble des objectifs moraux envisageables.

d'importance ; l'altruiste pur est notamment indifférent au fait qu'il soit lui-même l'auteur de l'action qui a permis de réaliser l'objectif moral. Un altruiste pur préfère simplement la situation où le bien public/moral est réalisé à la situation dans laquelle le bien public/moral n'est pas réalisé.

A côté de cet altruisme « pur », les économistes ont distingué une autre manière par laquelle les individus sont motivés à réaliser un objectif moral : faire le bien leur procure un « sentiment chaleureux » (*warm glow*) agréable. Ce serait donc *dans le but de ressentir* ce sentiment chaleureux agréable que les personnes choisissent de poursuivre des objectifs moraux – au prix, parfois, d'un sacrifice d'une partie de leurs intérêts matériels. La particularité de cette motivation par le « sentiment chaleureux », par rapport à la motivation par altruisme « pur », provient du fait que sa force est fortement influencée par la *manière* dont l'objectif moral est réalisé. En effet, il n'est possible de ressentir le « sentiment chaleureux » d'avoir bien agi que si l'on est soi-même l'initiateur de l'action morale. Une personne motivée à agir dans le but de ressentir un tel sentiment chaleureux désire que le bien public soit réalisé *par elle-même*, par le biais d'une contribution *de sa part*. Sa motivation est donc, dans le sens général et non économique du terme, un désir égoïste ; le but poursuivi est d'obtenir *pour soi* la récompense du « sentiment chaleureux » d'avoir fait le bien.

La plupart des personnes partagent les deux motivations : elles désirent à la fois que le bien public soit réalisé – indépendamment de leur propre contribution – *et* qu'elles aient pu elles-mêmes contribuer à la réalisation de ce bien – ce qui leur procure un sentiment agréable d'avoir bien agi. Elles possèdent donc à la fois des motivations altruistes et des motivations égoïstes. Les économistes qualifient ces personnes d'altruistes « impurs ».

Du point de vue formel, la théorie de l'altruisme « impur » représente les motivations des individus – constituées d'un mélange d'altruisme « pur » et d'un désir de ressentir un « sentiment chaleureux » – de la manière suivante : les individus ont chacun une fortune w_i qu'ils dépensent soit pour leur consommation privée x_i , soit pour faire don g_i pour un bien public G . Leur fonction d'utilité a ainsi pour variables *la consommation privée* x_i , la réalisation du bien public G et la participation individuelle au bien public g_i . Elle prend donc la forme suivante :

L'altruisme « impur » $U_i = U_i(x_i, G, g_i)$

L'individu i a une fonction d'utilité qui a pour variable à la fois le bien public lui-même (G) et le don *fait par i* (g_i) pour ce bien public. Dans cette formule, il est possible d'isoler l'effet de l'altruisme pur et celui du « sentiment chaleureux ». L'altruiste « pur » est un individu qui, à côté du souci de sa consommation privée x_i , se soucie exclusivement de la provision ou non-provision du bien public G . Sa fonction d'utilité varie donc uniquement en fonction de la provision du bien public G . Il n'accorde pas d'importance au fait supplémentaire d'avoir lui-même contribué à la réalisation de G .

L'altruiste « pur » $U_i = U_i(x_i, G)$

A l'inverse, l'égoïste motivé par « le sentiment chaleureux » accorde uniquement de l'importance au fait de contribuer *lui-même* au bien public – puisque ce n'est que de cette manière qu'il peut ressentir le sentiment chaleureux lié à l'accomplissement de bonnes actions. Formellement, cela implique que sa fonction d'utilité ne varie qu'en fonction de sa propre contribution g_i et ne varie pas en fonction de la réalisation du bien public G :

Le « sentiment chaleureux » $U_i = U_i(x_i, g_i)$

On le voit, la fonction d'utilité de l'individu motivé par le « sentiment chaleureux » varie uniquement en fonction du montant du don que lui-même attribue au bien-public et ne varie pas en fonction de la quantité totale de bien public G .

En résumé, un individu qui se soucie exclusivement du bien public, sans accorder d'importance au fait qu'il ait participé lui-même à sa réalisation, est un individu purement altruiste. Un individu qui, de l'autre côté, ne se soucie du bien public que *dans la mesure* où il participe lui-même à sa réalisation est un individu « égoïste », dans le sens qu'il agit dans le but de ressentir le « sentiment chaleureux » de bien agir. Pour finir, l'individu qui se soucie à la fois du fait que le bien public soit réalisé *et* de ressentir le sentiment d'avoir bien agi est un « altruiste impur ».³⁶

Nous avons distingué d'un côté le désir de réaliser un objectif collectif ou moral et, de l'autre côté, le désir de ressentir le « sentiment chaleureux » accompagnant la réalisation de

³⁶ « Quand $U_i = U_i(x_i, G)$, l'individu ne se soucie pas du tout du gain privé et peut ainsi être conçu comme purement altruiste. De la même manière, lorsque $U_i = U_i(x_i, g_i)$, l'individu est motivé à donner uniquement pour le "sentiment chaleureux" et est de ce fait purement égoïste. Quand aussi bien G que g sont des arguments de la fonction, la personne est impurement altruiste. » [“(W)hen $U_i = U_i(x_i, G)$ the individual cares nothing for the private gift per se, hence can be thought of as purely altruistic. Likewise, when $U_i = U_i(x_i, g_i)$ the individual is motivated to give only by warm-glow, hence is purely egoistic. When both G and g are arguments, the person is impurely altruistic” (ANDREONI, 1990).]

l'objectif. Nous avons vu que le premier désir est insensible alors que le second est sensible au fait que l'individu lui-même soit l'auteur de la bonne action. Nous avons obtenu le critère de distinction empirique suivant : pour savoir si le sentiment de satisfaction est ou non le véritable motif (égoïste) de l'action morale, il faut voir si la valeur que l'individu attribue à la réalisation de l'action morale baisse dès lors qu'il n'est pas lui-même l'auteur de l'action morale. Si tel est le cas – i.e. si la valeur qu'il attribue diminue lorsque l'action morale est réalisée indépendamment de sa propre contribution – alors l'individu est motivé par le désir égoïste de « ressentir un sentiment chaleureux ». Si, par contre, la valeur qu'il attribue à l'action morale ne baisse pas dès lors qu'il n'est pas lui-même l'auteur de l'action, alors sa motivation n'est pas égoïste. – même si, de fait, il ressent *après coup* un sentiment de satisfaction d'avoir bien agi. Dans ce second cas, le sentiment de satisfaction est *l'effet* plutôt que *la cause* de l'action morale.

2.3.2. LE DÉSIR D'AVOIR OU DE SUSCITER UNE IMAGE POSITIVE DE SOI

Nous avons opéré une distinction entre, d'un côté, la motivation morale et, de l'autre, le désir de ressentir la joie ou « sentiment chaleureux » d'avoir fait quelque chose de bien, joie qui n'est accessible qu'à ceux qui sont eux-mêmes les acteurs de la réalisation d'objectifs moraux. Il existe encore un autre désir proche mais néanmoins différent de la motivation morale, à savoir le « désir d'avoir une image morale positive de soi-même » ou « désir de réputation morale ». Nous allons passer en revue deux manières dont les économistes ont représenté ce « désir d'avoir une image positive de soi-même ». Dans la première, ce désir d'image n'est pas considéré comme différent de la motivation morale « d'être une personne vertueuse ». Dans la seconde représentation, par contre, le désir d'image et la motivation morale sont considérés comme étant distincts. Nous soutiendrons que la seconde option est correcte, puisqu'il existe un critère empirique de distinction entre les deux types de motivation. L'analyse de la distinction entre la motivation proprement morale et le désir d'avoir ou de donner une image morale positive de soi – que nous pouvons caractériser comme un désir de « réputation morale » – sera l'objet du prochain chapitre.

3. LE DÉSIR DE RÉPUTATION MORALE

Dans ce chapitre, nous nous emploierons à présenter le fonctionnement de ce désir de « réputation morale » et nous montrerons qu'il est différent de la motivation morale. Nous procéderons pour ce faire de la manière suivante : en premier lieu, nous présenterons les conceptions de la motivation morale qui l'identifient au désir d'avoir une bonne image morale ou une bonne réputation morale (3.1.). Dans un deuxième temps, nous présenterons les conceptions qui, à l'inverse, considèrent que la motivation morale est différente du désir de réputation morale (3.2.). Dans la troisième section, nous avancerons des arguments en faveur de la seconde thèse, à savoir qu'il existe une différence – bien que difficile à saisir – entre les deux types de motivations (3.3.). Pour finir, nous répondrons à une objection (kantienne) qui pourrait être formulée contre la distinction que nous proposons (3.4.)

3.1. UN DÉSIR IDENTIQUE À LA MOTIVATION MORALE

Certains économistes semblent considérer la motivation morale et le désir « d'avoir une image morale positive de soi-même » comme étant identiques. Il nous semble que c'est le cas de de Brettke, Kverdockk et Nyborg dans leur article « An Economic Model of Moral Motivation ». Ces économistes décrivent la motivation morale dans les termes suivants : les consommateurs – ou travailleurs – préfèrent se considérer eux-mêmes comme étant des individus socialement responsables (BETTKE, KVERNDORKK, & NYBORG , 2003). La motivation morale correspond donc au désir de se percevoir soi-même comme étant une personne qui agit pour le bien commun. Formellement, les auteurs traduisent cette motivation de deux manières différentes dans deux articles, mais ces variations importent peu pour notre propos (BETTKE, KVERNDORKK, & NYBORG , 2003; 2004). Ils introduisent le « souci d'avoir une image positive de soi » dans la fonction d'utilité des individus de la manière suivante:

Le désir d'avoir une image positive de soi :

$$U_i = u(x_i, l_i, G, I_i)$$

ou

$$U_i = u(x_i, l_i, G) + S_i$$

Dans les deux cas, la fonction d'utilité de i varie en fonction de la situation économique de la personne x_i , de son temps libre l_i et de l'existence d'un bien public G . Les variables l_i et S_i représentent quant-à-elles le souci de la personne d'avoir une image morale positive d'elle-même. Plus la personne agit d'une manière qui est conforme à l'idéal moral qu'elle souhaite atteindre, plus son désir d'avoir une image morale positive d'elle-même est satisfait. Par exemple, si notre individu est partisan d'une morale utilitariste, alors l'image morale qu'il a de lui-même sera améliorée s'il agit d'une manière qui maximise le plus grand bonheur du plus grand nombre. Autrement dit, plus l'action de la personne se rapproche de l'idéal moral qu'elle s'est fixée, plus la valeur de la variable S_i ou l_i augmente et plus l'utilité U_i de la personne augmente.

Nous voyons que le désir d'avoir une image morale positive de soi-même, dénoté par I ou S , ne donnera pas lieu à la même action selon la théorie morale acceptée. Ainsi, un partisan de l'utilitarisme ou un partisan de la protection inconditionnelle des droits fondamentaux auront des préférences I ou S qui ne seront pas satisfaites de la même manière. Un individu utilitariste qui veut avoir une image morale positive de lui-même préférera les états du monde qui maximisent le bonheur du plus grand nombre alors qu'un individu déontologiste préférera les états du monde dans lesquels les droits fondamentaux sont protégés. Le contenu du désir d'avoir une image positive de soi dépend donc de la doctrine morale normative adoptée par les individus composant la société.

Ce qu'il nous faut retenir pour notre propos, c'est que la motivation morale n'est pas distinguée, dans la théorie de Bettke et Cie, du désir d'avoir une image morale positive de soi-même. Désirer avoir une image morale positive de soi-même n'est pas différent, selon eux, de vouloir être une personne qui fait le bien, qui agit justement ou qui est vertueuse. Pour l'instant, cela semble plausible.

3.2. UN DÉSIR DIFFÉRENT DE LA MOTIVATION MORALE

D'autres économistes établissent de leur côté une différence entre la motivation morale et le souci d'avoir une image positive de soi. Ils distinguent ainsi la motivation *intrinsèque* pour la valeur morale d'une action et la motivation *instrumentale* du souci de sa propre image. Tirole et Benabou (2006), par exemple, séparent le désir intrinsèque de

sacrifier une part de bien-être matériel à une cause morale, du désir *extrinsèque* ou égoïste d'avoir une image morale positive de soi-même.

Comment distinguer la motivation morale *intrinsèque* du désir *extrinsèque* d'avoir une image morale positive de soi-même ? A première vue, la tâche semble difficile : toute action extérieurement conforme à la morale semble pouvoir être expliquée *aussi bien* par la motivation morale que par le désir d'avoir une image morale positive de soi. Il semble, en effet, qu'à chaque fois que l'on peut expliquer une action contraire à l'intérêt personnel à l'aide de la motivation morale, on peut *aussi* l'expliquer par le fait que la personne a le désir d'avoir une image morale positive d'elle-même.

Considérons l'exemple d'une personne qui fait don d'une part non négligeable de sa fortune à une œuvre de charité. Il est possible que sa motivation soit purement morale : elle cherche à faire le bien, à faire ce qui est juste ou à être une personne vertueuse. Toutefois, il est toujours possible, également, qu'elle soit motivée par le désir d'avoir une image morale positive d'elle-même ; en faisant don de sa fortune, elle acquiert l'image d'elle-même comme étant une bonne personne, une personne qui « fait le bien », une personne vertueuse. Ainsi, pour toute action conforme aux exigences morales, nous ne pouvons jamais exclure que cette action soit en fait motivée par le « désir intéressé d'avoir une image morale positive de soi », plutôt que par la volonté désintéressée de faire le bien.³⁷

Il ne semble donc pas possible de séparer *dans la réalité* les effets d'une motivation proprement morale des effets d'un désir d'avoir une image morale positive de soi-même. Pourtant, Tirole et Bénabou proposent une manière convaincante de séparer les deux types de motivations en fonction de leurs *effets observables* sur le comportement. Pour comprendre comment ils procèdent, il nous faut commencer par présenter le « désir d'avoir une image *morale* positive de soi-même », puis nous le distinguerons de la motivation morale désintéressée.

³⁷ Notons toutefois, comme le fait remarquer Jon Elster, qu'il est douteux qu'une personne puisse être *consciemment* motivée à faire l'action morale *dans le but d'avoir une image morale positive d'elle-même*. Il est impossible qu'une personne fasse l'action morale *dans le but conscient d'être une personne qui agit par souci de la morale*, puisqu'une telle personne sait qu'elle n'agit pas par souci de la morale mais en vue de parfaire de sa propre image. « [L]'idée d'une bonne action dont la motivation *efficace* et *consciente* serait l'approbation intime est incohérente. » (ELSTER, 2009, p. 28).

3.2.1. LA VARIÉTÉS DES DÉSIRS LIÉS À L'IMAGE MORALE DE SOI

Le désir d'avoir une image morale positive de soi est considéré par Tirole et Bénabou comme étant une *variante* d'un autre désir, à savoir le désir de *susciter chez autrui* une image morale positive de soi. Autrement dit, le souci d'avoir soi-même une image morale positive de soi fonctionne de la même manière que le souci de susciter *chez autrui* une image morale positive de soi. La seule différence est la suivante : alors que le spectateur ou *évaluateur* de notre action est *quelqu'un d'autre* dans le second cas, ce spectateur est *nous-mêmes* dans le premier cas. Nous adoptons à notre propre égard l'attitude d'un spectateur externe et nous formons une conception de nous-mêmes et de notre propre valeur sur cette base.³⁸

Tirole et Bénabou distinguent quatre types de désirs « d'image morale » en fonction de deux variables: en premier lieu, ils distinguent les désirs d'image en fonction du fait qu'ils portent sur l'image morale que nous avons de nous-mêmes ou sur l'image morale que les *autres* ont de nous-mêmes. En second lieu, ils distinguent les désirs d'image selon qu'ils consistent à *promouvoir une image positive* ou à *éviter de donner* une image négative. Dans un tableau à double-entrée, cela donne lieu à quatre variantes du « désir d'image» concernant nos propres aptitudes morales.

	Soi-même	Autrui
Image positive Par ex. « être une personne prosociale »	Désir d'avoir <i>soi-même</i> une image morale <i>positive</i> de soi	Désir de susciter <i>chez autrui</i> une image morale <i>positive</i> de soi
Image négative Par ex. « ne pas être une personne cupide »	Désir de ne pas avoir de <i>soi-même</i> une image morale <i>négative</i>	Désir de ne pas susciter chez <i>autrui</i> une image morale <i>négative</i>

Ces quatre désirs sont formalisés par Tirole et Bénabou de la manière suivante :

Le désir de donner à autrui une image « prosociale » de soi-même

$\mu_a = x\gamma_a$ où x dénote la probabilité d'être vu et γ_a le désir d'apparaître à autrui comme *quelqu'un de « prosocial »*

³⁸ « De la même manière que les individus se soucient de l'opinion des autres, ils se soucient de l'image qu'ils ont d'eux-mêmes. » [“[A]s much as people care about the opinion others have of them, they care about their self-image”. (TIROLE & BENABOU, 2006, p. 1653).]

Le désir de ne pas apparaître à autrui comme « cupide »

$\mu_v = x\gamma_v$ où x dénote toujours la probabilité d'être vu et γ_v le désir de *ne pas apparaître comme cupide*.

Le désir d'avoir une image « prosociale » de soi-même

$\mu_a = x\gamma_a$ où x dénote la probabilité de se souvenir de son action et γ_a le désir d'apparaître à nous-mêmes comme étant quelqu'un de « prosocial »

Le désir de ne pas apparaître à soi-même comme « cupide »

$\mu_v = x\gamma_v$ où x dénote la probabilité de se souvenir de son action et γ_v le désir de *ne pas apparaître à soi-même comme étant quelqu'un de cupide*.

Pour Tirole et Bénabou, le désir que « je me perçoive moi-même comme étant une bonne personne » est donc une variante du désir que « les autres me perçoivent comme étant une bonne personne » et le désir « de ne pas apparaître à soi-même comme étant cupide » est une variante du désir « de ne pas apparaître à autrui comme étant cupide ». La seule différence entre les deux types de désirs provient de la nature des probabilités liées à leur satisfaction : alors que le désir de susciter *chez autrui* une certaine image morale de soi est sensible à la probabilité d'être vu par autrui, le désir d'avoir soi-même une certaine image morale de soi est sensible à la probabilité de se souvenir de son propre comportement. Etant donné que c'est, selon les deux économistes, par le souvenir de nos propres actions que nous nous forgeons une conception positive ou négative de nous-mêmes, plus la probabilité de nous souvenir de notre action est élevée, plus il est important pour nous de bien agir.

3.2.2. LA PARTICULARITÉ DES DÉSIRS LIÉS À L'IMAGE MORALE DE SOI

L'ensemble de ces désirs concernant notre image ou réputation morale ont pour particularité que la motivation à bien agir *diminue* (plutôt qu'elle n'augmente) lorsque des récompenses sont rattachées aux bonnes actions. Autrement dit, plus le comportement moral est récompensé, que ce soit par des gains monétaires ou d'autres types de récompenses, moins les personnes sont motivées à contribuer à l'objectif moral. Ce processus s'explique de la manière suivante :

(i) Le désir de ne pas apparaître à autrui comme étant quelqu'un de cupide

Le fait qu'une récompense monétaire se trouve rattachée à l'accomplissement de l'action moralement appropriée diminue la motivation à contribuer à l'objectif moral, car la personne craint que, étant donnée la présence de la récompense, son comportement soit interprété comme motivé par la cupidité. Étant donné que la personne désire ne pas apparaître comme cupide aux yeux d'autrui, la récompense monétaire rattachée à l'accomplissement de l'action morale réduit plutôt qu'elle n'augmente la motivation de la personne à contribuer aux objectifs moraux ou collectifs.

(ii) Le désir d'apparaître à autrui comme étant quelqu'un de « prosocial »

Le fait qu'une récompense soit rattachée à l'action morale fait perdre une occasion de donner à autrui l'image d'être quelqu'un qui se soucie particulièrement du bien social. Plus la récompense est élevée, moins le comportement sera interprété par les autres comme étant motivé par le seul souci du bien public et moins le comportement moral sera intéressant pour celui qui désire donner une image morale positive de lui-même à autrui.

(iii) Le désir de ne pas apparaître à soi-même comme étant quelqu'un de cupide

A nouveau, l'introduction de *récompenses* monétaires favorables à l'accomplissement de l'action morale peut avoir un impact négatif sur le taux de contribution d'une personne à l'objectif moral. En effet, plus la récompense monétaire est élevée, plus la motivation monétaire devient évidente, plus l'individu interprète son propre comportement comme étant motivé par la cupidité plutôt que par la seule considération de la valeur morale de l'action.

(iv) Le désir d'apparaître à soi-même comme étant quelqu'un de « prosocial »

Même si l'individu n'interprète pas encore son propre comportement comme étant cupide, l'introduction des récompenses amenuise quand même le caractère « prosocial » de son comportement, diminuant de ce fait le désir d'agir moralement. L'introduction de la récompense supprime, à nouveau, une occasion pour l'individu d'avoir une image de lui-même, au travers de la mémoire de ses propres actions, comme étant quelqu'un de soucieux du bien public.

Les économistes ont constaté les effets de ces différents désirs « d'image morale » sur le taux de dons de sang, par exemple ; rattacher une récompense monétaire au don de sang

a tendance à faire diminuer plutôt qu'à faire augmenter les contributions totales. Ils expliquent ce phénomène de la manière suivante : les personnes qui accordent une grande importance à leur image morale cessent d'être motivées à donner leur sang lorsque elles sont récompensées pour leur bonne action, car elles craignent que leur comportement soit interprété (par elle-même ou autrui) comme étant motivé par la cupidité. Dans les données empiriques récoltées par les économistes, il s'avère que les femmes accordent particulièrement de l'importance au fait de donner (et d'avoir) une image d'elles-mêmes comme étant des êtres altruistes et bienveillants, alors que les hommes restent plutôt indifférents à la présence ou à l'absence de gains monétaires récompensant l'action morale (TIROLE & BENABOU, 2006, p. 1666). Autrement dit, les donations de sang par des femmes diminuent *fortement* avec l'introduction de récompenses alors que les donations de sang par des hommes restent plutôt constantes malgré la présence de récompenses. Cela montre que les femmes, plus que les hommes, se soucient de donner et d'avoir une image morale positive d'elles-mêmes.

Ainsi les désirs portant sur l'image morale de soi ont un effet visible lorsque des *récompenses non morales*, notamment monétaires, sont rattachées à l'accomplissement de l'action morale ; au lieu que l'instauration de récompenses *augmente* la contribution des individus à l'objectif moral, c'est le contraire qui est observé ; l'instauration des récompenses fait *diminuer* les contributions individuelles. Cela s'explique par le fait que la présence de récompenses soit (i) fait disparaître une occasion pour les individus d'apparaître à eux-mêmes ou à autrui comme étant motivés par des objectifs sociaux, soit (ii) risque au contraire de faire apparaître leur comportement à leurs propres yeux ou aux yeux d'autrui comme étant motivé par la cupidité.

Nous avons, dans cette section, présenté le fonctionnement du « désir de réputation morale » susceptible de motiver ou de démotiver une personne à accomplir l'action morale. Avant de conclure cette première partie de la recherche, axée sur la présentation *des trois grands types* (égoïsme, altruisme, morale) de motivations humaines, nous devons encore clarifier la différence entre la motivation morale proprement dite, c'est-à-dire désintéressée, et le désir *intéressé* d'« avoir ou de donner une image morale positive de soi ». Nous tenterons de clarifier cette différence dans la sous-section suivante.

3.2.3. LE DÉSIR D'ÊTRE UNE PERSONNE VERTUEUSE ET LE DÉSIR D'AVOIR UNE IMAGE POSITIVE DE SOI

Nous avons qualifié de *motivations morales* trois désirs différents : les désirs (a) de réaliser le *bien*, (b) de faire ce qui est *juste* et (c) d'être une personne vertueuse. Le désir de réaliser le bien et de faire ce qui est juste ne posent pas de problèmes conceptuels particuliers. Ils correspondent de manière peu problématique à ce que nous avons qualifié de « motivation morale désintéressée ».

Par contre, le désir d' « être une personne vertueuse » est difficile à distinguer du désir d' « avoir une image morale positive de soi-même ». En quoi, en effet, le désir d'être une personne vertueuse, caractérisant la motivation morale *désintéressée*, est-il différent du désir *intéressé* d'avoir une image morale positive de soi-même ? Ne s'agit-il pas tout simplement, de la même motivation ? Désirer « être une personne vertueuse » ou désirer « avoir une image morale positive de soi-même », n'est-ce pas la même chose ?

Sans entrer pour le moment dans le débat sur *l'évaluation* morale de ces différentes motivations (désintéressées et intéressées), nous proposons de les distinguer de la façon suivante : le désir désintéressé d' « être une personne vertueuse » partage avec les motivations morales de réaliser le bien ou de faire ce qui est juste, la propriété de ne pas être sensible à la présence d'une récompense non morales – notamment monétaire ; la force de la motivation morale d'une personne qui cherche à réaliser le bien, à faire ce qui est juste, ou à être une personne vertueuse, n'est pas affectée par le fait que l'action morale soit récompensée par des gains non moraux. Au contraire, si une personne accorde véritablement une valeur désintéressée à la morale, alors sa motivation à faire l'action morale devrait être *renforcée* par la présence d'une récompense monétaire ; en effet, au lieu d'avoir une seule motivation (morale), la personne possède *deux motivations* indépendantes qui la motivent à faire l'action morale. Sa motivation morale est donc renforcée par la motivation non morale.

Par contre, dans le cas du souci d' « avoir une image morale positive de soi-même », la présence d'une récompense non morale a un impact *négatif* sur la motivation à faire l'action morale ; avec le désir d'avoir une image positive de soi, plus la récompense est élevée,

moins nous sommes assurés d'être des individus sincèrement soucieux de morale, moins nous serons motivés à faire l'action morale.³⁹

Dans la suite de ce travail, nous considérerons donc que les deux types de motivations sont différents : d'un côté, la motivation morale consistant à vouloir être une personne vertueuse. Cette motivation est insensible à la présence de récompenses monétaires encourageant le comportement moral. De l'autre côté, la motivation consistant à vouloir avoir soi-même une image morale positive de soi ; cette motivation n'est qu'une variante de la motivation consistant à vouloir susciter une image morale positive de soi-même *chez autrui*. Cette motivation est fortement sensible à la présence de récompenses non morales encourageant le comportement moral : plus les récompenses augmentent, *moins* la personne est motivée à faire l'action morale, car plus grand est le risque que son comportement soit mal interprété.

3.2.4. NOTE SUR LA CONCEPTION KANTIENNE DE LA MOTIVATION MORALE

Dans la section précédente, nous avons considéré comme distincts la motivation *désintéressée* d'être une personne vertueuse et le désir *intéressé* d'avoir une image morale positive de soi. On pourrait nous reprocher que cette distinction n'est pas compatible avec une représentation largement répandue de la valeur morale des actions : la théorie kantienne. Nous soutenons pourtant que cette distinction, entre la motivation morale et le

³⁹ Il nous semble que la distinction entre la motivation morale *désintéressée* d'être une personne vertueuse et le désir *intéressé* d'avoir une image morale positive de soi-même pourrait être expliquée également par l'opposition suivante : la première motivation est identique au désir d'être *digne* d'approbation alors que la seconde est identique au désir d'être simplement *approuvé* (par soi-même ou autrui). Cette distinction entre le fait de désirer *mériter* l'approbation morale et le fait de désirer seulement *obtenir* l'approbation morale expliquerait d'ailleurs pourquoi la première motivation, mais non la seconde, est insensible à la présence d'une récompense non morale, notamment monétaire. A noter que celui qui se soucie d'être *digne d'approbation* peut, bien sûr, également apprécier le fait *d'être* approuvé – mais seulement lorsque cela est mérité. A l'inverse, celui qui se soucie uniquement *d'être approuvé* ne se soucie pas que cette approbation soit méritée ou non. Il désire uniquement donner une image morale positive de lui-même.

Ces distinctions rappellent bien entendu la discussion d'Adam Smith sur les différentes formes du désir d'approbation. Selon Smith, le désir d'obtenir l'approbation *méritée* – qui se distingue à la fois du désir d'obtenir l'approbation (cher à Mandeville) et du désir de *mériter* l'approbation (cher aux Stoïciens) – constitue le principal moteur de l'action humaine. Il semblerait même que ce désir *d'approbation méritée* soit considéré par Adam Smith, contrairement aux interprétations traditionnelles, comme étant à l'origine de l'échange et du commerce (BEE, (à paraître)). Si cette interprétation d'Adam Smith est correcte, cela signifierait que, selon le philosophe, la société commerçante dépend, dès son origine, des aptitudes morales des citoyens. En effet, c'est le souci du *mérite*, le souci de l'approbation *méritée* des autres, qui explique l'apparition de la société marchande. Dans une société composée exclusivement d'égoïstes, désireux d'obtenir des avantages indépendamment du mérite, la division du travail n'aurait simplement pas été possible. Ce n'est donc pas l'intérêt personnel, mais les sentiments moraux, qui sont à l'origine de la société commerçante.

désir d' « avoir une image morale positive de soi-même », est compatible même avec la conception kantienne de la valeur morale des actions.⁴⁰

Voilà ce que soutiendrait notre objecteur contre la distinction : dans la théorie kantienne, la valeur morale d'une action dépend fondamentalement de la motivation dont elle résulte. Si un agent accomplit son devoir moral dans le but d'accomplir son devoir, alors son action a une valeur morale positive. Par contre, s'il accomplit son devoir moral en vue de réaliser un objectif non moral, alors son action n'a pas de valeur.

[P]our qu'une action soit moralement bonne, il ne suffit pas qu'elle soit conforme à la loi morale, mais il faut qu'elle soit faite en vue de cette loi (KANT, 2007 (1785)).⁴¹

Il semblerait donc bien que, dans la théorie kantienne, une *bonne personne* ou une personne *vertueuse* est une personne qui accomplit son devoir *uniquement* parce qu'il s'agit de son devoir. Une personne vertueuse a donc raison, dans la théorie kantienne, de considérer que la présence d'une récompense *non morale* (par exemple monétaire) *entache* la valeur morale de son action. Elle a ainsi également raison d'être moins motivée à faire son devoir moral lorsque l'accomplissement de son devoir est récompensé par un gain non moral.

Par exemple, si un directeur d'entreprise décide d'investir dans de l'énergie renouvelable pour la *seule* raison qu'il s'agit de l'action correcte du point de vue moral, alors son action a une valeur morale positive. Par contre, s'il investit dans l'énergie renouvelable *également* parce que cela lui permet de faire plus de profit, la théorie kantienne implique, selon notre objecteur, que la valeur morale de son action se trouve amoindrie. Autrement dit, dans la théorie kantienne, seul un individu qui cherche à faire l'action morale *uniquement* par souci de la qualité morale de cette action est un individu véritablement vertueux. Or, cela signifie qu'une personne véritablement vertueuse devrait refuser de réaliser l'action morale si celle-ci est récompensée par des gains monétaires. La motivation morale *désintéressée* d' « être une personne vertueuse » est donc très sensible, dans la

⁴⁰ Nous rejoignons par notre argument la thèse de Dubbink et Liedekerke, selon laquelle la morale kantienne ne condamne pas les actions faites *à la fois* par souci de la morale *et* par souci de son propre intérêt (DUBBINK & van LIEDEKERKE, 2019).

⁴¹ „Denn bei dem, was moralisch gut sein soll, ist es nicht genug, dass es dem sittlichen Gesetze gemäss sei, sondern es muss auch um desselben willen geschehen“ (KANT, 2016 (1785), p. 14).

théorie kantienne, à la présence de gains non moraux, notamment monétaires. En conséquence, selon notre objecteur, la distinction que nous avons proposée dans la section précédente, entre, d'un côté, le désir d'être une personne vertueuse et, de l'autre, le désir d'avoir une image morale positive de soi, ne tient pas. La force motivationnelle de l'une comme de l'autre est diminuée lorsque des récompenses sont rattachées à l'action morale.

Cependant, cette objection repose sur une interprétation erronée de la théorie kantienne de la valeur morale des actions. Dans la théorie kantienne, le fait qu'une récompense non morale vienne *renforcer* la motivation globale, dont la motivation morale est la composante essentielle, n'affecte pas la valeur morale de l'action ou de la personne. L'important, pour Kant, est que la considération de la rectitude morale – i.e. conformité au devoir – de l'action soit *suffisante* pour motiver la personne à agir. Il n'est pas besoin que cette considération soit la *seule* qui le motive à agir.

Reprenons le cas du directeur d'entreprise. Si celui-ci est toujours motivé à investir dans le renouvelable même *en l'absence* d'un gain économique supplémentaire, le fait que sa motivation morale soit accompagnée d'une motivation économique ne diminue pas la valeur morale de son action. Si la motivation morale est suffisante pour motiver l'individu à agir, alors son action possède une valeur morale positive – même si cette motivation se trouve accompagnée, dans les faits, par d'autres motivations non morales. Par contre, si la considération de la qualité morale de l'action est *insuffisante* pour motiver l'individu à agir, c'est-à-dire que l'individu ne serait pas motivé à agir moralement en l'absence d'incitations économiques, alors, dans ce cas seulement, son action ne possède pas de valeur morale – ce qui n'implique pas encore que son action soit moralement *mauvaise*.

La présence d'une motivation *supplémentaire* à la seule motivation morale n'affecte donc pas nécessairement, dans la théorie kantienne, la valeur morale de l'action ou de la personne qui agit.⁴² Cela a énormément d'importance pour notre propos : un partisan de la morale kantienne, qui considère qu'une personne n'est vertueuse que si sa motivation

⁴² Comme le fait en outre remarquer Barbara Herman (1981), il existe même des circonstances dans lesquelles la motivation à faire son devoir moral ne *peut pas* avoir d'effet en l'absence d'une motivation non morale. C'est le cas notamment pour toutes les actions *moralement permises* : une personne décide d'accomplir l'action *après avoir vérifié* que son action est conforme au devoir (motif moral), mais le *but* visé par l'action moralement permise ne peut pas être uniquement de faire une action parce que cette action est moralement *permise*. Il est donc nécessaire que soit présente une motivation non morale pour que la motivation à « faire ce qui est juste » puisse causer l'action dans les situations où l'action n'est pas obligatoire mais seulement permise.

morale est *suffisante* pour motiver l'action, n'a pas besoin de soutenir que la présence d'une autre motivation, qui accompagne ou *complémente* la motivation morale, ait une répercussion négative sur la valeur morale de son action. Ainsi, un véritable kantien ne devrait pas considérer que la valeur morale de son action *diminue* en présence d'incitations économiques. Un entrepreneur kantien ne devrait pas être moins motivé à investir dans l'énergie renouvelable du fait que cet investissement lui permet aussi d'augmenter son profit ou du fait que, par hasard, le gouvernement a décidé de soutenir économiquement ce genre d'investissement. Aussi longtemps que la valeur morale de l'action *suffirait* à le motiver à investir dans le renouvelable, il peut considérer que son action possède une valeur morale positive.

Nous pouvons ajouter encore à ce commentaire que, pour Kant, la valeur morale d'une action n'est pas déterminée *exclusivement* par la motivation morale de l'individu : il faut, avant toute chose, que l'action soit conforme au devoir. Décider de ne pas faire une action pourtant conforme au devoir par peur que cela donne une image de soi comme étant motivé par l'argent ou la cupidité – à soi ou à autrui – revient à décider de ne pas faire son devoir *du tout*. Or, dans ce cas, non seulement l'action manque de valeur morale, puisque la motivation morale n'est pas suffisante, mais elle est surtout *contraire* au devoir – d'investir dans le renouvelable. Dans ce cas, notre action possède une valeur morale *négative* : elle n'est pas seulement l'action d'une personne dépourvue de vertu, mais elle est l'action d'une personne *vicieuse*.

La motivation des individus véritablement kantien devrait donc, dans tous les cas, rester inaffectée par la présence de récompenses non morales. Kant va d'ailleurs parfois même plus loin : il affirme que nous ne devons *jamais* décourager les comportements moraux, même lorsque l'on *soupçonne* que le motif n'est pas proprement moral, et même lorsque l'on *sait* que le motif n'est pas proprement moral. Ainsi, même si l'on *sait* que les motivations réelles d'un individu ne sont pas celles qu'il prétend, même si l'on sait que, sous couvert de faire le bien, un individu vise à satisfaire ses propres intérêts, il faut continuer à encourager son comportement conforme à la morale et il faut continuer à encourager, chez lui, l'usage *public* du langage de la vertu.⁴³

⁴³ François Calori explique la position kantienne en ces termes: « Au déploiement sans fard et cynique de l'exercice brut de la force, il vaut toujours mieux préférer celui qui se pare des oripeaux du droit et lui rend

4. CONCLUSION

Dans ce chapitre, nous avons distingué trois types de motivations humaines :

- L'égoïsme
- L'altruisme
- La motivation morale

Nous avons également passé en revue la manière dont ces motivations se trouvent modélisées en économie.

En ce qui concerne la motivation morale, nous avons pris soin de clarifier sa nature en la distinguant de deux motivations égoïstes qui s'en rapprochent : (1) le « sentiment chaleureux » procuré par le fait de bien agir et (2) le désir d'« avoir une image morale positive de soi ». Nous avons vu que la joie procurée par le fait de faire des bonnes actions ne motive une personne à poursuivre un objectif moral que si la personne elle-même contribue à la réalisation de cet objectif. De son côté, le désir d'avoir une image morale positive de soi a pour caractéristique que, lorsque le comportement moral est *récompensé* par des gains non moraux, la personne est *moins motivée* à faire l'action morale.

ainsi un hommage involontaire et salutaire, au moment où il le foule aux pieds. C'est ainsi, par exemple, que l'on peut interpréter l'invocation par les grandes puissances des principes des grands jurisconsultes au moment de justifier une déclaration de guerre, alors même qu'aucun de ces principes ne les a par ailleurs jamais conduits à abandonner une seule de leurs visées guerrières. Plutôt qu'en dénoncer l'hypocrisie, il faut y lire la puissance que l'idée de droit exerce, là même où elle n'est pas respectée, et craindre le moment où elle ne serait plus que tournée en dérision : « Cet hommage que tout État rend (du moins en paroles) au concept de droit prouve malgré tout que l'on peut trouver en l'homme une disposition morale plus haute encore, bien qu'en sommeil pour l'instant, grâce à laquelle il pourra un jour dominer le mauvais principe qui est en lui (et qu'il ne peut nier) et en espérer autant des autres. » [Ak VIII 355] » ou encore « Celui qui dénonce cyniquement et sans contrepartie l'apparence de la vertu s'interdit et interdit autour de lui la vertu véritable. Ce serait une trahison contre l'humanité. Mieux vaut affecter de croire à l'apparence de vertu, car c'est déjà croire d'une certaine manière à sa possibilité, et l'apparence est ici l'hommage du vice à la vertu. Mieux vaut de la petite monnaie que pas de monnaie du tout. » ou encore « il peut y avoir également un effet de moralisation par l'apparence, induit en l'autre par le respect que nous lui témoignons, quand bien même il ne le mériterait pas : « C'est donc un devoir de vertu que de jeter le voile de la philanthropie sur les fautes des autres, non seulement en adoucissant nos jugements, mais aussi en les passant sous silence : car les exemples de respect que nous donnons aux autres peuvent aussi susciter en eux l'effort pour le mériter également. C'est la raison pour laquelle la manie d'épier les mœurs d'autrui (allotrioepiscopia) constitue déjà par elle-même une curiosité offensante en matière de connaissance des hommes, à quoi chacun peut s'opposer à bon droit comme à une violation du respect qui lui est dû. [Ak VI 466.] » Il nous faut prendre ainsi conscience de la possibilité qu'une alchimie morale aboutisse à la conversion de la vile monnaie en or pur, « quand bien même c'est avec une perte considérable ». L'idée revient à plusieurs reprises dans le paragraphe : au jeu de la vertu, les hommes se laissent prendre et en viennent à les adopter véritablement. Celui qui joue le vertueux peut ne pas tromper l'autre, mais, sans le vouloir, il peut produire en lui-même des effets profonds : « Dans la mesure où des hommes jouent ces rôles, les vertus dont ils ont ainsi simplement, pour un certain temps, produit l'apparence se sont peu à peu véritablement éveillées et s'intègrent dans la disposition d'esprit. [Ak VII 151] ». (CALORI, 2011).

Par contraste, la force de la motivation proprement morale ne dépend ni du fait d'être soi-même l'auteur de la bonne action, ni du fait que l'action soit récompensée par des gains non moraux. Autrement dit, une personne véritablement soucieuse de morale ne se soucie ni du fait que l'action morale soit *la sienne*, ni du fait que l'action morale soit récompensée par des gains non moraux. Une personne qui veut faire le bien, faire ce qui est juste ou qui cherche à être une personne vertueuse est *indifférente* au fait d'être l'auteur de l'action morale et au fait d'être récompensée pour son action ; sa motivation morale n'est influencée ni positivement, ni négativement, par les propriétés non morales de l'action.⁴⁴

⁴⁴ Bien entendu, la plupart des personnes partagent plusieurs motivations, la motivation morale ne représentant que l'une des motivations parmi d'autres. La caractérisation proposée sert uniquement à *distinguer* les effets de la motivation morale sur le comportement par rapport aux effets d'autres motivations.

LA MOTIVATION MORALE EN MÉTAÉTHIQUE

1. LE PROBLÈME DE LA MOTIVATION MORALE DANS LA MÉTAÉTHIQUE CONTEMPORAINE

Le présent travail porte sur le besoin, pour toute société, que les individus soient capables de renoncer à leurs intérêts propres lorsque les satisfaire s'oppose à la réalisation de ce qu'ils jugent être le bien, la justice ou la vertu. Autrement dit, la question consiste à déterminer si la réalisation d'une société dépend du fait que ses membres soient capables de motivation proprement morale.

Or, il existe un débat métaéthique sur la question de l'existence et de la nature d'une motivation spécifiquement morale. Il nous paraît donc opportun de présenter et de discuter ce débat. Nous expliquerons néanmoins les raisons pour lesquelles cette discussion ne nous aide pas particulièrement dans notre recherche ; en premier lieu, (1) le débat est durablement enlisé. En second lieu, (2) la question de la motivation morale émerge dans le débat opposant le cognitivisme et le non-cognitivisme, qui n'est pas l'objet de notre recherche. Pour finir, (3) même si la solution s'avérait nous être favorable, elle ne nous offrirait pas une conclusion suffisamment forte pour étayer notre thèse, à savoir qu'il existe une motivation morale catégorique, universelle et toute-puissante, susceptible de remplacer l'intérêt personnel comme ressource motivationnelle fondamentale dans nos théories politiques.

1.1. LE CONSTAT DE FAIT

Le problème de la motivation morale en métaéthique contemporaine consiste à expliquer un phénomène empirique largement attesté et admis par l'ensemble des positions du débat. Ce constat empirique concerne la relation entre, d'un côté, les jugements moraux et, de l'autre, la motivation à agir. Il s'agit du constat suivant : en général, un jugement moral est accompagné d'une motivation à agir en conformité avec ce jugement. Ce constat peut être décomposé en deux observations empiriques distinctes.

(i) Les individus sont motivés à agir d'après leurs jugements moraux.

De manière générale, nous pouvons observer que les jugements moraux affectent la motivation à agir. Une personne qui juge que l'avortement est immoral sera motivée, en général, à adopter un certain comportement à l'égard de l'avortement, par exemple à critiquer les femmes qui subissent un avortement ou à manifester publiquement contre l'avortement. Il est en outre probable qu'une telle personne refuserait de subir elle-même un avortement.

Plus encore, nous considérons en général qu'un individu qui n'est pas motivé à agir d'après son jugement moral n'est pas *sincère* dans son jugement. Par exemple, si Georges *affirme* qu'il est immoral de manger de la viande, mais néanmoins mange régulièrement de la viande, sans éprouver de culpabilité, nous considérerions que ses affirmations sont *hypocrites*. Bien qu'il *affirme* qu'il est immoral de manger de la viande, il ne le *pense pas* vraiment. Autrement dit, il n'a pas de véritable conviction morale.

(ii) Les changements de convictions morales donnent lieu à des changements de motivation et de comportement

La seconde observation empirique, concernant la relation entre jugements moraux et motivation à agir, porte sur les *changements* de convictions morales. Dans ce contexte, l'observation est la suivante : lors d'un *changement* de jugement moral, on observe de manière fiable et régulière un changement dans les motivations à agir.

Considérons l'exemple suivant : d'abord, Christine croit que l'avortement est immoral. Son comportement est guidé par ce jugement: elle formule des reproches aux femmes qui avortent, elle descend dans la rue manifester contre l'avortement, elle vote pour les partis conservateurs. Après avoir discuté avec des amis et étudié la question, elle change son jugement: elle ne considère plus que l'avortement est immoral. On peut alors raisonnablement s'attendre à ce qu'elle cesse de manifester dans la rue, ne formule plus de reproches aux femmes qui avortent et, éventuellement, ne vote plus pour les partis conservateurs. Le changement dans ses convictions morales s'accompagne d'un changement dans ses motivations et dans son comportement (ROSATI, 2016).

Nous constatons donc deux phénomènes : en premier lieu, l'observation que, en général, le jugement moral est accompagné d'une motivation à agir et que, s'il ne l'est pas,

le jugement est considéré comme hypocrite. En second lieu, l'observation que des changements dans les convictions morales donnent lieu, de manière fiable et régulière, à des changements dans la motivation à agir et dans le comportement.

Le constat empirique de ce double phénomène doit pouvoir être expliqué par toute théorie métaéthique. C'est à ce niveau que les oppositions théoriques apparaissent.

1.2. L'EXPLICATION INTERNALISTE

Le problème consiste ainsi à expliquer l'observation empirique d'une relation entre les jugements moraux et la motivation à agir. Une explication immédiate de cette observation est que les jugements moraux sont *nécessairement* motivants, qu'il existe une connexion *nécessaire* entre les jugements moraux et la motivation à agir. C'est ce que soutiennent les partisans de *l'internalisme du jugement moral*. Selon cette doctrine, une personne qui n'éprouve pas de motivation à agir conformément à ce qu'elle affirme être sa conviction morale n'a tout simplement pas de réelle conviction morale. Elle ne pense pas véritablement ce qu'elle dit. Une personne qui affirme « il est immoral de manger de la viande » sans être motivée à renoncer à la consommation de viande fait un mauvais usage des concepts moraux, car elle n'exprime pas, par l'usage de ces concepts, un véritable jugement moral, une conviction morale sincère. Selon l'internalisme du jugement moral, exprimer un jugement moral sincère implique nécessairement d'être motivé à agir.

Plus formellement, l'internalisme du jugement moral, dans sa version la plus basique, se traduit de la façon suivante (ROSATI, 2016) :

L'internalisme du jugement moral	Nécessairement, si le sujet S juge que l'action X est moralement bonne, alors S est motivé à faire X et si S juge que l'action X est immorale, alors S est motivé à renoncer à faire X.

Quelques remarques clarificatrices sont nécessaires avant de continuer. En premier lieu, la relation de *nécessité* contenue dans la thèse internaliste est celle de la *nécessité conceptuelle*. Selon cette interprétation, la notion de « jugement moral » implique conceptuellement celle de « motivation ». Il est *contradictoire* de dire simultanément que « S juge que l'action X est moralement bonne » et que « S n'est pas motivé à faire X ». Si une personne attribue à S un jugement moral, sans lui attribuer la motivation correspondante,

cette personne ne comprend pas la signification des termes qu'elle utilise. Elle est conceptuellement confuse.

L'interprétation conceptuelle s'oppose notamment à une conception *causale* de la nécessité. Si la nécessité exprimée dans l'internalisme était causale, cela signifierait que les lois de la nature sont telles que, lorsque S porte un jugement moral, cela a pour effet que S est motivé à agir de manière correspondante. Autrement dit, selon la loi causale, tout jugement moral cause une motivation à agir correspondante. Dans ce cas, celui qui attribue un jugement moral à S sans lui attribuer la motivation correspondante n'est pas conceptuellement confus, mais il ignore une loi causale ; il n'a pas connaissance du fait (contingent) que tout jugement moral, dans le monde actuel, cause une motivation à agir. Nous retiendrons l'interprétation *conceptuelle* de l'internalisme du jugement moral, car il s'agit de l'interprétation retenue dans le débat métaéthique au sujet du non-cognitivism (JAQUET & COVA, 2019).

Il est également utile de préciser que l'internalisme du jugement moral peut être *fort* ou *faible*. Dans sa version *forte*, l'internalisme consiste à affirmer que si un individu émet un jugement moral, alors cet individu sera motivé à agir en fonction de son jugement, indépendamment de ses autres motivations. La motivation morale issue du jugement moral ne peut pas être anéantie par des motivations contraires. Dans sa version *faible*, l'internalisme consiste par contre à affirmer que, si un individu émet un jugement moral, alors cet individu sera motivé à agir par son jugement uniquement *en l'absence de motivations contraires plus fortes*. Ainsi, dans sa variante faible, l'internalisme du jugement affirme seulement qu'un jugement moral motive à agir *dans une certaine mesure (pro tanto)* – i.e. en l'absence de motivation contraire plus forte. A nouveau, c'est la variante *faible* qui est le plus souvent retenue dans le débat métaéthique (LEMAIRE, 2004, p. 938).

1.2.1. LES CONSÉQUENCES MÉTAÉTHIQUES DE L'EXPLICATION INTERNALISTE

La thèse de l'internalisme du jugement moral, selon laquelle il existe une relation de *nécessité conceptuelle* entre les jugements moraux et la motivation à agir, est employée en métaéthique dans le cadre d'un débat bien spécifique, à savoir le débat entre le cognitivism et le non-cognitivism.

Les partisans du cognitivisme soutiennent que les énoncés moraux sont des énoncés descriptifs qui possèdent une valeur de vérité et qui expriment des croyances, alors que les partisans du non-cognitivisme soutiennent que les énoncés moraux sont des énoncés non descriptifs qui sont dépourvus de valeur de vérité et expriment des états mentaux non cognitifs.

Dans ce débat, la thèse de l'internalisme du jugement moral est utilisée comme prémisse pour défendre le non-cognitivisme à l'aide du raisonnement suivant (SMITH, 2009; JAQUET & COVA, 2019) :

(1) Par nécessité conceptuelle, si le sujet S juge que l'action X est moralement bonne, alors S est motivé à faire X.	Internalisme du jugement moral
(2) Par nécessité conceptuelle, si le sujet S est motivé à faire X, alors S se trouve dans un état mental conatif.	Thèse classique (Humienne) de la motivation
(3) Par nécessité conceptuelle, si le sujet S juge que X est moralement bonne, alors S se trouve dans un état mental conatif.	Conclusion

La meilleure explication de la conclusion (3), i.e. du fait que, lorsque S émet un jugement moral, il se trouve dans un état mental conatif, consiste dans le fait que le jugement moral est un état conatif. Or, il s'agit précisément de la thèse non-cognitiviste. Si un jugement moral est un état mental conatif, comme le soutient un non-cognitiviste, cela explique assez trivialement pourquoi tout jugement moral est nécessairement accompagné d'un état mental conatif ; ils ne sont qu'un seul et même état mental.

La première prémisse de l'argument en faveur du non-cognitivisme est la thèse de l'internalisme du jugement moral que nous avons présentée précédemment. La seconde prémisse formule une thèse largement acceptée en philosophie au sujet de la nature de la motivation à agir ; cette thèse affirme que toute motivation à agir – morale ou non – provient nécessairement d'un état mental conatif, tel qu'un désir ou toute autre *pro-attitude*, et jamais uniquement d'un état mental cognitif, tel qu'une croyance. Selon cette

thèse, il existe donc au moins deux types d'états mentaux – les états mentaux conatifs et les états mentaux cognitifs – et seuls les premiers sont capables de motiver à agir.

Selon cette théorie classique de la motivation, une croyance (ou tout état mental cognitif) est un état mental qui a pour fonction de *représenter* le monde. De ce fait, la croyance est un état mental qui s'adapte à la manière dont le monde *est* ; la croyance est un état mental qui s'ajuste au monde. On parle alors d'un état mental ayant une direction d'ajustement esprit-monde⁴⁵ : si le monde est différent de la manière dont la croyance le représente, alors la « faute » appartient à la croyance et c'est le contenu de cette croyance qui doit être adapté.

La catégorie des états mentaux *conatifs* cette fois réunit l'ensemble des désirs ou *pro-attitudes*. Un état mental conatif est un état mental dont la fonction ne consiste pas à *représenter le monde*, mais plutôt à nous *disposer à modifier le monde* (JAQUET & NAAR, 2019, p. 26). La direction d'ajustement d'un tel état mental est inverse à celui de la croyance ; lorsque nous avons un désir ou une pro-attitude, c'est le monde qui doit s'ajuster à ce désir. Autrement dit, si le monde ne s'accorde pas à un désir, le désir ne disparaît pas. Au contraire, la présence d'un désir insatisfait motive à modifier le monde pour le satisfaire.

Du fait que la croyance a pour fonction de *représenter* le monde alors que le désir a pour fonction de nous *disposer à agir sur le monde*, les philosophes considèrent en général que la présence d'un désir est nécessaire pour qu'une personne soit motivée à agir. Une croyance, à elle seule, ne peut pas exercer de force motivationnelle, elle ne peut pas pousser quelqu'un à agir – puisque cela n'est pas sa fonction. Par exemple, il ne suffit pas que je crois que l'eau du lac est fraîche pour être motivée à m'y baigner. Pour être motivée à me baigner dans le lac, il est nécessaire que j'aie le désir de me rafraîchir – et je crois que la baignade dans de l'eau fraîche est un bon moyen de satisfaire mon désir.

Cette thèse psychologique, selon laquelle il est nécessaire d'avoir un désir pour être motivé à agir, s'appelle la conception « humienne » de la motivation à agir. Or, accepter la conception « humienne » de la motivation – prémisse (2) du raisonnement précédent – ainsi que l'internalisme du jugement moral – prémisse (1) – implique d'accepter que toute personne qui fait un jugement moral se trouve dans un état mental conatif – conclusion (3).

⁴⁵ Cette notion de « direction d'ajustement » a été introduite par la philosophe Elizabeth Anscombe (ANSCOMBE E., 1957).

Or, cette conclusion offre un argument de poids pour considérer que tout jugement moral est un état mental conatif. Accepter l'internalisme du jugement moral ainsi que la conception humienne de la motivation permet donc de favoriser la thèse métaéthique du non-cognitivism moral. Les jugements moraux ne représentent pas le monde, ne sont ni vrais ni faux, et ne sont que des préférences ou désirs individuels.

1.2.2. LES CONTRE-EXEMPLES

Afin d'éviter la conclusion non-cognitivist, deux options sont disponibles : soit nier l'internalisme du jugement moral, soit nier la thèse humienne de la motivation. De nombreux philosophes favorisent la seconde option ; ils rejettent la thèse humienne de la motivation (MCDOWELL, 1979; MCNAUGHTON, 1988; DANCY, 1993; NAGEL, 1970; SCANLON, 1998; SHAFER-LANDAU, 2005). Cela leur permet d'adopter une conception *cognitivist* des jugements moraux, tout en maintenant que, nécessairement, ces jugements motivent à agir. Ces philosophes soutiennent ainsi qu'au moins certains jugements cognitifs motivent par eux-mêmes à agir, c'est-à-dire en l'absence d'un état mental conatif indépendant.⁴⁶

Ces approches *cognitivist* de l'internalisme du jugement moral rencontrent cependant *des contre-exemples*. En effet, il existe des situations dans lesquelles des personnes émettent un jugement moral sincère, mais ne sont cependant pas motivées à agir, pas même *dans une certaine mesure*. Ces situations sont les suivantes : (1) les cas de déficiences pathologiques de la motivation, (2) les cas de faiblesse de la volonté et (3) des cas de « troubles de la perception morale ».

(1) Les cas de déficiences pathologiques de la motivation

En cas de déficiences ou maladies motivationnelles, telles que la dépression (STOCKER, 1979), l'apathie (DAMASIO, 1994) ou la psychopathie (CIMA, TONNAER, & HAUSER, 2010), les personnes n'éprouvent tout simplement *aucune* motivation à agir. Elles

⁴⁶ Chez Michael Smith, par exemple, la motivation morale s'explique de la manière suivante : le jugement moral (cognitif) de S concerne ce que S serait motivé à faire si S était pleinement rationnel. Or, par définition, une personne rationnelle est motivée à agir d'après son jugement au sujet de ce qu'elle serait motivée à faire si elle était pleinement rationnelle : « [U]n agent qui juge [...] qu'elle serait motivée [à agir d'une certaine manière] si elle était pleinement rationnelle [...] est pratiquement irrationnelle si elle n'est pas motivée à agir de cette manière » "[A]n agent who judges [...] that she would be motivated [to act in a certain way] if she were rational [...] is practically irrational if she is not motivated to act accordingly." (SMITH, 1994, p. 62)]. En conséquence, toute personne rationnelle est motivée (*pro tanto*) à agir d'après son jugement moral.

possèdent des capacités cognitives intactes, mais leur système émotif et motivationnel est endommagé. Dans ces situations, les personnes sont capables de former un jugement moral (cognitif), mais elles n'éprouvent pas de motivation – pas même *pro tanto*.

Pour les partisans du non-cognitivism, ces cas ne sont pas problématiques ; selon eux, les dépressifs, apathiques et psychopathes sont tout simplement incapables de former de véritables jugements moraux. A l'inverse, pour les internalistes cognitivistes, la situation est plus compliquée : dans leur perspective, il n'y a pas de raison de penser que les dépressifs, apathiques et psychopathes n'émettent pas de véritables jugements moraux, puisque leurs capacités cognitives fonctionnent parfaitement. Ils portent donc bien de véritables jugements moraux, mais ne sont pas motivés à agir.

(2) Les cas de faiblesse de la volonté

Les cas de faiblesse de la volonté représentent des situations courantes dans lesquelles une personne considère qu'il serait moralement approprié de faire une action, elle juge qu'il n'y a pas de considération plus forte en défaveur de l'action, mais elle ne parvient néanmoins pas à se motiver à agir conformément à son jugement moral. Elle juge que « toutes choses considérées », elle doit agir en fonction de son jugement moral, mais elle ne parvient pas à aligner sa motivation à son jugement. Une telle personne souffre de faiblesse de la volonté.

A nouveau, ce type de contre-exemples n'est pas problématique pour les partisans du non-cognitivism. En effet, pour le non-cognitivism, la faiblesse de la volonté n'existe tout simplement pas. Lorsque S affirme que X est une action moralement bonne, il exprime simplement une pro-attitude à l'égard de X. Si S agit finalement contre cette pro-attitude, cela signifie simplement qu'il avait un désir contraire plus fort d'agir autrement, mais cela ne signifie pas que S est irrationnel. Pour les internalistes cognitivistes, en revanche, il faut admettre qu'il existe des cas de faiblesse de la volonté, c'est-à-dire des cas où la motivation ne s'aligne pas à la croyance au sujet de la meilleure chose à faire. Dans ces situations, les personnes sont motivées à agir de manière irrationnelle – i.e. contre leur meilleur jugement. Or, à nouveau, cela est problématique, car les compétences cognitives des personnes faibles de volonté sont intactes, mais ces personnes ne sont néanmoins pas motivées à agir selon leur jugement.

(3) Les cas de « troubles perceptuels »

Pour certains réalistes cognitivistes, les situations dans lesquelles la motivation n'accompagne pas le jugement sont des situations de « troubles perceptuels ». Si une personne S juge que X est moralement bonne, mais qu'elle n'est pas motivée à faire X, cela signifie que S n'a pas « perçu » correctement les valeurs morales. David McNaughton soutient, par exemple que, lorsqu'une personne porte un jugement moral mais, néanmoins, n'est pas motivée à agir conformément à son jugement, la personne ne « voit pas » (contrairement à la personne vertueuse) que les considérations morales sont plus importantes que les autres considérations. Elle ne perçoit pas les considérations morales comme étant *prépondérantes* (MCNAUGHTON, 1988, p. 145).

Par exemple, Jean juge qu'il est immoral de porter atteinte au bien-être d'autrui. Toutefois, contrairement aux personnes dont le « sens moral » fonctionne bien, Jean ne voit pas que cela implique qu'il est inapproprié pour lui de faire du mal à autrui, même lorsque cela lui procure du plaisir. Jean ne « voit pas » que les considérations morales ont plus d'importance que toutes les autres considérations, notamment celles concernant le plaisir. Jean perçoit les valeurs morales, mais il ne perçoit pas correctement *la force* des considérations morales.

Pour répondre aux différents contre-exemples susmentionnés, des variantes plus complexes de l'internalisme cognitiviste du jugement moral ont été proposées. Selon ces variantes plus complexes, il n'est pas le cas que tout jugement moral constitue *dans toutes circonstances* une motivation à agir (*pro tanto*). Au contraire, un jugement moral ne constitue une motivation à agir (*pro tanto*) que dans *certaines circonstances*, lorsque, si l'on peut dire, « tout fonctionne bien » pour l'auteur du jugement.

Les conditions dans lesquelles le jugement moral constitue une motivation (*pro tanto*) à agir sont les suivantes (JAQUET & COVA, 2019, p. 6):

- A. L'auteur du jugement moral est psychologiquement normal (sans apathie, dépression, psychopathie)
- B. L'auteur du jugement moral est pratiquement rationnel (sans faiblesse de la volonté)
- C. Le sens moral de l'auteur du jugement moral fonctionne correctement (l'auteur ne souffre pas de troubles de « perception »)

Etant donné que cet internalisme complexe spécifie des *conditions* dans lesquelles le jugement moral implique la motivation à agir correspondante, il est qualifié d' « internalisme conditionnel du jugement moral ». Il s'oppose à l'internalisme simple selon lequel le jugement moral donne lieu *sous n'importe quelle* condition à une motivation à agir (*pro tanto*) correspondante.

L'internalisme conditionnel du jugement moral peut être formulé de la manière suivante :

L'internalisme conditionnel du jugement moral	Par nécessité conceptuelle ⁴⁷ , si le sujet S juge que l'action X est moralement bonne, alors S est motivé à faire X, et si S juge que l'action X est immorale, alors S est motivé à renoncer à faire X, <i>toutes choses égales par ailleurs</i> – i.e. en l'absence de troubles psychologiques, de faiblesse de la volonté et de « troubles de la perception morale ».

Pour résumer notre discussion jusqu'à ce point, nous avons vu, d'abord, que le défi pour toute doctrine métaéthique consiste à expliquer un double constat empirique concernant la relation entre les jugements moraux et la motivation à agir : en premier lieu, le constat qu'un jugement moral est la plupart du temps accompagné d'une motivation à agir et, en second lieu, le constat qu'un changement de jugement moral est de manière fiable et régulière accompagné d'un changement de motivation à agir. Nous avons vu ensuite qu'il existe une doctrine métaéthique qui parvient à expliquer de façon très directe cette connexion étroite entre les jugements moraux et la motivation ; il s'agit de la thèse de l'internalisme du jugement moral, selon laquelle il existe une *connexion nécessaire* entre un jugement moral et la motivation à agir correspondante. Cette thèse internaliste est en général accompagnée d'une thèse non cognitiviste. Lorsque cela n'est pas le cas, et que

⁴⁷ La notion de nécessité exprimée ici reste celle de la *nécessité conceptuelle*. L'idée est que, lorsque nous attribuons un jugement moral à une personne, nous voulons dire que cette personne *serait motivée à agir* si elle ne souffrait pas d'un problème motivationnel ou perceptuel. Lorsque nous disons « S juge que X est moralement bonne », nous affirmons que « S serait motivé à faire X si S se trouvait dans de bonnes conditions ». Le lien conceptuel se situe entre le jugement moral et la motivation *conditionnelle* plutôt qu'entre le jugement moral et la motivation *actuelle* (SMITH, 1994, p. 61). L'intuition sous-jacente à cette forme d'internalisme du jugement est qu'« il semble irrationnel pour un agent de valoriser quelque chose (ou de juger que quelque chose a de la valeur) et de ne pas être motivé au moins minimalement à agir en conformité avec ce qu'il valorise. » (ROCHELEAU-HOULE, 2013).

l'internalisme est combiné à une métaéthique cognitiviste, il devient nécessaire d'adopter une forme *conditionnelle* de l'internalisme, selon laquelle le jugement moral motive à agir dans des circonstances idéales. Le problème est que l'internalisme, même dans sa forme conditionnelle, fait face à une objection, à savoir l'objection externaliste.

1.2.3. L'OBJECTION EXTERNALISTE

L'internalisme du jugement moral combiné avec la théorie humienne de la motivation à agir mène au non-cognitivism moral. Autrement dit, s'il existe une relation nécessaire entre un jugement moral et la motivation à agir, et que la motivation à agir provient d'un état non cognitif, alors il existe une relation nécessaire entre un jugement moral et un état mental conatif, relation qui peut fort bien être celle de l'identité – comme l'affirment les non-cognitivistes.

Pour éviter la conclusion non cognitiviste, certains partisans du cognitivisme moral contestent la prémisse internaliste (1) du raisonnement précédent. Ils adoptent ainsi une position appelée « l'externalisme du jugement moral » (BRINK, 1989; FOOT, 1972, FRANKENA, 1958; RAILTON, 2016). Selon l'externalisme du jugement moral, la relation fiable et régulière observée entre le jugement moral et la motivation est purement contingente – i.e. elle dépend de la présence de désirs *indépendants* du jugement moral. La régularité et la fiabilité de la relation entre jugement moral et motivation s'expliquent alors par le fait que les désirs dont dépend la motivation à agir conformément à la morale sont largement répartis au sein de la population, mais elle ne s'explique pas par l'existence d'un lien nécessaire (conceptuel) entre le jugement moral et la motivation à agir. Par exemple, le désir d'avoir une image (morale) positive de soi joue un rôle primordial dans l'explication externaliste de la régularité et de la fiabilité avec laquelle les jugements moraux motivent à agir (LEMAIRE, 2008, p. 118).

Sur le plan formel, l'externalisme du jugement moral consiste, d'un côté, à nier l'internalisme moral et, de l'autre, à affirmer que la motivation pour accomplir l'action morale est fournie par un désir indépendant du jugement moral – tel que, par exemple, le désir d'avoir ou donner une image morale positive de soi. En affirmant la nécessité d'un désir indépendant du jugement moral, l'externalisme se distingue des théories qui, bien que niant l'internalisme, considèrent néanmoins que la motivation pour l'action morale ne

provient pas des désirs. Par exemple, un calviniste ne croit pas qu'il existe un lien nécessaire entre le jugement moral et la motivation. Il refuse donc l'internalisme. Toutefois, il ne croit pas non plus que la connexion entre jugement moral et motivation s'opère par l'intermédiaire d'un désir indépendant du jugement. Au contraire, selon le calviniste, la connexion entre le jugement et la motivation s'opère par l'intermédiaire de la grâce divine : Dieu assure que, chez les élus, le jugement moral s'accompagne d'une motivation.⁴⁸ L'externalisme, en affirmant la nécessité d'avoir un désir indépendant du jugement moral pour pouvoir être motivé à agir d'après son jugement, s'oppose donc aussi bien à l'internalisme qu'à la doctrine calviniste.

L'externalisme consiste ainsi, d'un côté, à nier l'internalisme et, de l'autre côté, à affirmer la nécessité de la présence d'un désir indépendant du jugement moral.

Externalisme du jugement moral	(a) Il est faux que: nécessairement, si S juge que X est une action moralement bonne, alors S est motivé à faire X et que si S juge que X est immorale, alors S est motivé à renoncer à faire X. (b) Pour que S soit motivé à faire X, S doit avoir un désir – indépendant du jugement moral – que l'action X permet de satisfaire.

S'il est faux que, nécessairement, un jugement moral s'accompagne d'une motivation à agir correspondante, alors il ne s'ensuit plus que tout jugement moral s'accompagne d'un état mental conatif. L'argument en faveur du non-cognitivisme, que nous avons présenté dans la section précédente, ne fonctionne plus.

Il ne suffit pourtant pas de simplement *affirmer* qu'une thèse est fausse pour contester un raisonnement. Il faut encore avancer des arguments qui prouvent que la thèse est fausse, ou, du moins, qui donnent des raisons de croire que la thèse est fausse.

Le principal argument des partisans de l'externalisme du jugement moral contre l'internalisme est le suivant : il est possible qu'il existe un individu amoraliste, c'est-à-dire un individu qui juge qu'une action est moralement bonne, mais qui n'est absolument pas

⁴⁸ Cet exemple représentatif nous a été suggéré par Laurent Jaffro.

motivé à faire l'action en question – bien qu'il soit pleinement rationnel, psychologiquement normal et que sa perception morale fonctionne bien (BRINK, 1997).

L'argument de la possibilité de l'amoraliste	Il est possible qu'il existe un individu qui juge que X est une action moralement bonne/obligatoire, mais qui n'est pas motivé à faire X – même si l'individu est psychologiquement normal, rationnel, et que sa « vision » morale fonctionne correctement.

Ainsi, nous pouvons parfaitement envisager qu'Andreas, un ami parfaitement rationnel, psychologiquement normal et dont le sens moral fonctionne bien, juge qu'il serait moralement approprié pour lui de faire un don à une œuvre d'entraide aux victimes de l'ouragan Chomsky, mais qu'Andreas ne soit *pas du tout* motivé à faire ce don. Il est tout à fait concevable qu'Andreas affirme « oui, il serait bon du point de vue moral d'aider les populations touchées par une catastrophe naturelle telle que l'ouragan Chomsky », mais qu'il ajoute « et alors ? Pourquoi devrais-je me sentir concerné ? ».⁴⁹

Or, s'il est possible qu'il existe un tel individu amoraliste, alors l'internalisme du jugement moral est faux. En effet, pour l'internalisme du jugement moral, *il est conceptuellement impossible* qu'un individu pleinement rationnel, psychologiquement normal et sans défaut de perception, juge qu'une action est moralement bonne mais ne soit pas motivé – pas même *pro tanto* – à l'accomplir.

1.3. UN CONFLIT ENLISÉ

Les partisans de l'internalisme du jugement moral ne sont néanmoins pas très impressionnés par les exemples d'individus amoralistes qui sont avancés contre leur position. En effet, les internalistes, en réponse, nient tout simplement que l'amoraliste use de concepts moraux pour exprimer de véritables *jugements moraux*. Selon eux, l'amoraliste fait un usage «entre guillemets» des concepts moraux ; il *parle* comme quelqu'un qui a des convictions morales, mais en fait il n'en a pas. Lorsqu'il *affirme* « il serait bon de faire un don à l'œuvre d'entraide pour les victimes de l'ouragan Chomsky » sans éprouver de motivation,

⁴⁹ Rappelons que l'internaliste peut tout à fait accepter l'existence d'un individu qui juge qu'il est bon de faire un don, mais qui n'est pas motivé *toutes choses considérées* à faire un don, d'un individu qui n'a pas *l'intention* de faire un bon. Ce que l'internaliste ne peut pas accepter, c'est la possibilité qu'une personne émette un jugement moral et n'éprouve *aucune* motivation à agir, pas même une motivation *pro tanto*.

il ne comprend pas véritablement ce qu'il dit ou il ne le dit pas sincèrement. En ce sens, contrairement à ce que prétendent les externalistes, un véritable amoraliste – i.e. quelqu'un qui *a la conviction* sincère qu'une action est bonne ou mauvaise, qui ne souffre pas d'une déficience psychologique, de faiblesse de la volonté ou de trouble perceptuel, mais qui n'a pas la motivation correspondante – est un individu conceptuellement impossible.⁵⁰

Les externalistes ne se contentent cependant pas de cette réponse. Ils rétorquent que l'amoraliste a de véritables *pensées* morales; l'amoraliste ne se contente pas de faire usage du *langage* moral sans avoir les pensées correspondantes, mais il a de véritables convictions morales qu'il exprime sincèrement.

Comme on le voit, le débat est enlisé entre deux définitions opposées, basées sur des intuitions opposées, au sujet de la nature des jugements moraux. D'un côté, les internalistes partagent l'intuition qu'un jugement moral sincère s'accompagne d'une motivation à agir. De l'autre côté, les externalistes partagent l'intuition que la figure de l'amoraliste ne représente pas une impossibilité conceptuelle; selon eux, le jugement moral ne s'accompagne pas nécessairement d'une motivation à agir.

Ce constat d'enlèvement du débat sur la motivation morale est largement partagé: c'est la conclusion notamment de l'article de la *Stanford Encyclopedia of Philosophy* sur le thème de la motivation morale, qui considère que le débat a atteint une « impasse ».

Les débats entre les internalistes et les externalistes se concentrent sur la figure de "l'amoraliste" – la personne qui en apparence émet des jugements moraux, mais qui n'est pas motivée à agir en conformité avec eux. De multiples efforts ont été produits pour répondre au problème de l'amoraliste, et ces efforts ont mené au développement d'un grand nombre de versions de l'internalisme motivationnel du jugement moral. Généralement, les internalistes ont insisté sur le fait que l'amoraliste est une impossibilité conceptuelle [...]. Les externalistes, bien entendu, maintiennent que l'amoraliste n'est pas une impossibilité conceptuelle. [...] A ce point de la dialectique, les internalistes et les externalistes

⁵⁰ Chez M. Smith, par exemple : « [L]e *mieux* que l'on puisse dire au sujet des amoralistes est qu'ils *essaient* de faire des jugements moraux, mais échouent. » ["[T]he *very best* we can say about amoralists is that they try to make moral judgments but fail" (SMITH, 1994, p. 68)].

produisent des arguments supplémentaires, dans le but de surmonter ce qui semble être désormais une impasse.⁵¹

Pour sortir de l'impasse, plusieurs philosophes proposent de recourir aux expériences empiriques (JAQUET & COVA, 2019). Les expériences doivent déterminer si les locuteurs compétents ont l'intuition que les jugements moraux impliquent ou n'impliquent pas nécessairement la motivation à agir. Autrement dit, on cherche à déterminer si les locuteurs compétents ont des intuitions internalistes ou externalistes. L'expérimentation consiste ainsi à dépeindre un individu qui émet un jugement moral quelconque et à demander au sujet de l'expérimentation s'il pense que l'individu éprouve ou n'éprouve pas nécessairement une motivation correspondante à agir.

Plus concrètement, une expérience empirique de ce type a été menée avec le cas suivant :

Anne regarde un reportage consacré à la famine au Soudan dans lequel on voit des gens qui meurent de faim et cherchent désespérément de quoi se nourrir. Anne ne souffre pas de troubles mentaux et fonctionne normalement : elle n'est ni dépressive, ni apathique, ni émotionnellement perturbée, ni psychopathe, ni quoi que ce soit de ce genre. Alors qu'elle regarde le reportage, Anne n'est pas motivée du tout, pas le moins du monde, à donner de l'argent à ceux qui sont en train de mourir de faim.

Question : Se peut-il qu'Anne pense qu'il est moralement requis de donner une part de son argent à ceux qui meurent de faim, quand bien même elle n'est pas le moins du monde motivée à le faire ? »⁵²

79% des participants ont répondu que le scénario était possible ; il est possible qu'Anne juge qu'il est moralement requis d'aider financièrement ceux qui meurent de faim

⁵¹ "Debates between internalists and externalists often center on the figure of the "amoralist"—the person who apparently makes moral judgments, while remaining wholly unmoved to comply with them. Various efforts have been made to respond to the problem of the amoralist, and these efforts have led to the development of numerous versions of motivational judgment internalism. Generally, internalists have insisted that the amoralist is a conceptual impossibility. [...] Externalists, of course, maintain that the amoralist is not a conceptual impossibility. [...] At this point in the dialectic, internalists and externalists tend to produce additional arguments in an effort to overcome what seems an impasse." (ROSATI, 2016).

⁵² Jaquet et Cova (2019) reprennent dans leur article les résultats des recherches menées par Strandberg et Björklund (2013).

sans pourtant être motivée à le faire. Cette étude tend à montrer que les locuteurs compétents non philosophes ont des intuitions externalistes plutôt qu'internalistes.

Selon nous, la démarche empirique n'est toutefois pas adaptée pour sortir de l'impasse actuelle du débat. Nous expliquons pourquoi dans la section suivante.

1.4. L'ORIGINE PROBABLE DE L'ENLISEMENT

Plutôt que de chercher une solution au débat par le biais d'expériences empiriques, nous préférons fournir une possible *explication* de la divergence d'opinion concernant la nature motivante ou non des jugements moraux, afin de montrer quelle est, selon nous, la marche à suivre adaptée pour s'extraire de l'impasse décrite préalablement.

David Brink a proposé une explication de l'enlèvement du débat sur la motivation morale qui nous paraît convaincante. Il suggère que l'opposition provient d'une différence dans la *définition* ou *caractérisation* de ce qu'est un jugement moral, par opposition à d'autres types de jugement. De ce point de vue, les internalistes ont tendance à caractériser les jugements moraux de façon purement *fonctionnelle* ; les jugements moraux sont des jugements auxquels les individus accordent une importance toute particulière – importance qu'il s'agit ensuite d'expliquer. Les jugements moraux, dans cette perspective, ne se distinguent pas des autres jugements par leur *contenu* spécifique, mais par leur *importance* ou par leur *rôle* dans la constitution de la personnalité individuelle. Est moral tout jugement auquel la personne accorde une importance particulière.

Cette importance accordée aux jugements moraux est ensuite expliquée, par exemple, par le fait que le jugement moral affecte la conception fondamentale qu'une personne se fait d'elle-même ou par le fait que le jugement porte sur les désirs *profonds* – plutôt que seulement *apparents* – de la personne (KORSGAARD, 1996). Or, si les jugements moraux concernent la personnalité ou les désirs profonds d'un individu, il semble effectivement qu'une personne ne peut pas émettre un jugement moral – affectant sa personnalité profonde ou ses désirs profonds – sans être motivée à agir par son jugement.

De l'autre côté, les externalistes proposent en général une caractérisation *substantielle* des jugements moraux. Les jugements moraux, selon eux, ne se distinguent pas des autres jugements par leur importance ou par leur rôle dans la vie d'un individu, mais plutôt par le fait que les jugements moraux possèdent un *objet* spécifique. Par exemple, certains

externalistes soutiennent qu'un jugement est moral s'il concerne la manière dont le bien-être des personnes est affecté. Tout jugement qui concerne la manière dont le bien-être des personnes est affecté est un jugement moral et tout jugement moral porte sur le bien-être des personnes. C'est le *contenu* du jugement moral qui permet de le caractériser comme *moral*, plutôt que sa *fonction* dans l'existence d'une personne ou son *importance* pour la personne.

Parfois, le débat entre les internalistes et les externalistes dépend du fait que les deux parties emploient des critères différents pour identifier les jugements comme étant des jugements moraux. Souvent, les internalistes emploient des critères fonctionnels – en gros, ces jugements sont des jugements moraux que l'évaluateur considère comme particulièrement importants – alors que les externalistes emploient des critères substantiels – en gros, ces jugements sont des jugements moraux qui concernent un certain sujet, par exemple, qui concernent le bien-être des parties affectées.⁵³

Comme le souligne Brink, si nous adoptons une caractérisation *fonctionnelle* du jugement moral, alors l'internalisme du jugement moral semble intuitivement plus plausible. Si lorsque S juge que « X est moralement bonne », il émet un jugement sur quelque chose auquel S accorde une importance toute particulière, alors il semble s'ensuivre immédiatement que S est affecté dans sa motivation par son jugement. Par contre, si lorsque S juge que X est moralement bonne, il émet un jugement sur la manière dont le bien-être d'autres personnes est affecté, alors il semble tout à fait envisageable que son jugement n'affecte aucunement sa motivation. Il est tout à fait envisageable que S soit totalement indifférent à la manière dont son action affecte le bien-être d'autrui.

Si juger que « X est moralement bonne », c'est juger que « X favorise le bien-être général », alors il est plus simple de considérer qu'un locuteur peut juger que X est moralement bonne, mais n'être pas motivé à agir. Si un jugement est *moral* en vertu du fait que l'évaluateur lui accorde une importance particulière, alors il semble en effet irrationnel de ne pas être motivé à agir selon ce jugement.

⁵³ "Sometimes the debate between internalists and externalists depends on the two employing different criteria for identifying judgments as moral judgements. Often, internalists employ functional criteria – roughly, those judgements are moral judgements that the appraiser treats as fundamentally important – whereas externalists employ a contentful criteria – roughly, those judgements are moral judgements that concern certain sorts of matters, for instance, having to do with the welfare of affected parties." (BRINK, 1997, p. 21).

Pour savoir si l'internalisme ou l'externalisme est vrai, il ne nous semble pas utile de mener des enquêtes empiriques ; en effet, à chaque fois que le jugement moral sera caractérisé de manière substantielle, les intuitions favoriseront certainement l'externalisme – comme dans l'exemple d'Anne qui est indifférente à la faim dans le monde. A chaque fois, par contre, que le jugement moral sera caractérisé de façon *fonctionnelle*, par exemple parce que S désirerait s'il ne commettait pas d'erreur au sujet de ses désirs les plus profonds – alors les intuitions favoriseront certainement l'internalisme du jugement moral. Il nous semble donc que ni les intuitions internalistes ni les intuitions externalistes ne peuvent servir de *prémisses* dans un débat métaéthique – pour ou contre le cognitivisme, par exemple. Les thèses internalistes ou externalistes sont plutôt les *conséquences* d'une position métaéthique préalable – voire, comme nous le soutiendrons, les conséquences d'une doctrine *normative* particulière. Si les jugements moraux concernent ce que l'individu désirerait dans certaines conditions idéales, ou concernent la conception fondamentale que la personne a d'elle-même, alors les jugements moraux seront nécessairement motivants. Si les jugements moraux concernent plutôt le bien-être, le plaisir, la souffrance ou le bonheur des autres personnes, ou les règles qui peuvent être acceptées par tous, alors il semble tout à fait possible que S porte un jugement moral qui n'affecte aucunement sa motivation.

Nous pensons donc qu'il n'est pas possible de prendre position pour ou contre l'internalisme du jugement moral avant de disposer d'une caractérisation de la nature et du contenu de la motivation morale. Or, étant donné que, selon nous, seule une caractérisation *substantielle* de la motivation morale est pertinente, comme nous le verrons dans le dernier chapitre, notre position aura pour corollaire de favoriser indirectement une forme d'externalisme du jugement moral.

1.5. L'OBJET DE LA CONTROVERSE COMPARÉ À L'OBJET DU PROBLÈME QUI NOUS INTÉRESSE

Nous avons vu, dans la section précédente, que l'opposition entre l'internalisme et l'externalisme du jugement moral ne peut pas être résolue par des arguments *indépendants* de la recherche métaéthique, c'est-à-dire par le biais d'intuitions au sujet de la possibilité ou de l'impossibilité de la figure de l'amoraliste. A cela nous ajouterons que le débat sur la motivation morale tel qu'il est formulé dans le débat métaéthique contemporain ne joue pas de rôle primordial dans la question que nous traitons pour deux raisons :

(1) En premier lieu, la défense de l'internalisme et de l'externalisme est utilisée dans le débat opposant le cognitivisme et le non-cognitivisme, qui n'est pas le débat qui nous intéresse dans notre travail.

(2) En second lieu, même s'il était possible de montrer que l'internalisme du jugement moral est vrai, cela ne nous aiderait que modérément. En effet, l'internalisme du jugement moral a pour conséquence uniquement que les personnes qui émettent un jugement moral ont une motivation *pro tanto* à agir d'après leur jugement, c'est-à-dire *en l'absence d'une motivation contraire plus forte*. De ce fait, nous pouvons conclure qu'il existe une motivation spécifiquement morale, mais nous ne pouvons rien conclure de plus. Notamment, nous ne pouvons pas conclure que, lors d'un conflit entre l'intérêt personnel et la morale, les individus sont capables de surmonter leur intérêt pour satisfaire les exigences morales. L'internalisme est tout à fait compatible avec le fait que *dans notre monde*, la motivation morale est tellement faible qu'elle disparaît complètement à chaque fois qu'un désir contraire (même très faible) est présent. L'internalisme du jugement moral ne permet ainsi aucunement de conclure, à lui seul, que la motivation morale pourrait remplacer l'intérêt personnel comme ressource motivationnelle fiable pour le fonctionnement de la société. Or, c'est précisément ce que nous cherchons à démontrer ; que la motivation morale est suffisamment répandue et puissante pour remplacer ou surmonter l'intérêt personnel lorsque les deux motivations s'opposent. Donc la thèse de l'internalisme du jugement moral, même si elle était vraie, ne serait pas suffisante pour notre thèse.

DEUXIÈME PARTIE

**Le sophisme moraliste :
L'argument des économistes**

L'ACCUSATION DE SOPHISME MORALISTE EN ÉCONOMIE

Dans le présent chapitre, nous étudierons l'accusation de « sophisme moraliste » soulevée par des économistes de différentes traditions. Nous répondrons à cette accusation en montrant que l'alternative préconisée par les économistes, consistant à utiliser l'intérêt personnel pour susciter chez les individus un comportement conforme à la morale, n'est pas plus « réaliste » que la solution « moraliste » consistant à s'appuyer directement sur les motivations morales des citoyens et acteurs économiques.

1. L'HYPOTHÈSE ÉGOÏSTE EN ÉCONOMIE

Comme Martin Hollis le formule de manière convaincante, toute théorie politique offre une recette pour préparer un « sirop socialisant » dont la fonction est de rendre l'être humain, doté d'une nature spécifique, capable de vivre en société.

Les recettes de la Bonne Société [...] ont produit de nombreux plats fameux en théorie politique [...] [L]a formule magique pour le sirop socialisant varie en fonction de l'analyse de la nature humaine. Par exemple, si les hommes sont essentiellement des égoïstes cupides poursuivant la richesse, la gloire et l'honneur, alors le sirop consistera en un mélange de répression par la peur et de récompense pour la coopération. Si les hommes sont nés libres, égaux et bons, ils ont alors seulement besoin d'être mijotés dans l'éducation éclairée au sein de l'institution démocratique. Si les hommes sont par nature les enfants pécheurs de Dieu, alors un chef conservateur va faire infuser son mélange à l'aide de notion telle que la loi, l'autorité, la tradition, la propriété et le patriotisme, le tout teinté d'une méfiance à l'égard de la raison.⁵⁴

⁵⁴ "Recipes for the Good Society [...] have produced many classic dishes in political theory. [...] [T]he magic formula for the socializing syrup varies with the analysis of human nature. For instance, if men are essentially greedy egoists in pursuit of riches, fame and honour, then the syrup will be a blend of repression through fear and reward for cooperation. If men are born free, equal and good, they need only to be stewed in Enlightened education amid democratic institution. If men are by nature the sinful children of God, then a

La théorie économique néo-classique repose elle aussi sur une certaine conception de la nature humaine : *L'Homme économique*. *L'Homme économique* possède deux caractéristiques distinctes : il est (1) rationnel et (2) égoïste (ou intéressé).

La *rationalité* en économie concerne exclusivement la manière dont les êtres humains *organisent* leurs préférences ou leurs désirs. Les conditions de rationalité les mieux acceptées sont la transitivité⁵⁵ et la complétude⁵⁶. *L'Homme économique* est donc, en tant qu'être rationnel, un individu dont les choix répondent à des préférences transitives et complètes.

La rationalité concerne la *forme* des désirs ou préférences. *L'égoïsme*, de son côté, concerne le *contenu* des préférences. Dans la théorie économique, l'égoïsme se définit comme le désir, pour un individu, d'améliorer son propre bien-être matériel, sans égard pour le bien-être matériel d'autrui. Cet égoïsme se caractérise par le fait que les préférences d'une personne ne varient que d'après la manière dont le bien-être matériel de cette personne elle-même est affecté. Ce souci exclusif de soi s'exprime de ce fait par une *indifférence* à la manière dont le bien-être matériel des autres personnes est affecté. Ainsi, le fait que le bien-être d'une personne A se trouve amélioré ou détérioré n'affecte pas la manière dont une autre personne B évalue les différentes options. Sur le plan formel, cette indifférence à autrui s'exprime par la condition de « séparation des fonctions d'utilités individuelles », dans le sens que l'échelle des préférences de chacun est indépendante de celle des autres. *L'Homme économique* est ainsi un individu qui est motivé à agir *exclusivement* par le désir d'améliorer son propre bien-être, sans se soucier du bien-être d'autrui.

L'intérêt personnel, ou «égoïsme» exclut aussi bien l'altruisme [i.e. désir d'améliorer la situation d'autrui] que la motivation morale [i.e. désir de faire le bien, ou de faire ce qui est juste ou d'être une personne vertueuse]. Les individus égoïstes ne sacrifient jamais une partie de leur intérêt pour promouvoir les intérêts de quelqu'un d'autre ou pour promouvoir

conservative chef will distil his brew from notions like law, authority, tradition, property and patriotism, tinged with distrust of reason." (HOLLIS, 1977, p. 1).

⁵⁵ Un ordre de préférences est transitif s'il respecte la clause suivante : si l'individu préfère l'option A à l'option B et l'option B à l'option C, alors il préfère A à C.

⁵⁶ Un ordre de préférence est complet s'il respecte la clause suivante : pour toute pair d'options A et B, soit l'individu préfère A à B, ou préfère B à A, ou est indifférent entre les deux. Aucune option n'est *incommensurable* à une autre.

une valeur supérieure, comme par exemple la vérité, la justice, la nation, l'épanouissement de la société ou la préservation de la nature. Ils ne se soucient du bien d'autrui ou des valeurs impersonnelles uniquement dans la mesure où les préserver ou les promouvoir permet de satisfaire leurs propres intérêts.

1.1. L'ÉGOÏSME MÉTHODOLOGIQUE

La représentation de l'être humain comme étant parfaitement *rationnel* a reçu de nombreuses critiques du fait que, dans la réalité, les êtres humains ne sont pas aussi rationnels que le prétend la théorie.⁵⁷ De notre côté, nous allons nous intéresser exclusivement à l'hypothèse de *l'égoïsme* sur laquelle repose une grande partie de la théorie économique. Nous allons commencer par poser la question suivante : étant donné que, depuis le début de la science économique, la majorité des économistes admet que l'hypothèse d'une nature humaine exclusivement égoïste est fausse, i.e. ne correspond pas à la manière dont sont motivés les êtres humains dans la réalité, pourquoi a-t-on quand même élaboré une science entière sur la base de cette hypothèse ? Pourquoi, en d'autres termes, représente-t-on les êtres humains comme étant motivés exclusivement par l'intérêt personnel, tout en sachant pertinemment que, dans la réalité, ils sont également altruistes et désintéressés ?

Il est possible de justifier ce choix d'une hypothèse *fausse* sur le plan descriptif de plusieurs manières. Nous en avons retenu quatre principales : (i) l'exigence de simplicité ou d'économie théorique, (ii) la formulation d'une première approximation, (iii) la possibilité de représenter mathématiquement le comportement et (iv) la nécessité méthodologique de « faire l'économie de la vertu ». La quatrième justification, fondée sur l'argument de « l'égoïsme méthodologique », sera l'objet de notre étude ultérieure.

Une première raison théorique (i) concerne la simplicité et l'esthétique d'une théorie qui parvient à expliquer le comportement humain sur la base d'une motivation unique. Si une théorie parvient à rendre compte (de manière non triviale) de l'ensemble des comportements humains par le désir de richesse matérielle, alors cette théorie doit être préférée à toute théorie qui aurait besoin, pour expliquer les mêmes phénomènes, de faire appel à un plus grand nombre de motivations. Cet argument a été vivement critiqué,

⁵⁷ On peut citer, par exemple, les critiques de Herbert Simon (1957), de Daniel Kahneman (2011) ou encore Jon Elster (2010).

notamment par Jon Elster dans son livre «Le Désintéressement». La simplicité d'une théorie n'est en effet une vertu que si la théorie est d'abord vraie, c'est-à-dire que si les hypothèses qu'elle contient sont vérifiées par l'observation. Or, l'hypothèse de l'intérêt personnel n'est pas vérifiée (non trivialement) dans les faits (ELSTER, 2009, p. 15).

Un deuxième argument théorique (ii) consiste à admettre que l'hypothèse de l'égoïsme généralisé ne correspond pas à la réalité, mais que néanmoins elle permet d'expliquer et de prévoir de manière approximative le comportement des individus.⁵⁸ Une objection qui a été faite à cet argument, notamment par Martin Hollis et Edward J. Neil (1975) est que les économistes, contrairement aux scientifiques dans les autres sciences naturelles, ne savent pas exactement ce que contiennent les clauses «toutes choses égales par ailleurs» qui accompagnent les « lois » de leur science. Ils ne savent donc pas, lorsque leur théorie n'est pas vérifiée dans la réalité, pourquoi elle ne l'est pas ; est-ce parce que l'hypothèse est fausse, i.e. les êtres humains ne s'intéressent pas exclusivement à leur propre situation matérielle, ou est-ce parce que les clauses « toutes choses égales par ailleurs » ne sont pas satisfaites ?⁵⁹

La discussion autour de cet argument est très intéressante, mais nous ne nous y arrêterons pas dans ce travail pour deux raisons : en premier lieu, nous pensons qu'il existe un meilleur argument en faveur de l'hypothèse de l'égoïsme, qui sera le centre de notre attention par la suite. En second lieu, même s'il est possible de considérer la science économique comme étant la science *d'un facteur causal unique*, i.e. le désir d'enrichissement, à l'exclusion de tous les autres facteurs causaux, vaguement définis, et de considérer que ses généralisations ne sont valables que dans la sphère particulière de la

⁵⁸ On parle alors de science *inexacte*. Le caractère inexact de la science économique a été interprété de différentes manières par les économistes. L'inexactitude peut concerner (1) l'existence d'une *marge d'erreur* dans les prévisions, (2) le caractère probabiliste des prévisions, (3) le caractère *contrefactuel* des lois de l'économie et (4) la présence de clauses *toutes choses égales par ailleurs* (*ceteris paribus*) (HAUSMAN, 1992, p. 128). Nous suivons Hausman pour considérer que la dernière interprétation est la plus convaincante.

⁵⁹ « On ne peut pas distinguer le fait que le modèle échoue à satisfaire les variables réelles observées du fait que « toutes les choses » ne sont pas « égales par ailleurs ». Le premier cas réfute le modèle, alors que le second cas ne le réfute pas ». [“We cannot distinguish the failure of a model to agree with the true values of observed variables from the failure of 'other things' to be 'equal'. The former refutes the model; the latter does not” (HOLLIS & NELL, 1975, p. 34)] Dans la même direction, Daniel Hausman, dans son livre « La science inexacte et séparée de l'économie » (HAUSMAN, 1992), défend l'idée que, de la difficulté à formuler des prédictions testables, la théorie économique devrait renoncer à la méthodologie positiviste ou poppérienne et justifier ses principes sur des bases intuitives. Il considère ainsi que la méthodologie employée *dans la pratique* par les économistes est adaptée à leur discipline, mais que les économistes se trompent en voulant défendre leur pratique à l'aide de critères positivistes ou poppériens.

production et de l'échange de biens matériels, où le désir d'enrichissement logiquement domine, cela ne suffit pas pour justifier les prétentions *d'étendre* l'analyse économique à d'autres activités – citoyenneté, recherche scientifique, assurances, gouvernement, etc. Autrement dit, même si l'on admet que l'hypothèse de l'égoïsme universel offre une première approximation acceptable du comportement dans le cadre de l'échange économique, il est difficile de considérer que cette hypothèse offre une bonne première approximation dans d'autres domaines de l'activité humaine. Pour justifier l'extension de l'hypothèse égoïste à de nouveaux domaines, il faut faire appel à un autre argument que celui de la première approximation. Nous pensons que le dernier argument que nous verrons, celui de la nécessité de « faire l'économie de la vertu », offre la justification recherchée.

Un troisième argument théorique (iii) avancé en faveur de l'hypothèse de l'égoïsme économique, formalisé par la séparation des ordres de préférences individuels, est le suivant : un avantage de la séparation des préférences individuelles provient du fait que cette séparation permet de représenter mathématiquement la motivation de chaque individu. Si les préférences de chaque individu dépendaient également des préférences des autres individus, cela impliquerait que les fonctions d'utilité dépendraient les unes des autres, ce qui rendrait leur représentation mathématique d'autant plus compliquée. Pour obtenir une représentation mathématique fonctionnelle du comportement humain, il vaut mieux considérer les fonctions d'utilité, représentant les préférences au sujet du bien-être matériel, comme étant indépendantes les unes des autres. Toutefois, à nouveau, comme dans le cas de l'argument de la simplicité, la possibilité de représenter mathématiquement un choix ou un comportement n'est un avantage que si la description est d'abord *vraie*, ou, au moins, s'il n'y a pas de raisons de penser que la description est fausse.

Le dernier argument (iv) est celui qui va nous intéresser dans la suite de cette présentation. Il ne s'agit plus, cette fois, d'un argument qui concerne la nature *descriptive* de la théorie économique, mais qui concerne la nature *prescriptive* ou *normative* de la théorie économique.

En effet, les économistes n'ont pas seulement pour fonction, dans la société, de décrire les phénomènes économiques. Ils sont également chargés de concevoir différentes institutions dans le but d'apporter des solutions à des défis collectifs ou d'améliorer la

situation collective de la société, notamment en termes de richesse économique. Les économistes sont, par exemple, chargés de construire le système fiscal, les assurances sociales, les institutions de prévoyance pour la retraite, les investissements dans les infrastructures communes, etc.⁶⁰

Dans cette démarche, ils doivent envisager des solutions *réalisables* à des défis collectifs. Or, cette exigence de « réalisabilité » des solutions proposées conduit les économistes à adopter la norme méthodologique suivante : pour être *réalisable* une solution devrait pouvoir fonctionner sans supposer que les individus sont particulièrement et spontanément altruistes, généreux, soucieux du bien commun, etc. C'est pourquoi les économistes réfléchissent à des solutions institutionnelles qui «font l'économie de la vertu», c'est-à-dire qui obtiennent des résultats socialement bénéfiques sans avoir besoin de faire appel à autre chose qu'à l'égoïsme des individus.

Quelles sont les raisons [...] qui favorisent l'hypothèse égoïste ? [...] [L]es économistes s'occupent non seulement d'expliquer les comportements mais également d'inventer les institutions comme la sécurité sociale ou le système fiscal. Dans cette dernière tâche, ils s'imposent souvent pour exigence, à la suite de David Hume, de *faire l'économie de la vertu*. Sans forcément supposer que les gens ne soient pas vertueux, altruistes ou désintéressés, ils refusent de faire reposer les institutions sur l'hypothèse qu'ils le sont (ELSTER, 2010, p. 16).

Il s'agit donc, pour les économistes, d'obtenir des résultats socialement bénéfiques tout en faisant l'hypothèse que les individus ne sont motivés à agir que par égoïsme. On appelle «mécanisation de la vertu» cette manière d'organiser les institutions sociales en vue de diriger ou guider les individus égoïstes pour qu'ils réalisent le bien commun – sans jamais directement le vouloir.⁶¹

⁶⁰ Dès son origine, l'économie politique a été conçue comme une discipline dont le but consiste non seulement à décrire les faits, mais aussi à permettre au législateur d'augmenter la richesse de la nation. Cette conception de l'économie politique est, par exemple, explicitement acceptée par Adam Smith: « Adam Smith, dans l'introduction au livre IV de la Richesse des nations, définit l'économie politique comme « une branche des connaissances du législateur et de l'homme d'Etat, [qui] se propose deux objets distincts », il s'agit à la fois de « procurer au peuple » ou « le mettre en état de se procurer lui-même » un revenu ou une subsistance abondante » et « de fournir à l'État ou à la communauté un revenu suffisant pour le service public. » La finalité de l'économie est ainsi « d'enrichir à la fois le peuple et le souverain », (vol. 2, p. 11). » (SHIRINE, 2008, p. 56).

⁶¹ Une objection qui émerge régulièrement contre notre propos est que, en fait, la théorie économique ne repose pas sur l'hypothèse égoïste, contrairement à ce que nous prétendons. Selon cette objection, nous confondons *maximisation de l'utilité* et égoïsme. Or, selon l'objection, qui a été soulevée également par

1.2. LE SOPHISME MORALISTE SELON LES ÉCONOMISTES

Selon les économistes, au lieu d'espérer ainsi naïvement que les hommes se soucieront de morale, il vaut mieux partir de l'hypothèse que les hommes se soucieront exclusivement de leurs propres intérêts. Il faut « faire l'économie de la vertu » dans l'organisation de la vie en société.⁶²

La conformité à la morale, même pour une morale que les agents considèrent comme « vraie », ne peut simplement pas être considérée comme acquise. 'Si les hommes étaient des anges', a tendance à argumenter l'économiste, les choses pourraient être différentes ; mais dans le monde réel il faut mettre de côté les conceptions héroïques de la nature humaine et s'occuper du comportement humain tel qu'il est, avec notamment tous ses défauts. Pour ce faire, l'économiste dispose d'un intérêt particulier pour les institutions qui, comme le disent souvent les économistes, « font l'économie de la vertu ». ⁶³

Hausman (2011, p. 22) contre Sen, la maximisation de l'utilité ne concerne *aucunement* le contenu des préférences et n'implique donc pas l'égoïsme.

Pour répondre à cette critique, faisons d'abord remarquer que nous ne considérons pas que l'égoïsme des individus soit représenté par la maximisation de l'utilité. En effet, nous sommes d'accord que la maximisation de l'utilité – condition purement formelle – n'implique pas l'égoïsme. En revanche, nous avons différentes raisons de penser que de nombreux modèles économiques admettent l'hypothèse de l'égoïsme *en plus* de celle de la maximisation. Ces raisons sont les suivantes : (1) la séparation des fonctions d'utilité, (2) l'attribution d'un contenu substantiel aux motivations des acteurs économiques (acheteurs, consommateurs, producteurs, etc.) (HAUSMAN & MCPHERSON, 2007, p. 236), (3) la manière dont les économistes décrivent leur propre discipline, (4) l'existence de modèles économiques « alternatifs » qui incluent des considérations tels que l'équité, l'altruisme ou même l'envie.

Notons finalement que, même si les économistes n'admettaient pas, en fait, dans leur travail quotidien, cette hypothèse, notre étude resterait importante. En effet, la thèse selon laquelle une doctrine politique « réaliste » se doit de « faire l'économie de la vertu » a été défendue non seulement pas des économistes, mais également par des philosophes. Cette thèse ne se cantonne ainsi pas à la sphère économique et mérite ainsi de toute façon notre attention.

⁶² Pour certains économistes, l'hypothèse égoïste est tout simplement fausse, mais elle est adoptée pour des fins méthodologiques. Pour d'autres, comme Geoffrey Brennan (1993, p. 133), l'hypothèse est vraie dans une large mesure, les êtres humains étant principalement égoïstes. Dans les deux cas, ces économistes favorisent, pour des raisons de *réalisme pratique*, les modèles de société qui font l'économie de la vertu. Notons encore que notre argument affecte uniquement les économistes qui adoptent la conception égoïste de la nature humaine, mais il n'affecte pas les économistes qui n'adoptent pas cette conception.

⁶³ "Compliance with morality, even one that the agents regard as 'true', simply cannot be taken for granted. 'If men were angels', the economist is inclined to argue, things might be different; but in the real world we must determinedly set aside heroic conceptions of human nature and deal with human behavior as it is, warts and all. To do so commits the economist to a particular interest in institutions, which, as economists often put it, 'economize' on virtue. [...]" (BRENNAN, 1993, p. 137).

Considérons par exemple le cas du comportement moral des entreprises.⁶⁴ La thèse des économistes partisans de l'hypothèse de l'égoïsme est la suivante : il ne faut pas supposer que les entreprises choisiront spontanément d'adapter leur comportement pour atteindre des objectifs moraux. Il est même naïf d'avoir une telle espérance. Au lieu de cela, un homme d'Etat – conseillé par l'économiste – doit se donner pour but de *faire converger* l'intérêt de l'entreprise et la réalisation des objectifs moraux. Pour ce faire, il doit *organiser* les incitations économiques et les sanctions légales afin qu'il devienne avantageux pour l'entreprise de bien agir. Il doit, en quelque sorte, rendre la morale *intéressante* ou *profitable* pour l'entreprise.

Selon les économistes, si un homme d'Etat ne procède pas de cette manière, i.e. s'il ne met pas en place des incitations et sanctions garantissant que le comportement moral soit intéressant pour une entreprise à la recherche du meilleur profit, alors notre homme d'Etat commet un *sophisme moraliste*. En effet, l'homme d'État croit que les entreprises *se comporteront* d'une certaine manière simplement parce que les entreprises *ont le devoir moral* d'agir de cette manière. Or, une telle inférence est fallacieuse. Dans les termes de l'économiste Ingo Pies :

Un sophisme moraliste apparaît lorsque l'on attend des entreprises qu'elles agissent différemment dans une situation où les incitations restent inchangées, sans se préoccuper de modifier les incitations afin que le changement de comportement devienne *nécessaire* (et même attrayant) pour l'entreprise. La différence d'approche dont il s'agit peut être exprimée formellement de la manière suivante: changement d'attitude versus changement des conditions.⁶⁵

Au lieu d'espérer naïvement transformer les attitudes et motivations des entreprises, il faut plutôt garantir le changement de comportement à l'aide d'incitations et de sanctions qui affectent l'intérêt personnel, i.e. en modifiant les *conditions* de choix plutôt que les *motivations*.

⁶⁴ Nous admettons qu'une entreprise est une entité morale à part entière, susceptible d'être titulaire de droits et de devoirs moraux en lieu et place des personnes physiques qui la composent.

⁶⁵ „Ein moralistischer Denkfehler liegt dort vor, wo man von Unternehmen erwartet, dass sie sich angesichts unveränderter Anreize anders verhalten sollen, anstatt dafür zu sorgen, dass eine veränderte Anreizsituation es für sie unabdingbar (und sogar attraktiv) macht, sich anders zu verhalten. Die Alternative, um die es hier geht, lässt sich auf die Formel bringen: Gesinnungswandel versus Bedingungswandel.“ (PIES, 2014, p. 36).

Nous l'avons vu, selon l'économiste Ingo Pies, un homme d'Etat qui a une conception trop angélique ou optimiste de la nature humaine commet un « sophisme moraliste ». Pour mieux comprendre de quoi il s'agit, nous suivrons l'économiste en établissant une distinction entre trois niveaux différents de la théorie politique : (1) le niveau de la description des motivations et comportements humains réels, (2) le niveau *de mise en application* des principes par les récompenses et sanctions économiques ou légales et (3) le niveau de la *formulation* des principes ou objectifs moraux idéaux.

Un *sophisme moraliste* consiste, selon Pies, à opérer une transition directe, dans le raisonnement et dans l'organisation de la société, du niveau le plus haut, celui des principes (3), au niveau le plus bas, celui des comportements individuels (1), en ignorant complètement le niveau intermédiaire de la mise en place d'incitations et de sanctions (2).

Le sophisme moraliste [...] consiste à poursuivre un idéal moral approprié, mais à vouloir le réaliser au mauvais niveau. Il consiste ainsi à vouloir directement sauter du troisième niveau au premier niveau, sans passer par le niveau intermédiaire, en adressant simplement aux acteurs économiques un appel au changement, acteurs dont on attend que – par seul souci moral de bien agir – ils adaptent leur comportement.⁶⁶

Simplement demander aux entreprises de se comporter moralement ne suffit pas. Par exemple, il ne suffit pas d'informer les entreprises du fait qu'elles ont le devoir moral d'offrir à leurs employés des conditions de travail sécurisées, de fournir à leurs clients toutes les informations pertinentes sur leurs produits, de contribuer fiscalement à faire fonctionner l'économie du pays dans lequel elles opèrent, pour que les entreprises soient effectivement motivées, sans incitations ni sanctions, à respecter ces exigences. Pour les motiver à agir d'une manière qui soit moralement appropriée, il faut s'assurer que cela sert leurs intérêts.

Ce raisonnement ne vaut pas uniquement pour les entreprises, mais pour l'ensemble des acteurs économiques – travailleurs, chômeurs, consommateurs, fournisseurs, etc. A l'égard de l'ensemble de ces acteurs, une attitude réaliste requiert de ne pas s'attendre à ce

⁶⁶ „Der moralistische Fehlschluss [...] besteht darin, das richtige Ideal auf der falschen Ebene einzufordern. Hier wird direkt – unter Auslassung von Ebene 2 – versucht, von Ebene 3 auf Ebene 1 zu springen, und zwar mittels der appellativ vorgetragenen Aufforderung, wirtschaftliche Akteure sollten – um eines moralischen Anliegens willen – ihr Verhalten ändern.“ (PIES, 2014, p. 35).

que les acteurs économiques adoptent un comportement moral lorsque cela est contraire à leurs intérêts économiques.

Pour éviter de commettre un sophisme moraliste [...] il est conseillé de partir du principe que les acteurs économiques ne vont pas modifier leur comportement spontanément, mais qu'ils vont le modifier seulement lorsque les incitations appropriées existent et qu'il est assuré, au niveau institutionnel, que les manières morales de se comporter ne seront pas pénalisées par la compétition économique (PIES, 2014, p. 36).⁶⁷

C'est donc pour éviter de supposer trop naïvement que les acteurs économiques se soucieront spontanément de morale que les économistes adoptent l'hypothèse selon laquelle l'ensemble de ces acteurs sont purement égoïstes. Cela leur permet d'envisager l'organisation des incitations et sanctions économiques d'une manière telle que le comportement moral devienne intéressant même pour des personnes motivées exclusivement par l'enrichissement personnel.⁶⁸

1.3. L'ÉCONOMIE COMME SCIENCE DE « MÉCANISATION DE LA VERTU »

Le projet d'organiser une société dans le but d'obtenir un comportement moral par le biais d'incitations et de sanctions peut être caractérisé comme un projet de « mécanisation de la vertu ». « Mécaniser la vertu » consiste dans ce sens à organiser la société de telle manière que même si tous les individus sont égoïstes, le résultat des actions individuelles est positif pour l'ensemble de la société. Autrement dit, « mécaniser la vertu » signifie obtenir l'acte (extérieurement) vertueux sans *motivations* vertueuses. L'idée est que si une société fonctionne – i.e. obtient un résultat socialement bénéfique – alors même que tous les

⁶⁷ Un autre exemple de cette position est exprimé par Raynauld dans le passage suivant: « Ce sont les règles du jeu qui déterminent la qualité des résultats et non la motivation des acteurs. Cette conclusion s'applique tout autant au secteur public qu'au régime de marché. L'intervention gouvernementale contribue au bien-être collectif seulement dans la mesure où les règles du jeu dans le secteur public subordonnent les intérêts personnels à l'intérêt public » (RAYNAULD & STRINGER, 2003, pp. 55-56).

⁶⁸ Des philosophes adhèrent également à la thèse selon laquelle il faut organiser la société en « faisant l'économie de la vertu ». Christopher Hamel (2015, p. 104) cite à ce titre Shelley Burt (BURTT, 1990). Il s'agit également d'une position répandue parmi certains instituts de philosophie économique. Par exemple, l'institut de philosophie économique de l'université de Mannheim se trouve sous l'influence des pensées de Karl Homann (2014) et ses compères (HOMANN, KOSLOWSKI, & LUETGE, 2007), fervents défenseurs de l'approche de la « mécanisation de la vertu », dont fait également partie l'économiste Ingo Pies. L'objectif de « faire l'économie de la vertu » est donc bien vivant aujourd'hui, autant parmi les économistes que parmi les philosophes.

individus poursuivent leurs intérêts personnels, sans se soucier directement du bien commun, alors cette société est plus solide et durable que si elle a besoin, pour fonctionner, que les individus poursuivent intentionnellement le bien commun.

[L]es partisans de la mécanisation de la vertu prétendent [...] que c'est en réorganisant les institutions de la manière à rendre la vertu *intéressante* que la société libre a des chances de se maintenir (HAMEL, 2015, p. 117).

Une société qui fonctionne en l'absence de tout souci du bien commun de la part des acteurs qui la composent est plus stable et durable qu'une société qui a besoin, pour fonctionner, que les acteurs accordent de l'importance au bien commun.

Etant donné qu'elle repose sur l'hypothèse de l'égoïsme universel, la science économique représenterait donc la science par excellence de la *mécanisation* de la vertu. Elle serait basée sur une conception *réaliste* ou non utopiste de la nature humaine. Le rôle des économistes consisterait en conséquence à élaborer des institutions sociales permettant d'obtenir des comportements vertueux sur la base de motivations non vertueuses.

Nous pourrions penser que seuls quelques économistes isolés justifient l'hypothèse de l'égoïsme dans leur science à l'aide de l'exigence méthodologique de « faire l'économie de la vertu ». Cela n'est cependant pas le cas. En effet, cette conception de la science économique, comme science de la « mécanisation de la vertu », se retrouve dans les écrits de nombreux économistes d'horizons différents. Elle se retrouve par exemple chez le représentant français du courant néo-classique contemporain Jean Tirole dans son ouvrage *L'Economie du bien commun* :

Ce livre part [...] du principe suivant: que nous soyons homme politique, chef d'entreprise, salarié, chômeur, travailleur indépendant, haut fonctionnaire, agriculteur, chercheur, quelle que soit notre place dans la société, nous réagissons tous aux incitations auxquelles nous sommes confrontés. Ces incitations – matérielles ou sociales – et nos préférences combinées définissent le comportement que nous adoptons, un comportement qui peut aller à l'encontre de l'intérêt collectif. C'est pourquoi la recherche du bien commun passe en grande partie par la construction d'institutions visant à concilier autant que faire se peut l'intérêt individuel et l'intérêt général (TIROLE, 2016, p. 15).

Plus loin dans le livre, il affirme encore :

L'économie insiste sur la nécessité de faire en sorte que les objectifs individuels soient alignés sur les objectifs collectifs (TIROLE, 2016, p. 62).

Nous retrouvons la même conception dans un texte de l'économiste allemand Horst Siebert, qui fut conseiller pour le gouvernement allemand et pour le président de l'Union Européenne entre les années 1990 et 2007.

C'est le rôle de l'économiste de montrer comment les arrangements institutionnels peuvent diriger les décisions individuelles – des ménages et des entreprises – de telle manière que leurs décisions au final participent à la réalisation des objectifs de la société dans son ensemble. Le problème consiste à mettre en place les incitations appropriées.⁶⁹

Cette même représentation de l'économie apparaît aussi chez des économistes travaillant sur des thématiques comme celles de l'aide au développement ou de l'éradication de la pauvreté. On citera en exemple l'économiste Partha Dasgupta :

La théorie standard fournit un cadre conceptuel pour expliquer et développer des mécanismes d'allocation de ressources dans un monde où le comportement moral ne peut pas être tenu pour acquis à moins que des incitations suffisantes soient mises en place. [...] Il n'y a absolument aucune légitimité dans l'affirmation selon laquelle la théorie standard fait l'hypothèse d'agents motivés exclusivement par la cupidité, ou plus particulièrement que la théorie prouve la supériorité économique d'une société dans laquelle la cupidité est encouragée.⁷⁰

Comme l'affirme clairement Partha Dasgupta dans cette citation, le but de l'hypothèse de l'égoïsme ne consiste pas à promouvoir ou favoriser une société d'égoïstes, mais plutôt à garantir que la société puisse continuer de fonctionner même si tout le monde était ou

⁶⁹ "[I]t is the economist's role to point out how institutional arrangements can steer the decisions of individuals – households and firms – so that their decisions eventually end up in contributing to the target of society as a whole. This is an issue of setting the right incentives." (SIEBERT, 2011, p. 164).

⁷⁰ "The standard theory provides a framework for explaining and further developing resource allocation mechanisms in a world where moral behavior cannot be taken for granted unless sufficient incentives are provided. [...] There is absolutely no substance to the claim that the standard theory hypothesizes greed on the part of moral agents, and that in particular it demonstrates the economic superiority of a society in which greed is encouraged." (DASGUPTA, 2011, p. 34).

devenait égoïste. En assurant un alignement des intérêts privés sur l'intérêt collectif, les économistes offrent l'assurance que le bien collectif sera réalisé, même si aucun acteur économique ne poursuit intentionnellement cet objectif.

Nous retrouvons également cette conception de l'économie comme science de la « mécanisation de la vertu » chez Peter Boettke, représentant contemporain de « l'école autrichienne », héritière de la pensée de Friedrich Hayek et de la théorie du choix public de James Buchanan.

L'idée centrale de cette approche de l'économie [i.e. l'approche « autrichienne »] repose sur deux observations concernant la société commerciale: (1) les individus poursuivent leurs propres intérêts, et (2) un ordre social complexe aligne les intérêts individuel avec l'intérêt général (BOETTKE, 2012 , p. xvii).

L'idée que la science économique permet d'organiser la société en faisant « l'économie de la vertu », c'est-à-dire sans attendre des êtres humains qu'ils soient disposés spontanément à sacrifier leurs intérêts personnels pour le bien commun, est donc une idée largement répandue parmi les économistes.

Bien entendu, une conclusion centrale de la recherche économique est que l'institution du marché – i.e. la propriété privée, l'offre et la demande, l'incitation par le profit et les pertes – offre la meilleure solution au problème de «mécanisation de la vertu».

[L]a spécification institutionnelle d'une économie de marché basée sur la propriété privée, guidée par les prix et disciplinée par le profit et les pertes, amènera tout comportement motivé par l'intérêt personnel à promouvoir la coopération sociale.⁷¹

Grâce au marché on obtient la prospérité générale, l'amélioration générale des conditions de vie, sur la base de motifs purement égoïstes. Le marché constituerait donc la

⁷¹ "The institutional specification of private property market economy guided by price signals and disciplined by profit and loss accounting will steer self-interested behavior in the direction of social cooperation." (BOETTKE, 2012 , p. 7). On retrouve la même conception du marché dans encore d'autres textes: « Le marché peut donner d'excellents résultats en dépit des motivations égoïstes des participants. Même si chacun cherche à vendre ses services à un prix élevé, à acheter au plus bas prix possible, à travailler le moins fort possible, le marché permet de maximiser le bien-être collectif. Cela tient aux règles de fonctionnement qui contraignent les acteurs à servir l'intérêt collectif s'ils veulent maximiser leur propre bien-être. Les règles du jeu du marché subordonnent l'intérêt collectif, si certaines conditions sont respectées. » (RAYNAULD & STRINGER, 2003, p. 56)

meilleure institution pour promouvoir le bien-être général à l'aide de la seule motivation de l'intérêt personnel.

2. LES CONSÉQUENCES SOCIALEMENT DANGEREUSES DE L'HYPOTHÈSE DE L'ÉGOÏSME

Dans cette section, nous allons exposer différents risques liés au fait de baser des décisions politiques et l'organisation de la société sur une conception purement égoïste de la nature humaine. Nous allons passer en revue quatre conséquences socialement dangereuses susceptibles d'être provoquées par l'adoption de l'hypothèse de l'égoïsme dans une théorie politique : (1) l'effet contreproductif des incitatifs économiques, (2) l'affaiblissement, voire la disparition, des compétences de discernement moral, (3) les coûts de la mise en place d'un appareil de surveillance et (4) le transfert de responsabilité morale des acteurs économiques vers l'Etat.⁷²

2.1. L'EFFET CONTREPRODUCTIF DES INCITATIFS ÉCONOMIQUES

Un premier danger lié à l'objectif de «mécaniser la vertu» par le recours à des incitatifs économiques provient du fait que ces incitatifs peuvent avoir un effet *contreproductif* sur le nombre total de bonnes ou mauvaises actions. En effet, des expériences empiriques ont montré que l'introduction d'incitations pécuniaires insuffisamment élevées fait parfois augmenter le nombre d'actes immoraux – plutôt que de le faire diminuer – et parfois diminuer le nombre d'actes vertueux – plutôt que de le faire augmenter.

Pour illustrer cet effet contreproductif des incitatifs économiques, nous allons distinguer deux types de situations : (1) le cas des *sanctions* économiques censées *décourager* les actes immoraux et (2) le cas des *récompenses* économiques censées *encourager* les actes vertueux. Nous verrons que le mécanisme en jeu et les effets socialement négatifs ne sont pas identiques dans les deux cas.

⁷² Un argument puissant contre l'hypothèse de l'égoïsme consisterait à montrer que, étant donné que les êtres humains sont en fait altruistes et soucieux de morale, un système politique qui *ne tient pas compte* des dispositions altruistes et morales des citoyens est inférieur, sur le plan pratique, à un système fondé sur l'altruisme et la motivation morale. Je remercie Christine Clavien pour avoir soulevé ce point.

2.1.1. L'ACHAT D'UN DROIT DE FAIRE DES ACTIONS IMMORALES

Le problème avec l'introduction d'une *sanction* économique pour décourager des actes immoraux provient du fait que la sanction risque de faire disparaître, dans la conscience des personnes, le caractère immoral de l'acte. Le risque que la valeur morale négative des actions disparaisse de la conscience des acteurs apparaît dès lors qu'une personne peut *acheter le droit* de commettre une action immorale. La sanction *monétaire* rattachée à une action immorale peut donner le sentiment qu'il est possible d'acheter un *droit* d'accomplir une action qui, en l'absence de la sanction, était considérée comme immorale. Autrement dit, lorsqu'une sanction pécuniaire lui est associée, l'action passe du statut d'action *moralement interdite* au statut d'action *moralement permise*. En outre, la perception *morale* de l'action laisse place à une perception *économique* de cette même action. La conséquence, empiriquement observée, est qu'en donnant la possibilité d'acheter un « droit de faire le mal », le nombre d'actions immorales augmente plutôt qu'il ne diminue.

Une expérience empirique, souvent mentionnée sur cette thématique, est celle d'une école israélienne qui a décidé d'introduire une sanction monétaire pour punir les parents qui venaient chercher leurs enfants en retard (GNEEZY & RUSTICHINI, 2000). Pour contraindre ces parents peu scrupuleux à l'égard des enseignants, la direction a exigé le paiement d'une pénalité économique en cas de retard. A sa grande surprise, cependant, la pénalité a eu l'effet contraire à celui attendu. Au lieu d'encourager la ponctualité, elle a déculpabilisé les retardataires, qui avaient désormais la possibilité de s'offrir le *droit* d'être en retard. Non seulement les parents qui arrivaient déjà régulièrement en retard, avant l'introduction de la pénalité, sont arrivés encore plus souvent en retard, mais certains parents qui, auparavant, n'arrivaient pas en retard, ont commencé à arriver en retard. La quantité et la régularité des retards ont donc augmenté.⁷³

La pénalité a ainsi supprimé chez les parents toute conscience du caractère irrespectueux de leur comportement à l'égard des enseignants, en leur permettant de se payer le « droit d'être en retard ».

⁷³ Une autre étude des mêmes auteurs montre que les incitations ont des effets contreproductifs sur les dons de charité et la performance d'étudiant-e-s (GNEEZY & RUSTICHINI, 2000). Dans certains cas, l'introduction d'une récompense monétaire, qu'elle soit faible ou élevée, a un effet contreproductif. Dans d'autres cas, une faible récompense a un effet contreproductif, alors qu'une récompense élevée améliore bien la performance par rapport à l'absence de récompense.

La sanction pécuniaire semble avoir supprimé chez les parents l'idée d'obligation morale de ne pas déranger les employés et elle les a conduits à considérer le fait d'être en retard comme simplement un bien (*commodity*) supplémentaire à acheter.⁷⁴

Nous pourrions penser que la situation présentée n'est pas aussi problématique qu'il n'y paraît. En effet, il semble qu'il suffit de supprimer à nouveau la pénalité pécuniaire pour que le taux de retards revienne à son niveau initial, c'est-à-dire au niveau antérieur à l'introduction de la pénalité. C'est d'ailleurs par ce raisonnement, en espérant retrouver le niveau initial de retards, que la direction de l'école a décidé de supprimer la pénalité. Malheureusement, au lieu de retrouver son niveau initial, la quantité de retards a encore augmenté ; après la suppression de la pénalité économique, le nombre de cas de retards devint encore plus élevé que lorsque la pénalité était appliquée.⁷⁵

Comment expliquer ce processus ? Selon nous, et c'est cela qui rend vraiment dangereux l'usage des pénalités économiques pour décourager les comportements immoraux, lorsqu'une sanction pécuniaire est introduite, la valeur morale d'une action est *remplacée* par sa valeur économique. La motivation morale est remplacée par une motivation économique. Le problème vient du fait que, lorsque l'on supprime ensuite la pénalité économique, l'action ne retrouve pas, dans la conscience des personnes, une valeur morale. Au contraire, une fois la pénalité économique supprimée, l'action ne possède plus ni valeur morale ni valeur économique. Plus rien ne semble donc pouvoir retenir les personnes de commettre les actions immorales : ces actions ne possèdent plus, aux yeux de personnes, une valeur morale négative et ne présentent plus non plus de coût économique.

⁷⁴ "The fine seems to have undermined the parents' sense of ethical obligation to avoid inconveniencing the teachers and led them to think of lateness as just another commodity they could purchase." (BOWLES, 2008, p. 1605).

⁷⁵ L'introduction des incitatifs économiques et des modèles de concurrence a eu des effets contreproductifs également dans la sphère de la scolarité, comme le montre une étude sur la réforme du système scolaire suédois. Dans ce cadre, l'application du modèle de l'entreprise au système de scolarité a eu la conséquence suivante : alors que les élèves ont reçus de meilleures notes de leurs enseignants, leur niveau s'est effondré en comparaison internationale. Cela s'explique par le fait que les enseignants étaient encouragés, par le contexte compétitif, à surévaluer leurs élèves (WENNSTRÖM, 2017, p. 26). De leur côté, les élèves ont été incités à considérer les notes ou évaluations comme un outil de sélection dans l'éducation supérieure, comme une monnaie dont le principal objectif est de susciter la comparaison et la concurrence avec autrui (WENNSTRÖM, 2017, p. 30). La valeur intrinsèque de la connaissance a disparu des esprits. La conclusion de l'étude est que « les incitations seules peuvent générer de puissants cadrages psychologiques qui réduisent la saillance des valeurs morales et offrent des justifications pour un comportement immoral. » ["[I]ncentives alone can provide powerful psychological frames that reduce the salience of moral values and offer justification for immoral behavior" (WENNSTRÖM, 2017, p. 13).]

Schématiquement, nous proposons de décrire le déroulement de la situation par le tableau suivant :

	Première étape Aucune pénalité	Deuxième étape Introduction d'une pénalité	Troisième étape Suppression de la pénalité
Valeur du retard	Valeur morale	Valeur économique	Aucune valeur
Nombre de retards	x	y où $y > x$	z où $z > y$

Dans la première situation, où aucune pénalité économique n'existe, la seule valeur du retard est un déficit moral – dont se soucient plus ou moins les parents. Dans la seconde situation, une pénalité monétaire est introduite, transformant la valeur morale négative de l'action immorale en coût simplement économique. Le nombre d'actions problématiques augmente. Dans la dernière situation, lorsque la pénalité économique est supprimée, l'action ne retrouve pas de valeur morale négative, mais elle n'a plus non plus de coût économique. L'action immorale étant devenue « gratuite » sous tous les aspects, le nombre de cas augmente encore. Ainsi, l'introduction d'une pénalité économique risque de manière irréversible de faire augmenter les cas d'actes immoraux.⁷⁶

2.1.2. LA DIMINUTION DES ACTIONS VERTUEUSES : LE CAS DU DON DE SANG

Nous avons vu, dans la section précédente, les effets possiblement contreproductifs des pénalités économiques pour décourager des comportements immoraux. Dans la présente section, nous considérons les effets possiblement contreproductifs des incitations

⁷⁶ De cette analyse, les économistes tirent, par exemple, la conclusion suivante: il ne faut pas considérer que la sanction purement pécuniaire puisse remplacer les mesures d'emprisonnement, du fait que les deux types de sanctions transmettent un message différent : l'emprisonnement explicitement signifie à celui qui a violé la norme que son acte est immoral, alors que la sanction économique donne l'impression que l'acteur participe à une forme de transaction économique (FEHR & FALK, 2002). De notre côté, nous souhaitons faire porter cette réflexion sur les mesures visant à protéger l'environnement : les précédentes considérations soulèvent des questions par exemple sur le caractère approprié ou non de l'attribution de «droits de polluer», puisque l'idée de pouvoir acheter un «droit de polluer» donne l'idée que, parce que nous avons payé ce droit, l'action de polluer cesse d'être moralement problématique. Cela risque donc de faire augmenter plutôt que diminuer les actes de pollution, notamment si le prix est insuffisamment élevé.

économiques visant à *encourager* des comportements moralement louables. Dans le cas qui nous intéresse désormais, il ne s'agit plus d'*éviter* une action mauvaise, mais d'*encourager* une bonne action. Dans ce cadre, le problème que pose l'introduction d'encouragements économiques, qui ont pour but de rendre l'action morale *intéressante*, est que parfois les personnes sont motivées à faire des actions précisément par le fait qu'il s'agit d'actions *désintéressées*. De ce fait, lorsque les actions désintéressées sont transformées en actions intéressées, les individus cessent par la même occasion d'être motivés à les réaliser.⁷⁷ Pour illustrer ce point, nous allons nous utiliser l'exemple classique des *dons de sang*.

Les expériences empiriques ont montré que rattacher une récompense économique à l'action de donner son sang ne fait pas augmenter, mais a tendance plutôt à faire diminuer, le nombre de cas de dons de sang (TIROLE & BENABOU, 2006, p. 1652). Le risque ici est que la présence d'une récompense donne l'impression que l'action de donner son sang est motivée par la cupidité plutôt que par le seul souci moral de faire le bien.

La présence d'incitations extrinsèques [i.e. non liées à la valeur morale de l'action] détruit la valeur des bonnes actions en termes de *réputation*, en créant des doutes quant à la mesure dans laquelle ces actions sont accomplies pour les incitatifs plutôt que pour elles-mêmes.⁷⁸

Or, du fait que de nombreuses personnes – selon l'étude, surtout les femmes – désirent avoir et donner une image morale positive d'elles-mêmes, cela a pour conséquence que leur motivation à donner le sang diminue lorsque leur bonne action est « achetée » par le biais d'une récompense monétaire. Ces personnes perdent une occasion de donner une image morale positive d'elles-mêmes ou craignent de donner une image morale négative d'elles-mêmes.⁷⁹ C'est pourquoi le nombre de bonnes actions diminue fortement.⁸⁰

⁷⁷ Nous retrouvons le désir de réputation morale étudié dans la première partie.

⁷⁸ "The presence of extrinsic incentives spoils the reputational value of good deeds, creating doubt about the extent to which they were performed for the incentives rather than for themselves." (TIROLE & BENABOU, 2006, p. 1654)

⁷⁹ Un effet similaire a été observé en Suisse sur le taux de participation aux votations. L'introduction du vote par correspondance a supprimé une occasion de se montrer publiquement en train d'exercer son devoir civique, ce qui a fait diminuer le taux de participation. Bien que le vote par correspondance permette de faire diminuer les coûts (en temps) du vote, il ne permet plus de savoir qui a accompli et qui n'a pas accompli son devoir civique (TIROLE & BENABOU, 2006, p. 1653). Cet effet a aussi été observé pour des situations de passagers clandestins : le fait que le resquilleur doive payer une amende diminue la force de motivation issue de la désapprobation sociale des autres personnes, en transformant l'acte immoral en une forme de transaction économique (URI, MEIER, & REY-BIEL, 2011, p. 201).

Il existe des situations dans lesquelles la récompense économique présente un effet *positif* sur la motivation à bien agir. C'est le cas, par exemple, de la motivation à s'investir dans ses études, à arrêter de fumer ou à perdre du poids. Toutefois, même dans ces cas, l'effet positif de l'encouragement économique n'est pas entièrement satisfaisant : il s'avère que son effet positif est limité au court terme. Sur le long terme, l'effet des récompenses économiques semble plutôt contreproductif : la personne est moins motivée à agir suite au retrait de l'incitatif économique qu'elle ne l'était avant l'introduction des incitatifs économiques.

Parce que l'effet standard de l'incitation n'existe plus sur le long terme (on définit le long terme par la situation après le retrait de l'incitation), l'effort sera inférieur à ce qu'il était avant que l'incitation ne soit introduite.⁸¹

Ainsi, même lorsque les incitatifs économiques fonctionnent sur le court terme, ils risquent d'avoir un effet contreproductif sur le long-terme.

Dans cette section, nous avons étudié les effets possiblement contreproductifs des incitatifs économiques dans deux types de situations : dans le premier type, ces incitatifs visent à *décourager* les actes immoraux. Nous avons vu que, dans ce cas, la présence d'un coût économique *supprime* la conscience du caractère immoral de l'acte avec un effet irréversible. Dans le second type de situation, la présence d'une récompense économique supprime une opportunité d'avoir ou de donner une image morale positive de soi, ce qui diminue le nombre d'actes vertueux. Nous avons vu également que, même lorsqu'une récompense économique peut avoir un effet positif sur le court-terme, rien ne garantit un effet positif sur le long-terme, c'est-à-dire lorsque l'encouragement est ensuite retiré.⁸²

⁸⁰ Sur la base de cette analyse, il est possible de tirer la recommandation suivante: les incitations économiques visant à récompenser les actes *vertueux* devraient plutôt être octroyées d'une manière *privée*, c'est-à-dire invisible au public. Par exemple, il vaut mieux encourager économiquement l'achat d'un réfrigérateur plus économe, ou d'un système de chauffage plus économe, que d'encourager économiquement l'achat d'une voiture électrique (URI, MEIER, & REY-BIEL, 2011, p. 202).

⁸¹ "Because the standard incentive effect is gone in the long run (we define the long run as after the incentives are removed), effort will be lower than it was before extrinsic incentives were offered." (URI, MEIER, & REY-BIEL, 2011, p. 194).

⁸² Les économistes sont tout à faits conscients de ce problème. C'est pourquoi ils tentent d'adapter leurs modèles incitatifs à chaque situation particulière. Toutefois, plus il est nécessaire d'étudier les cas particuliers pour mettre en place les structures incitatives adaptées, plus il devient difficile de voir l'avantage d'un tel instrument – en termes de besoins d'information – par rapport à une économie directement planifiée.

2.2. LA DISPARITION DES COMPÉTENCES DE DISCERNEMENT MORAL

Le second danger, lié au premier, provient du fait qu'en attendant des acteurs économiques qu'ils se contentent de suivre les incitations que le gouvernement a mis en place, ces acteurs cessent de réfléchir par eux-mêmes à la valeur morale de leur action. En effet, s'il est de notoriété publique que les incitations économiques ont pour fonction d'inciter aux comportements moralement appropriés, alors les acteurs économiques pourraient se contenter de suivre les incitatifs économiques, sans procéder par eux-mêmes à une réflexion sur la valeur morale de leur action.⁸³ Du fait qu'ils se contentent de suivre ainsi les incitatifs économiques, sans évaluer par eux-mêmes la valeur positive ou négative de leur action, les acteurs économiques risquent de perdre la capacité de discerner par eux-mêmes ce qu'il est approprié ou inapproprié de faire.

Pire encore, on peut arriver à une situation où les individus croient que c'est toujours en suivant les incitations économiques qu'ils accomplissent l'action vertueuse.

En affirmant l'exclusivité de ses intérêts personnels, le Sujet économique, dans la théorie libérale, n'a pas en effet conscience de nuire aux intérêts des autres sujets économiques. Il sait même que c'est en poursuivant son intérêt de producteur et de consommateur qu'il assure finalement la prospérité générale (FYOT, 1952, p. 18).

Il est vrai que, si les incitations économiques ont pour fonction d'indiquer le chemin à suivre pour réaliser le bien commun, alors il est même du *devoir moral* de chaque citoyen de suivre les incitations. Cette conception du rôle des incitations économiques, censées guider les intérêts particuliers vers l'intérêt général par l'institution du marché libre, a certainement contribué à forger l'idée, dans l'esprit des dirigeants d'entreprise, que leur enrichissement personnel non seulement n'est pas immoral, mais représente même une action moralement

Notons encore, comme nous l'a fait remarquer Christine Clavien, que notre argument présente une portée limitée : il montre que la mise en place d'incitatifs *peut* avoir des effets contre-productifs. Il ne permet pas d'affirmer, ni que cela soit *toujours* le cas, ni que, *globalement*, il y ait plus de désavantages que d'avantages à utiliser les incitations économiques. Seule une étude empirique globale – difficile à réaliser – permettrait d'évaluer l'effet cumulé positif ou négatif des incitations économiques. Nous pouvons néanmoins comprendre que, dans des domaines qui ne relèvent pas spécifiquement de l'économie, par exemple dans l'éducation, l'usage d'incitations économiques risque de transformer des motivations non économiques en motivations économiques, ce qui peut être dangereux.

⁸³ Ce problème a été soulevé par l'économiste néerlandais Burt Nooteboom (2014, p. 58).

louable. C'est du moins ce que soutient Joseph Stiglitz dans ce passage cité par Rudi Verburg :

La théorie de la main invisible [qui fournit la preuve que les intérêts privés concordant avec l'intérêt général] offrait un important soulagement aux dirigeants d'entreprise, puisqu'elle leur montrait qu'en poursuivant leur propre enrichissement, ils participaient à la prospérité générale. Ils devaient donc non seulement ne pas se sentir coupables, mais ils devaient au contraire se sentir fiers.⁸⁴

La théorie économique permet donc de soutenir l'idée que, en suivant aveuglément les incitations économiques mises en place par l'Etat, les acteurs économiques non seulement ne font pas de mal, mais ils agissent comme il est attendu d'eux qu'ils agissent.⁸⁵ Le fait que les acteurs économiques deviennent incapables de juger par eux-mêmes de la valeur morale de leur action et qu'ils soient même amenés à croire qu'ils accomplissent leur devoir moral en agissant pour leurs propres intérêts, présente plusieurs effets négatifs.

En premier lieu, il n'est pas garanti que les incitations économiques transmettent une information adéquate concernant la valeur morale de l'action. Au contraire, il peut se produire des situations dans lesquelles les incitations existantes poussent à adopter un comportement contraire à l'intérêt général. On peut penser, par exemple, au transport aérien souvent beaucoup moins onéreux que le transport ferroviaire, tout en étant beaucoup plus nuisible à l'environnement. De ce fait, si les acteurs économiques basent leurs décisions sur les incitations économiques, ils risquent d'être amenés à agir à l'encontre de l'intérêt général.

En second lieu, la théorie économique peut amener à considérer comme légitimes des actions immorales, du fait précisément que ces actions sont conformes aux intérêts des acteurs économiques. Autrement dit, le fait d'encourager les actions intéressées – censées être favorables au bien commun – peut amener à rendre acceptables des actions intéressées

⁸⁴ "The invisible hand theory was a great relief to CEOs, for it told them that by doing well (for themselves) they were doing good (for society). Not only should they feel no guilt they should feel pride." Citation tirée du livre *The Raaring Nineties* de Joseph Stiglitz (2003) (VERBURG, 2018, p. 2).

⁸⁵ « Il semble que les consommateurs et les entreprises ont en quelque sorte l'obligation de poursuivre leur intérêt personnel. Sinon les bénéfices du libre commerce ne seront pas réalisés ». ["It seems that consumers and firms are in some sense required to pursue their self-interest. Otherwise the economic benefits of free enterprise will not be realized." (MCMAHON, 1981, p. 252)].

même lorsque celles-ci ne contribuent pas au bien commun. Comme le remarquent Daniel Hausman et Michael McPherson :

Les économistes, en analysant les raisons intéressées de tricher et en explorant leurs implications, tout en ignorant et rejetant les raisons “morales”, peuvent finalement, de manière non intentionnelle, être amenés à justifier ces pratiques.⁸⁶

Du fait que la théorie économique justifie la poursuite de l'intérêt privé dans certains contextes, les acteurs risquent de considérer que la poursuite de l'intérêt privé est légitime dans n'importe quel contexte, même lorsque cela va à l'encontre de l'intérêt général.

Nous avons vu qu'il existe deux risques liés à cette manière d'interpréter les incitations économiques comme étant des « guides » pour faire ce qui est correct du point de vue moral ou collectif. Le premier risque provient du fait que les incitations peuvent indiquer une fausse route. Le second risque provient du fait que les acteurs économiques peuvent, du fait que la poursuite des intérêts privés est encouragée par la théorie, se sentir moralement justifiés à faire des actions qu'ils savent pourtant être moralement condamnables.

2.3. LE COÛT DE LA SURVEILLANCE ÉTATIQUE

Un problème supplémentaire provient des coûts, à la fois en termes économiques et en termes de liberté individuelle, du projet de « mécanisation de la vertu ». En effet, ce projet nécessite un système de sanctions et de récompenses capable de garantir que *tous les acteurs respectent en tout temps et en tout lieu les règles communes*. Un tel système nécessite donc un Etat omniscient et omnipotent, capable de sanctionner *toutes* les actions contraires au bien commun et de récompenser *toutes* les actions bénéfiques au bien commun. Or, une société dans laquelle la capacité de surveillance des comportements individuels par l'Etat est sans limite pose plusieurs difficultés – outre le fait qu'elle ne paraît pas désirable en termes de protection d'une sphère individuelle de liberté.⁸⁷

⁸⁶“Economists who analyze the self-interested reasons for cheating and who explore their implications while ignoring or dismissing “moral” reasons may, perhaps unintentionally, wind up justifying the practice” (HAUSMAN, 1992, p. 73).

⁸⁷ Bien entendu, nous pouvons envisager un État qui obtient le comportement moral par un conditionnement psychologique des individus, voire par la manipulation, en modifiant leurs croyances et préférences. Cependant, des solutions dans ce sens impliquent une atteinte à la liberté individuelle et à l'autonomie des acteurs que les économistes libéraux que nous étudions ne peuvent (en principe) pas accepter. Notons cependant qu'un économiste aussi libéral que Hayek semble accepter (à contrecœur) l'usage

En premier lieu, une surveillance totale requiert de disposer de matériel et de personnel permettant de garantir le contrôle. L'ensemble de ces investissements destinés à permettre d'assurer la surveillance des comportements économiques génère des coûts importants, assumés par les citoyens payant l'impôt. Il n'est donc pas assuré que la balance des coûts et bénéfices d'un tel système de surveillance soit positive pour la collectivité. En termes de richesse collective, il n'est donc pas certain qu'il soit avantageux de développer un appareil de surveillance et de sanction suffisamment vaste pour garantir la conformité de tout le monde en tout temps.

En deuxième lieu, un tel système total de surveillance, accompagné d'un système d'application des sanctions assurant que toute faute soit *effectivement* sanctionnée ne paraît pas réaliste. Dans le monde tel que nous le connaissons, il est impossible qu'un Etat soit capable de connaître et de punir tout écart par rapport aux normes régissant la vie économique.

Déjà au regard de normes légales, il est douteux que leur respect puissent être garanti par la seule contrainte extérieure. Sans un minimum d'honnêteté et de décence, un immense appareil de supervision et de dépistage étatique devient nécessaire. Pour forcer les citoyens à se conformer à la loi. Un tel appareil serait très coûteux [...] et très certainement très peu efficace.⁸⁸

de la manipulation et du conditionnement psychologique. En effet, selon lui, il faut encourager les gens à croire (faussement) que leur position sociale est le résultat de leurs propres efforts. Ce n'est que de cette manière que les individus adopteront l'attitude et les traits de caractères appropriés à l'économie de marché. « Etant donné la nature des marchés, telle qu'elle est décrite par Hayek, il existe un degré élevé d'ambiguïté morale dans le fait d'encourager les individus à adhérer à une croyance – que leur position et les ressources qu'ils contrôlent dans un marché sont ultimement le résultat de leur propre mérite et effort – qu'il admet comme étant « largement erronée » (HAYEK, 2006 (1960), p. 73). D'un côté, cette croyance doit être instillée par le biais d'une variété de mécanismes sociaux [...] parce qu'elle est « probablement, d'un point de vue pragmatique, l'incitatif le plus efficace vers l'action réussie (HAYEK, 2006 (1960), p. 73). De l'autre côté, sa « nature largement erronée » pourrait être contreproductive pour la légitimation du marché [...] » [“Given the nature of markets, as pointed out by Hayek, there is a high degree of moral ambiguity in favoring the adherence of individuals to a belief – that their position and the resources they control in a market society is ultimately the result of their own merit and effort – which he acknowledge to be “largely false” (HAYEK, 2006 (1960), p. 73). On the one hand, this belief should be instilled through a variety of social mechanisms [...] because it is “probably the pragmatically most effective incentive to successful action” (HAYEK, 2006 (1960), p. 73). On the other hand, its “largely false nature” could be counterproductive to the legitimacy of the market [...].” (RODRIGUES, 2013, p. 584).]

⁸⁸ „Schon im Hinblick auf Rechtsnormen muss bezweifelt werden, sie *allein* durch äusseren Zwang aufrechterhalten werden können. Ohne ein Minimum und Ehrlichkeit und Anstand wäre ein gigantischer Überwachungs- und Verfolgungsapparat erforderlich, um die Rechtstreue der Bürger zu erzwingen. Ein solcher Apparat wäre sehr kostspielig [...] und wahrscheinlich wenig effektiv.“ (BAYERTZ, 2004, p. 152)

Il est ainsi peu probable que l'Etat puisse garantir que toute action frauduleuse sera sanctionnée – sauf à développer un appareil étatique très envahissant et très coûteux économiquement.

On pourrait alors penser qu'un système de *récompenses* présente de meilleures chances de réussir qu'un système de sanctions. En effet, dans un système fonctionnant sur la *récompense*, ce sont les acteurs eux-mêmes qui doivent réclamer la récompense, en offrant spontanément et volontairement à l'Etat les preuves de leur mérite. La charge de la preuve repose ainsi sur l'acteur économique, qui peut choisir de rendre ses activités accessibles ou non à l'Etat. L'Etat n'a donc pas besoin de systématiquement surveiller les comportements, étant donné que ce sont les acteurs économiques eux-mêmes qui doivent fournir la preuve de leur bon comportement s'ils souhaitent recevoir la récompense. Les coûts économiques de la surveillance, ainsi que les coûts en termes de liberté individuelle, sont grandement réduits – en comparaison avec un système basé sur les sanctions.

Cependant, même un système de récompenses présente son lot de difficultés. En premier lieu, un tel système est également très coûteux économiquement pour la collectivité. En effet, pour que l'acteur économique – par exemple une entreprise – ait avantage à adopter le comportement récompensé, il faut que la récompense soit d'un montant supérieur à ce que l'acteur économique pourrait gagner par la fraude. Cela signifie que, plus une fraude est avantageuse, plus il faut offrir une récompense importante pour garantir que l'entreprise ne choisira pas l'action frauduleuse. Or, dans de nombreux cas, il y a des risques que la mise en place de la récompense ne soit pas avantageuse pour la collectivité : il faut récompenser l'action moralement conforme par une récompense supérieure à ce que coûterait le comportement frauduleux. Cela n'est donc pas avantageux.

En second lieu, d'un point de vue plus conceptuel, il est douteux que n'importe quel type d'action puisse être l'objet d'une *récompense*. En effet, les récompenses ne concernent que les actions *facultatives*, hautement *encouragées* mais non obligatoires. Les récompenses ne sont pas les conséquences légales appropriées pour les actions *obligatoires* ou *interdites*. Par exemple, dans le cas de l'interdiction de corruption, il ne fait pas sens de dire que l'État doit récompenser les entreprises qui ne pratiquent pas la corruption. Il doit plutôt *punir* ou sanctionner les entreprises qui pratiquent la corruption. De la même manière, une entreprise a l'*obligation* de fournir une information complète sur ses produits ; il ne paraît

ainsi pas approprié d'offrir des *récompenses* aux entreprises qui transmettent effectivement une information complète sur leurs produits. En conclusion, l'effet légal approprié à une *obligation* est la *sanction* en cas de non-conformité, plutôt que la *récompense* en cas de conformité.

Dans cette section, nous avons vu qu'un système de sanctions et récompenses capable d'*assurer* ou *garantir* qu'il est *toujours* avantageux pour *tous* les acteurs économiques d'agir en conformité avec l'intérêt général est un système non seulement irréaliste, mais en outre peu désirable, au vu des coûts économiques importants qu'il implique et au vu de l'intrusion totale de l'Etat dans le contrôle des activités économiques qu'il requiert.⁸⁹

2.4. LE TRANSFERT DE RESPONSABILITÉ MORALE

Selon la prochaine critique du projet de «mécanisation de la vertu», obtenir le comportement vertueux à l'aide d'incitations – i.e. taxes, sanctions, récompenses, etc. – engendre un déplacement de la responsabilité morale ; cette responsabilité ne revient plus à la personne qui a effectivement commis la faute, mais revient plutôt à l'État. En effet, sous l'hypothèse que les acteurs économiques poursuivent nécessairement leurs intérêts personnels, et qu'il est du rôle de l'Etat d'organiser les incitations pour diriger ces intérêts vers le bien commun, toute faute morale des acteurs économiques est considérée comme le

⁸⁹ Il est évident que les économistes ont conscience des effets incitatifs potentiellement contreproductifs de l'implémentation de certaines lois. Par exemple, une large littérature s'est développée pour évaluer le « taux optimal de taxation » (MIRREES, 1971), tenant compte du fait que les personnes susceptibles d'investir, c'est-à-dire les personnes fortunées, sont découragées par l'impôt à augmenter leur travail pour augmenter leur revenu, ce qui diminue la part globale de revenu lié à l'impôt pour la collectivité. Sur ce thème, voir par exemple le recueil d'articles réunis par Joel B. Slemrod (2000).

Selon nous, cependant, se concentrer sur la manière optimale de taxer la population, en termes de coûts et bénéfices pour la collectivité, transforme la discussion d'une manière qui, selon nous, est pernicieuse – comme nous y reviendrons dans la prochaine section. Par exemple, la question n'est plus de savoir quelle est la manière *correcte* ou *juste* de taxer les riches. La question devient plutôt : « De quelle manière ne pas susciter une incitation négative à l'égard des riches, qui les motiverait à diminuer leur travail et à augmenter leur loisir? ». La réponse à cette question semble être qu'il est avantageux de taxer les revenus moyens plutôt que les hauts revenus, car l'incitation à gagner davantage (marginale) y est plus forte. En effet, l'importance pour un individu riche de gagner une unité monétaire additionnelle est extrêmement faible – par rapport au gain en loisir, alors qu'elle est très importante pour des individus moins riches. La conclusion est donc la suivante : puisque l'augmentation de l'impôt n'a pas d'effet sur la motivation à travailler des personnes à revenus moyens, mais a un important effet sur les personnes à hauts revenus, alors il n'est pas souhaitable de taxer davantage les revenus élevés. L'impôt progressif est en conséquence inadapté.

Comme on le voit dans le raisonnement précédent, les considérations normatives sont largement laissées de côté. Or, nous rejoignons Avi-Yonah (2002) pour dire que la question la plus importante n'est pas celle de l'*effet* du taux d'impôt sur la motivation à contribuer, mais celle du fondement normatif de l'impôt – et de l'impôt progressif plus spécifiquement. L'impôt progressif est-il justifié uniquement pour sa contribution à la richesse collective?

résultat d'une mauvaise organisation des incitations. La faute revient donc à l'État – plutôt qu'à l'acteur économique. Daniel Hausman et Michael McPherson ont formulé une telle critique :

Dans la mesure où les économistes supposent que les seules raisons que nous devons chercher pour expliquer le comportement de la firme sont des raisons intéressées (*self-interested*) [...] leur analyse a tendance à excuser ou justifier le comportement de la firme et à attribuer la responsabilité pour la pollution non pas à la firme elle-même, mais au gouvernement pour avoir échoué à mettre en place les incitations nécessaires.⁹⁰

On peut constater les effets de ce transfert de responsabilité, par exemple, dans le discours ordinaire concernant les délocalisations d'entreprise. Dans ce contexte, il est commun d'entendre ou de lire que si les entreprises quittent une région ou ne s'y installent pas, cela est dû au manque « d'attractivité » de la fiscalité dans cette région – peu important les considérations de *justice*. Autrement dit, si une entreprise quitte un territoire dans le but de bénéficier d'une fiscalité plus avantageuse, créant de ce fait des difficultés économiques pour la région, la faute revient à l'État, qui a mal géré les incitations.

Par exemple, ce discours est très explicite dans un texte de la Fédération des Entreprises Romandes (FED) visant à défendre la réforme suisse de la fiscalité des entreprises (RIE III), soumise au vote populaire en février 2017.

Pour empêcher la délocalisation des entreprises [...] Genève devra offrir une fiscalité suffisamment attractive sur le plan international. Concrètement, elle devra combiner une baisse de l'impôt sur le bénéfice avec certaines mesures proposées par le projet de loi (KÄMPFEN, 2014).

Cette attribution de responsabilité est *indépendante* de la considération du caractère juste ou injuste du taux de fiscalité demandé par l'État et indépendante de la fonction que l'impôt est destiné à satisfaire au sein d'un État. En fait, dans cette approche, peu importe

⁹⁰ "To the degree that economists assume that the only reasons to be sought in explaining the firm's behavior are self-interested reasons [...] their analysis will tend to excuse or justify the firm's behavior and to locate responsibility for the pollution not with the firm but entirely with the government for failing to set the firm's incentives properly." (HAUSMANN, 2006, p. 73).

l'équité de l'impôt demandé, l'État commet une faute morale s'il ne propose pas une fiscalité suffisamment attractive et concurrentielle pour les entreprises.⁹¹

Le transfert de responsabilité des acteurs économiques vers l'État ne se remarque pas uniquement dans le discours ordinaire ou dans les débats politiques. Il est également présent précisément dans les écrits des économistes favorables à la « mécanisation de la vertu ». Par exemple, nous retrouvons dans le livre déjà amplement cité de Jean Tirole, *L'Économie du bien commun*, de nombreux passages attribuant la responsabilité à l'État pour des fautes morales commises par les acteurs économiques. Le premier exemple concerne les employeurs et les employés qui, selon Tirole, sont « poussés » par l'État à agir contre les intérêts de la communauté.

Les relations entre employeurs et employé sont médiocres en France... sauf quand il s'agit de s'entendre sur le dos de l'assurance chômage. Mais, comme toujours, les agents économiques réagissent aux incitations auxquelles ils sont confrontés. Les vrais coupables en la matière sont donc nos institutions, qui poussent à des manipulations concertées entre patronat et salariés au sein de l'entreprise (TIROLE, 2016, p. 332).

Comme il l'affirme explicitement dans ce passage, Jean Tirole considère que l'État est responsable du fait que les entreprises et les employés essaient de tirer avantage de l'assurance chômage à l'encontre de l'intérêt public. La faute morale est donc commise au

⁹¹ Globalement, la concurrence fiscale entre les États représente un cas de *dilemme du prisonnier* ; même si chaque État a avantage à baisser sa fiscalité, la concurrence fiscale est finalement désavantageuse pour tous les États. En effet, par la pression fiscale exercée par les entreprises, les pays ne sont plus capables d'assumer leurs prérogatives à l'intérieur de leurs frontières. Ils ne sont même plus capables d'assumer une forme de *souveraineté* à l'égard de la manière dont ils organisent la taxation. Par exemple, les études économiques montrent que, par la concurrence fiscale, les États deviennent de moins en moins capables d'assurer la provision de biens publics. Or, si la science économique peut être caractérisée comme la science de la compétition avantageuse – puisque la concurrence sur les *prix des marchandises* est avantageuse pour la collectivité – il est faux que *toute forme* de concurrence est avantageuse. Cela n'est pas le cas de la concurrence fiscale, qui est nuisible à la collectivité dans son ensemble. C'est pourquoi il convient de dissocier les situations de dilemme du prisonnier qui sont avantageuses pour la collectivité – comme c'est le cas de la concurrence sur les prix entre les entreprises – et les situations de dilemme du prisonnier qui sont désavantageuses pour la collectivité – comme c'est le cas pour la concurrence fiscale internationale. Il est ainsi possible de soutenir que les États ont des devoirs moraux les uns envers les autres – et non seulement envers leurs citoyens – étant donné que, par les mesures de baisses fiscales qu'ils décident de prendre à l'interne, ils affectent négativement la capacité des autres pays à exercer leur souveraineté. Ainsi, en vertu (1) d'un principe de souveraineté, (2) d'un principe de non-nuisance et (3) d'un principe d'assistance aux plus démunis, les États ont le devoir moral de mettre fin à la concurrence fiscale par des accords internationaux. (DIETSCH & CLAVEAU, 2008). Pour approfondir la réflexion sur la pression fiscale internationale, voir notamment (DIETSCH & RIXEN, 2014).

niveau de la mise en place des incitations économiques plutôt qu’au niveau de la prise de décision par les acteurs économiques.

D’autres passages du livre suggèrent la même conclusion. Un exemple particulièrement illustratif concerne la crise des subprimes de 2008 aux Etats-Unis. Si les banques ont émis des prêts extrêmement risqués, mettant en danger l’ensemble de la collectivité, cela relève de la responsabilité de l’État, qui a mal organisé les incitations légales et économiques.

Les émetteurs de prêts immobiliers ont toujours joué du désir des ménages d’accéder à la propriété pour leur vendre des contrats risqués. L’État fédéral américain aurait sans doute dû au minimum contribuer à rétablir la symétrie d’information entre émetteurs de prêts et emprunteurs [...] (TIROLE, 2016, p. 434).

La citation illustre bien notre propos : selon Tirole, les émetteurs de prêts ont toujours *joué – pour ne pas dire abusé* – du désir de propriété privée des citoyens. Cependant, ce constat n’amène pas l’économiste à attribuer la responsabilité morale de la crise aux émetteurs eux-mêmes – qui *abusent* les personnes. Au contraire, selon lui, la responsabilité morale revient à l’Etat qui n’a pas « rétabli la symétrie d’information » entre les émetteurs et les acheteurs.

Or, nous y reviendrons plus loin, une éthique du marché basée sur *la théorie des défaillances du marché* – théorie admise par Jean Tirole lui-même – a pour conséquence que les acteurs économiques ont le *devoir moral* de ne pas tirer profit d’une asymétrie d’information – même lorsque l’opportunité se présente et même lorsqu’elle n’est pas sanctionnée par la loi. Cela signifie, contrairement à ce que soutient Tirole, que la responsabilité morale peut être attribuée *intégralement* aux émetteurs de prêts, et cela en vertu d’une éthique *interne* au marché libre – et non pas en vertu d’un principe moral *externe* au marché.

En résumé, si nous adoptons la perspective d’un défenseur de la « mécanisation de la vertu », dans la lignée de Jean Tirole, il faut conclure que la fonction de l’État consiste à repérer et corriger les incitations économiques nocives pour le reste de l’économie :

L'enjeu [pour un État est] la chasse aux incitations encourageant les acteurs économiques à adopter des comportements nocifs pour le reste de l'économie (TIROLE, 2016, p. 448).

Si les acteurs économiques commettent des actes nocifs pour le reste de la collectivité, la responsabilité revient donc à l'État qui n'a pas adéquatement rempli sa fonction. La responsabilité morale est entièrement transférée, dans cette approche, de celui qui commet la faute morale à l'État qui n'organise pas correctement les incitations légales ou économiques.

3. LE MANQUE DE RÉALISME DE L'HYPOTHÈSE DE L'INTÉRÊT PERSONNEL

L'hypothèse selon laquelle les êtres humains sont incapables d'agir autrement que par intérêt personnel se justifie par l'exigence méthodologique de « faire l'économie de la vertu » dans la manière d'envisager, de construire et d'organiser les institutions politiques, sociales et économiques d'une société. Autrement dit, l'hypothèse que tous les hommes sont motivés exclusivement par l'intérêt personnel répond à une exigence de « réalisme » ou de « non-angélisme » dans la manière de représenter l'être humain dans les théories politiques.

Nous avons vu, dans le chapitre précédent, que l'adoption de cette hypothèse par les personnes chargées d'organiser la vie commune – notamment par les économistes-conseillers – risque potentiellement d'engendrer un ensemble de conséquences sociales fâcheuses. Toutefois, même si l'hypothèse ne s'accompagnait pas de tels risques pour la société, il resterait de bonnes raisons de mettre en doute sa pertinence. En effet, il est possible d'utiliser l'exigence méthodologique de « réalisme » à l'égard des motivations humaines *contre* l'hypothèse de l'égoïsme elle-même. En d'autres termes, le principal argument avancé en faveur de l'hypothèse de l'égoïsme peut être retourné contre cette même hypothèse. En effet, il est possible de soutenir que l'hypothèse de l'égoïsme, telle que nous la trouvons dans les théories économiques, est encore trop optimiste, irréaliste et naïve à l'égard des défaillances de la nature humaine. Cette objection de « manque de réalisme » prend quatre formes différentes, que nous nommerons selon les appellations

suivantes : (1) « Les motivations destructrices de la nature humaine », (2) « la morale des politiciens », (3) « l'égoïsme » et (4) « la norme morale fondamentale du marché ».

3.1. LES MOTIVATIONS HUMAINES DESTRUCTRICES

Dans cette section, nous allons étudier une première manière de retourner l'exigence de « réalisme » dans les théories politiques, exigence qui concerne la description des motivations humaines, contre l'hypothèse de l'égoïsme elle-même. Dans un premier temps, nous présenterons l'objection et, dans un second temps, nous étudierons une possibilité, pour les partisans de la « mécanisation de la vertu », de répondre.

3.1.1. L'EXCLUSION DE LA MALVEILLANCE HUMAINE

L'égoïsme de la théorie économique est défini comme une *indifférence* à l'égard du bien d'autrui, indifférence qui s'incarne mathématiquement dans l'indépendance des fonctions d'utilité des personnes. Le problème est alors le suivant : si une telle indifférence au bien d'autrui exclut certes toute forme de souci *positif* à l'égard du bien d'autrui, elle exclut également tout un ensemble de motivations humaines *négatives* à l'égard du bien d'autrui.⁹² Parmi ces motivations se trouvent notamment la jalousie, l'envie⁹³, la malveillance, le racisme, le nationalisme, la cruauté⁹⁴, le désir de domination, de violence, de

⁹² Les économistes sont tout à fait conscients que l'égoïsme, tel qu'il est défini en économie, i.e. comme simple indifférence à autrui, exclut aussi bien toute bienveillance que toute malveillance, comme le montre ce passage de Robert Sugden : « Je vais supposer la plupart du temps que les préférences des individus ne contiennent aucune motivation bienveillante et aucune motivation malveillante. » [“I shall usually assume that people's preferences are free of both benevolence and malevolence.” (SUGDEN, 1981, p. 33).

⁹³ L'envie se caractérise par le fait d'éprouver des sentiments négatifs à l'égard d'une autre personne pour la simple raison que la situation de cette personne est meilleure que la nôtre, sans que cette inégalité ne change quoi que ce soit à notre propre situation. L'envie, sous sa forme la plus sombre, à savoir « l'envie noire », peut pousser une personne à sacrifier ses propres intérêts dans l'unique but de porter atteinte à la personne enviée. Nous reprenons ainsi la définition proposée par John Rawls : « Nous pouvons [...] définir l'envie comme la tendance à éprouver de l'hostilité à la vue du plus grand bien des autres, même si leur condition plus favorisée que la nôtre n'ôte rien à nos propres avantages. Nous envions ceux dont la situation est supérieure à la nôtre [...] et nous voulons les priver de leurs plus grands avantages même s'il est nécessaire pour cela que nous renoncions nous-mêmes à quelque chose. » (RAWLS, 2009, p. 575). Il faut noter que toute forme d'hostilité à l'égard de la situation favorisée d'autrui n'est cependant pas de l'envie ; si cette hostilité est motivée par un jugement d'injustice, alors on parlera de *ressentiment* plutôt que d'*envie*. Le ressentiment, contrairement à l'envie, est un sentiment moral (RAWLS, 2009, p. 576).

⁹⁴ Pour une définition de la cruauté, qui la distingue à la fois de la brutalité et de la malveillance, nous nous appuyons sur la définition suivante : « La cruauté, le désir de faire souffrir [est] une disposition, qui, comme le fait remarquer Nietzsche, diffère de la brutalité dans la mesure où elle nécessite la même sensibilité que la bonté, alors que la brutalité est insensibilité. Dans les formes les plus courantes de la cruauté, l'agent tire son plaisir de l'impression qu'il a de provoquer lui-même douleur et frustration chez autrui, et il est clair qu'à travers ce comportement cruel, il tente d'exprimer sa puissance. La malveillance est à la fois très proche et

pouvoir, peut-être même un désir de « contre-justice »⁹⁵, etc. En définissant l'égoïsme par simplement de l'indifférence à autrui, la théorie économique ignore complètement l'existence de motivations *négatives* et *antisociales* chez l'être humain et fait ainsi reposer le fonctionnement de la société sur une conception encore trop généreuse de la nature humaine.

En conséquence, même s'il est vrai qu'une conception de la société reposant sur l'hypothèse de l'égoïsme apparaît comme moins « utopiste » qu'une conception qui dépendrait du fait que les individus se soucient de morale ou du bien commun, elle reste toujours trop « utopiste » du fait qu'elle ignore complètement les motivations les plus dangereuses de la nature humaine. Comme le fait remarquer Christopher Hamel, les partisans de la « mécanisation de la vertu » parviennent à offrir une conception de la société qui « fait l'économie de la vertu » uniquement parce qu'ils négligent un ensemble de motivations humaines potentiellement destructrices pour la vie sociale.

L'idée selon laquelle il y aurait une sorte d'harmonie préétablie entre les intérêts personnels les plus fondamentaux et le bien public [...] trahit un singulier optimisme, d'autant plus étrange qu'il vient d'une théoricienne [i.e. Shelley Burtt] revendiquant explicitement le « réalisme » contre le « mythe historique » de la vertu publique. Qu'est-ce qui permet à Burtt de supposer que les intérêts personnels des individus peuplant les sociétés démocratiques contemporaines se limiteront à la défense vigoureuse de leurs libertés contre les atteintes du pouvoir, ainsi qu'à l'enrichissement personnel seulement en tant qu'il contribue à la prospérité générale ? Dans la mesure où elle se garde bien de hasarder l'hypothèse selon laquelle les désirs les plus craints par les républicains du passé – l'ambition, l'avarice, la domination, la gloire personnelle – auraient disparu des sociétés contemporaines, on se demande par quel miracle leur pleine satisfaction pourrait avoir des effets favorables au bien commun (HAMEL, 2015, p. 116).

très différente de la cruauté, et on pourrait voir en elle une motivation ressemblant à l'envie, dans laquelle on désire simplement que le bonheur d'autrui n'existe pas. » (WILLIAMS, 2014, p. 2023).

⁹⁵ « La contre-justice [est une] disposition [consistant] à frustrer les fins de la justice, [parfois simplement] pour prendre du plaisir devant la frustration de la justice en tant que telle et la déception que ressentent les personnes de bonne volonté. À la limite, cette contre-justice peut constituer une esthétique presque désintéressée de l'horrible [...]. » (WILLIAMS, 2014, p. 2023).

Dès lors que nous admettons l'existence de motivations destructrices dans notre description de la nature humaine, afin de satisfaire pleinement à l'exigence de « réalisme », alors il devient beaucoup plus difficile de comprendre comment il est possible que l'absence totale de vertu puisse amener à un résultat collectivement avantageux.

3.1.2. L'INTÉRÊT PERSONNEL COMME FREIN AUX PASSIONS DESTRUCTRICES

Selon Samuel Scheffler, l'absence presque complète, dans les théories morales et politiques, des motivations peu louables de la nature humaine provient du fait que la philosophie morale, peut-être sous la pression d'une exigence de « scientificité », tend de plus en plus à vouloir offrir une conception simplifiée, plus économe, moins qualitative et plus quantitative, plus superficielle et moins profonde, de la nature humaine.

Comme nous le savons tous, les êtres humains sont motivés par des choses telles que la jalousie, la colère aveugle, et la haine personnelle; par le désir de vengeance, de perte, et de mort ; par des sentiments d'inutilité et de grandeur ; par le besoin de contrôle et par la peur de l'impuissance ; par le besoin d'amour et par la crainte du besoin ; par l'envie de détruire et le désir de réparer. Cependant, la philosophie morale s'est bien trop souvent contentée de travailler avec un répertoire très limité et hautement stylisé [de motivations], qui inclut les sentiments d'amour de soi et de sympathie, un possible sens du devoir, et peut-être encore quelques loyautés et attachements personnels.⁹⁶

De par l'étroitesse de son répertoire des motivations humaines, la philosophie morale et politique – dont est issue, historiquement, la théorie économique – ignore tout un ensemble de motivations humaines pourtant potentiellement nuisibles pour la société.

Il est sans doute vrai que la plupart du temps, la philosophie morale et politique, ainsi que la science économique, reposent sur un répertoire extrêmement appauvri des motivations humaines. Il ne s'agit cependant pas, selon nous, de la principale raison pour laquelle l'égoïsme est venu occuper une place aussi centrale dans la théorie économique. Il

⁹⁶ "As we all know, human beings are motivated by things like sexual jealousy, blind rage, and personal loathing; by desires for revenge, recognition, of loss, and of death; by feelings of worthlessness and grandiosity; by the need for control and the fear of powerlessness; by the need for love and the fear of needing; by the urge to destroy and the desire to repair. However, moral philosophy has too often been content to work with a very limited and highly stylized repertoire that typically includes feelings of self-love and sympathy, possibly a sense of duty, and beyond that perhaps a few personal loyalties and attachments." (SCHEFFLER, 1992, p. 8).

existe en réalité un puissant argument pour justifier l'accent porté sur l'intérêt personnel – considéré comme désir d'enrichissement matériel – par opposition aux motivations socialement encore plus dangereuses de la nature humaine. Cet argument a été avancé notoirement par les économistes et philosophes du 18^{ème} siècle, notamment Adam Smith et David Hume.

L'argument se présente sous la forme suivante: il est vrai que la conception de l'égoïsme dans la théorie économique exclut aussi bien les motivations immorales et antisociales des êtres humains que leurs motivations morales et pro-sociales. Cependant, c'est également cela qui fait sa force : en tant que motif puissant de l'être humain, le désir d'enrichissement matériel est susceptible, mieux que les motivations morales ou altruistes, de servir de rempart aux motivations antisociales. Ainsi, s'il est vrai que le désir d'enrichissement ne représente pas la motivation la plus louable de la nature humaine, il représente la motivation la plus efficace pour contrer les aspirations les plus mauvaises de l'être humain. C'est pourquoi Adam Smith et David Hume, adoptant une posture « pragmatiste » plutôt que « moraliste », considèrent qu'une société commerçante est une bonne solution pour résoudre les problèmes de sociabilité de l'homme. Dans les termes de Dennis Rasmussen :

Un commerce étendu pourrait bien être incompatible avec les vertus républicaines strictes, [Smith et Hume] le reconnaissent, mais ils estimaient également que la valorisation de l'intérêt personnel matériel aiderait à remplacer les passions dangereuses et sources de division telles que la xénophobie, l'intolérance religieuse ou la soif de gloire militaire.⁹⁷

Autrement dit, contre les motivations destructrices de l'homme, violentes et imprévisibles, il faut opposer une motivation tout aussi puissante, mais *calme* et *prévisible*. Le désir d'enrichissement, basé sur le calcul, exige calme et réflexion. En outre, en plus d'être calme et prévisible, le désir d'enrichissement constitue une motivation *puissante*, capable de réfréner les passions les plus dangereuses de la nature humaine. Ainsi, ce désir présente des qualités qui le rendent très intéressant pour être utilisé comme ressource

⁹⁷ "Extensive commerce might be incompatible with strict republican virtue, [Hume and Smith] acknowledged, but they also believed that a focus on material self-interest would help to replace dangerous and divisive passions such as xenophobia, religious intolerance, and the thirst for military glory." (RASMUSSEN, 2014, pp. 263-264).

motivationnelle fondamentale de la vie en société ; il est à la fois puissant, calme et prévisible.

Etant donné les attributs positifs de l'intérêt personnel ou de l'amour de l'argent, l'avarice a commencé à être saluée comme une passion calme et durable, capable de réfréner les passions sauvages. Cette transition idéologique et sémantique préfigurait l'idée que l'intérêt est le constituant fondamental de l'ordre et de la fortune.⁹⁸

En résumé, l'intérêt personnel a été considéré, dans l'histoire de la pensée économique, comme le meilleur rempart possible aux motivations humaines les plus destructrices. En envisageant une société fonctionnant exclusivement par intérêt personnel, les économistes n'ont ainsi pas ignoré le fait que les êtres humains possèdent des motivations malveillantes violentes. Au contraire, ils ont voulu envisager une société où le désir de richesse sert de rempart aux pires motivations humaines – rempart que les motivations plus « nobles », mais aussi plus « faibles », n'étaient pas capables d'offrir.

En outre, au-delà du rôle quelque peu « négatif » de frein aux passions violentes, la passion de l'intérêt personnel offrait également certains avantages positifs. En premier lieu, la pratique du commerce non seulement requiert de mettre un frein aux pires passions humaines, mais il favorise aussi un ensemble de dispositions désirables de caractère.

[Hume et Smith] ont soutenu que la société commerçante favorise le développement des vertus « bourgeoises » de loyauté, de bienséance et de coopération.⁹⁹

En second lieu, par le commerce – et par l'argent – les individus se libèrent de relations de dépendance à l'égard d'individus particuliers. Si certains considèrent comme regrettable le caractère « impersonnel » des relations commerciales, cette impersonnalité peut aussi

⁹⁸ "Given the positive attributes of material interest or love of money, avarice came to be acclaimed as a calm and persistent passion capable of taming the wild passions. This ideological and semantic drift heralded interest as the building block of order and wealth." (VERBURG, 2018, p. 14).

⁹⁹ "Moreover, they argued that commercial society helps to promote the "bourgeois" virtues of reliability, decency, cooperativeness, and so on – moral and social good that were comparatively lacking in pre-commercial societies." (RASMUSSEN, 2014, pp. 263-264). On retrouve une telle défense des vertus commerciales aujourd'hui par l'économiste Deirdre McCloskey (2006).

offrir un avantage ; l'impersonnalité évite que le bien-être matériel d'une personne dépende de la volonté arbitraire d'une personne particulière.¹⁰⁰

Loin de « contaminer » par des considérations instrumentales des relations personnelles qui étaient auparavant « pures », [Adam Smith] suggère que le commerce aide à libérer les relations personnelles de type de coercitions qui furent abondantes dans l'histoire de l'humanité.¹⁰¹

Il apparaît désormais clairement que la promotion de sociétés dans lesquelles l'égoïsme est la motivation dominante ne s'appuie pas uniquement sur l'exigence de « faire l'économie de la vertu », mais aussi sur l'exigence d'utiliser le désir de richesse pour réfréner les passions les plus dangereuses et pour encourager certaines vertus, tout en libérant les individus du joug de volontés arbitraires.

3.2. LA MORALE DES POLITICIENS

Dans la théorie économique néo-classique, il revient aux hommes et femmes politiques de s'assurer que, par la mise en place de règles et d'incitations, les intérêts individuels des acteurs économiques sont orientés de manière à réaliser les valeurs morales et le bien commun. Cette attribution des rôles présuppose que les acteurs politiques ne partagent pas les mêmes motivations que les acteurs économiques ; les premiers poursuivent des objectifs collectifs, alors que les seconds poursuivent des objectifs purement privés. Dans la théorie économique néo-classique, le politicien ou la politicienne

¹⁰⁰ Pierre-Yves Néron souligne que l'économiste Elisabeth Anderson et le philosophe républicain Philippe Pettit, eux aussi, soutiennent que le caractère impersonnel du marché offre l'opportunité de se libérer de la dépendance à la charité d'individus particuliers : « Comme le notent avec justesse Elisabeth Anderson et Philip Pettit, de telles interactions qui se fondent sur l'égoïsme du marchand participent peut-être à une validation de la poursuite d'intérêts étroits et personnels, mais ils permettent également aux personnes ordinaires de ne pas dépendre de manière servile de la « noblesse oblige » d'autrui. [...] Ce qui est important ici n'est donc pas tellement que les relations de marché se fondent sur l'égoïsme du marchand, mais surtout qu'elles ne se fondent pas sur la seule bienveillance de certains groupes de la société. » (NÉRON, 2016, pp. 198-199). En outre, il me semble que le livre du sociologue Georg Simmel sur l'argent attribue des avantages similaires aux relations commerciales: étant donné que les acteurs économiques sont « anonymes » et « interchangeables », et étant donné que l'argent peut servir à acheter une variété de choses différentes, l'individu se libère de sa relation *personnelle* à des personnes ou des choses particulières. Les relations hiérarchiques entre deux personnes sont remplacées par des relations égalitaires où seul l'argent, et non les caractéristiques personnelles, est important (SIMMEL, 1987).

¹⁰¹ "Far from "contaminating" previously pure personal relationships with instrumental concerns, [Adam Smith] suggests, commerce helps to free personal relationships from the type of coercion that have abounded throughout human history." (RASMUSSEN, 2014, p. 292).

se soucie exclusivement de réaliser le bien commun, alors qu'un dirigeant d'entreprise ne se soucie que de réaliser un profit personnel.

Les acteurs politiques – ainsi que les économistes-conseillers qui les accompagnent – sont ainsi décrits comme étant capables, dans la théorie économique, de renoncer à la recherche d'avantages personnels à la faveur du bien commun. Cette présupposition d'une vertu de l'acteur politique – au contraire de l'acteur économique – est nécessaire pour qu'il soit possible d'envisager un gouvernement qui organise, avec bienveillance, les règles et incitations d'une manière qui soit bénéfique à la collectivité.

Il existe toutefois une objection classique contre cette double description des motivations humaines. L'objection a donné naissance à un nouveau courant de la pensée économique, à savoir la théorie du « choix public ». ¹⁰² Selon cette objection, les économistes néo-classiques appliquent un « double standard ». D'un côté les acteurs économiques sont décrits comme mauvais et ne se préoccupant que de leur propre intérêt ; de l'autre côté, les acteurs politiques, et avec eux les économistes qui les conseillent, sont considérés comme motivés exclusivement par le souci du bien commun. Mais il n'y a pas de raison de penser que les politiciens ont des motivations plus nobles que les acteurs économiques.

L'analyse traditionnelle de l'intervention gouvernementale est un exemple manifeste d'application d'un double standard. On soumet le régime de marché à une analyse rigoureuse pour en déceler les faiblesses, mais on ne traite pas l'intervention gouvernementale sur le même pied. On préconise l'intervention en supposant toujours qu'elle donne les résultats escomptés. [...] Dans cette vision naïve de la réalité, on attribue au gouvernement la mission de veiller à l'intérêt public et on présume que toutes ses actions ont pour résultat le mieux-être de la collectivité. On se représente le gouvernement comme un despote bienveillant et omnipotent, dont toutes les actions seraient consacrées à l'intérêt public (RAYNAULD & STRINGER, 2003, p. 54).

L'analyse économique néo-classique ne fait en quelque sorte que déplacer le problème de naïveté ou d'angélisme à l'égard des motivations humaines. En effet, si cette analyse

¹⁰² L'ouvrage fondateur de ce courant est le livre de James Buchanan et Gordon Tullock *Le Calcul du Consentement* (BUCHANAN & TULLOCK, 1965).

satisfait à l'exigence de « faire l'économie de la vertu » dans la sphère des échanges économiques, ou même dans la sphère de la citoyenneté en générale, c'est uniquement en présumant que les règles et incitations de la société sont établies par un gouvernement bienveillant et vertueux. Or, il n'y a pas de raison de penser qu'une personne devient vertueuse, capable de renoncer à ses intérêts privés, du seul fait qu'elle occupe une fonction publique.

La nature d'un individu ne change pas du seul fait qu'il passe du secteur privé au secteur public. C'est le même individu, ayant les mêmes motivations et les mêmes préoccupations, mais effectuant un travail différent pour un nouvel employeur. Il n'y a aucune raison de croire qu'en passant au secteur public il cesse d'être mû par son intérêt personnel et n'agit qu'en fonction de l'intérêt public, au point de négliger ses propres intérêts (RAYNAULD & STRINGER, 2003, p. 55).¹⁰³

Puisque la sphère publique doit être considérée comme tout autant affectée par la défaillance morale de ses membres que la sphère privée, il convient également de mettre en place des règles d'action et des incitations pour le gouvernement. Ces règles et incitations doivent permettre, à nouveau, de garantir que les décisions égoïstes des acteurs publics soient conformes à la poursuite du bien commun.

Les règles doivent servir à imposer des contraintes au comportement des politiciens, bien qu'elles ne transforment pas la nature humaine. En d'autres termes, la supposition que les hommes sont malhonnêtes est maintenue, mais les règles du bon gouvernement au sein desquelles les hommes interagissent les uns avec les autres imposent une discipline à leur malhonnêteté, à tel point que les comportements malhonnêtes disparaissent complètement.¹⁰⁴

¹⁰³ Pour un exemple plus concret: « Pour quelle raisons les fonctionnaires et les politiciens seraient-ils des êtres différents des acteurs du secteur privé? Pourquoi l'administrateur d'un service de courrier privé différerait-il de l'administrateur des services de postes au point d'avoir des motivations différentes? Pour quelle raison l'enseignant d'une école privée différerait-il fondamentalement de l'enseignant d'une école publique? Par quelle vertu particulière l'employé de CFTM changerait-il subitement de nature du simple fait de passer à l'emploi de Radio-Québec? » (RAYNAULD & STRINGER, 2003, p. 55)

¹⁰⁴ "Rules must bind the behavior of politicians, even though they don't transform human nature. In other words, men remain presumed to be knaves, but the rules of good governance within which men interact with one another discipline their knavery to such an extent that knavish behavior is held in check to the point of nonexistence." (BOETTKE, 2012, p. 59). Peter Boettke est un défenseur contemporain de la théorie du choix public.

Selon les défenseurs de la « théorie du choix public », selon laquelle il faut *étendre* l'exigence de « faire l'économie de la vertu » à l'analyse de la sphère publique, sous peine à nouveau de faire preuve d'un angélisme naïf et potentiellement dangereux, le type d'institution qui permet le mieux d'assurer la conformité des choix politiques avec le bien commun est, à nouveau, l'institution du marché. Ils développent donc des modèles d'analyse dans lesquels les acteurs politiques ne sont pas des individus vertueux dont les choix seraient guidés par le souci du bien commun, mais des individus eux aussi égoïstes dont le seul but est d'obtenir le pouvoir et la gloire grâce à la « vente » d'un programme politique en l'échange d'un vote.

Les électeurs, les politiciens et les bureaucrates sont les intervenants du secteur public et ils cherchent à maximiser leur bien-être personnel à l'intérieur des contraintes qui leurs sont imposées. Les électeurs participent au processus en élisant des représentants politiques. Au moment du vote, ils se comportent par hypothèse comme des consommateurs. Ils votent pour le candidat qui leur propose les politiques les plus avantageuses ; ils appuient celui dont le programme contribue le plus à leur bien-être personnel. [...] Les politiciens sont en quelque sorte des entrepreneurs politiques : ils offrent à l'électorat des programmes dans le but d'obtenir des voix (RAYNAULD & STRINGER, 2003, p. 57).

S'il existe une différence entre l'acteur politique et l'acteur économique, ce n'est pas que le premier poursuit le bien commun et le second son intérêt purement individuel. Tous les deux poursuivent leurs intérêts propres. La différence provient uniquement du fait que, dans la sphère politique, l'intérêt individuel est constitué par le pouvoir et la gloire, alors que dans la sphère économique cet intérêt est constitué par la richesse matérielle.

En résumé, il est reproché aux économistes néo-classiques de limiter l'exigence de « faire l'économie de la vertu » à l'analyse de la seule sphère économique, sans l'étendre à l'analyse de la sphère publique. Selon les objecteurs, l'hypothèse de l'intérêt personnel exclusif doit être étendue à l'analyse de l'activité politique.¹⁰⁵

¹⁰⁵ Comme le fait remarquer Mark Nelson, cet argument à l'encontre du programme d' « économie de la vertu » rappelle un argument avancé par Bernard Williams à l'encontre de l'utilitarisme, à savoir l'argument de la « morale ésotérique » des politiciens (WILLIAMS, 2006, p. 109). Selon cet argument, il est indispensable, dans une société utilitariste, que seuls les politiciens soient motivés à agir *dans le but* de réaliser les objectifs utilitaristes. La masse populaire, de son côté, ne devrait pas être informée du programme utilitariste, car il est

Avant de procéder à la présentation de l'objection suivante, formulons une remarque: aussi longtemps qu'ils acceptent que l'exigence de « réalisme » à l'égard des motivations humaines requiert de faire l'hypothèse d'un égoïsme universel, les économistes néo-classiques devraient, par souci de cohérence, étendre leur analyse également à la sphère publique. Toutefois, cette extension de l'hypothèse de l'égoïsme vers l'activité politique ne permettra pas de répondre à toutes les autres objections contre l'usage de l'hypothèse de l'intérêt personnel dans les théories politiques. L'extension de l'hypothèse ne permet de répondre ni aux objections concernant les conséquences potentiellement néfastes de l'hypothèse sur la vie sociale, ni aux deux objections que nous présenterons dans les sections suivantes.

3.3. L'ÉGOCENTRISME

Les hommes ne sont rien. Et ne laissent aucune trace. Raffaele quittait Montepuccio et tous les hommes sur son passage enlevèrent leur chapeau et baissèrent la tête, conscients qu'à leur tour, ils ne tarderaient pas à disparaître et que cela ne ferait pas pleurer les oliviers.
(GAUDÉ, 2004, p. 228)

Il existe une autre version de l'objection selon laquelle l'hypothèse de l'égoïsme que l'on trouve dans la théorie économique est trop « naïve » ou « angélique » à l'égard des motivations humaines. Selon cette nouvelle objection, la théorie ne tient pas compte du fait que l'être humain est un individu hautement *égocentrique*. De ce fait, elle ne respecte pas non plus l'exigence d'adopter une conception *pessimiste-réaliste* de la nature humaine comme fondement d'un modèle politique.

3.3.1. L'ÉGOCENTRISME

Une personne égoцентриque est une personne qui se prend pour le centre du monde. Elle accorde à sa propre existence une valeur excessive et évalue les événements du monde

avantageux, pour le bien collectif, que la population ne poursuive pas directement ce bien. En conséquence, les politiciens et la masse du peuple ne sont pas soumis aux mêmes exigences morales, les uns devant poursuivre directement des objectifs utilitaristes et les autres non. Les politiciens sont ainsi soumis à leur propre morale, la *morale ésotérique des politiciens*.

à la lumière de ce qui se produit dans sa propre vie. En résumé, pour un égocentriste, son existence est la mesure, le critère d'évaluation de l'état du monde en général.

Du fait qu'il se considère comme la mesure de toute chose, l'égocentriste considère que les autres personnes ont le devoir de se soucier de ce qui lui arrive. Il croit que ce qui se produit dans son existence a de l'importance du point de vue d'autrui. En fait, plus un individu croit que son existence a de l'importance *pour les autres personnes*, plus il croit que sa propre vie compte *pour les autres*, plus un individu est égocentrique.

Il semblerait que nous soyons tous, dans une certaine mesure, égocentriques. Par exemple, le fait d'éprouver un sentiment d'injustice lorsque, faisant face à des événements tragiques dans nos existences, nous constatons que le reste du monde continue de vivre normalement, constitue une réaction fortement égocentrique. En effet, nous réagissons comme si la manière dont se déroule notre vie devait recevoir l'attention du reste de la population. La perte d'un être cher, une séparation amoureuse, l'apparition d'un handicap ou d'une maladie, l'expérience d'un désastre naturel, la perte d'un emploi, etc., troublent énormément une existence individuelle. Face à de tels événements, le fait que la vie des autres personnes continue son cheminement normal peut paraître incompréhensible, injuste, et susciter une souffrance supplémentaire. Comme si, de notre point de vue, ce qui se passe dans nos vies devait immanquablement préoccuper les autres personnes, même celles qui nous sont inconnues.

Il existe différents degrés d'égocentrisme en fonction de la manière dont un individu évalue l'importance de sa propre existence pour autrui. Au niveau le plus faible, un individu peut penser que sa vie est dépourvue de toute valeur pour autrui. Sa vie n'a de valeur que de son propre point de vue, une valeur *relative à son propre point de vue*, mais elle est sans valeur pour autrui – à l'exception, éventuellement, de ses proches. Un tel individu est faiblement égocentrique et n'accorde à sa propre vie qu'une faible valeur.

Comme le décrit bien Thomas Nagel, si nous, les êtres humains, accordions seulement une telle valeur relative à notre existence, cela signifierait que [...]

[...] nos vies et les choses qui ont de l'importance dans nos vies (ce qui comprend les vies des personnes auxquelles nous accordons de la valeur) ne seraient pas

considérées comme ayant de la valeur indépendamment du fait qu'elles sont les nôtres, ou indépendamment du fait qu'elles nous concernent personnellement.¹⁰⁶

Au niveau le plus élevé, cette fois, un individu peut considérer que son existence a de la valeur non seulement *de son propre point de vue*, mais aussi du point de vue des autres personnes – même du point de vue de personnes inconnues. Il considère que sa vie n'a pas de la valeur seulement *pour lui*, relative à son point de vue, mais que sa vie a de la valeur *pour les autres*, et même pour le monde en général. Cet individu croit que le monde dans son ensemble est meilleur ou moins bon en fonction de ce qui se produit dans son existence à lui. De ce fait, il considère que les autres personnes *devraient se soucier* de ce qui lui arrive, puisque ce qui lui arrive influence la valeur que possède le monde dans lequel elles vivent. Une telle personne est *hautement égocentrique*.

3.3.2. LA VALEUR DES INDIVIDUS DANS DIFFÉRENTS SYSTÈMES MORAUX

Nous avons défini l'égocentrisme comme étant une disposition à se considérer soi-même comme le centre du monde et à croire que ce qui se produit dans nos vies a de l'importance indépendamment de notre propre point de vue. Nous avons vu également qu'il existe au moins deux niveaux d'égocentrisme : un individu fortement égocentrique pense que sa vie devrait être source de préoccupation pour autrui alors qu'un individu peu égocentrique pense que sa vie ne possède pas de valeur indépendamment de son propre point de vue et qu'il est donc acceptable qu'autrui ne lui accorde pas d'importance.

La question qui se pose alors est la suivante : quel est le système moral qui reconnaît et attribue à l'individu égocentrique la valeur que cet individu croit posséder ? Autrement dit, quel système moral reconnaît à notre individu la haute estime qu'il a de sa propre existence ? Dans les prochaines lignes, nous allons tenter d'identifier, parmi quelques systèmes moraux existants, ceux qui sont adaptés à différents degrés d'égocentrisme. Nous allons ainsi étudier trois systèmes moraux qui chacun attribue à l'existence humaine une valeur différente. Nous verrons alors qu'un individu hautement égocentrique devrait

¹⁰⁶ "[...] our lives and what matters to us as we live them (including the lives of other people we care about) would not be regarded as mattering at all if considered apart from the fact that they are ours, or personally related to us." (NAGEL, 2008 p. 110).

préférer vivre dans une société régie par un système moral déontologique plutôt que de vivre dans une société régie par un système moral égoïste ou utilitariste.

La valeur attribuée à la vie d'une personne est déterminée par la manière dont il est considéré comme *acceptable* ou *inacceptable* de traiter cette personne. Comparons en ce sens trois doctrines morales qui définissent trois manières différentes dont il est moralement acceptable ou inacceptable de traiter les personnes : (1) l'égoïsme éthique, (2) l'utilitarisme et (3) la doctrine déontologique des droits fondamentaux. Nous nous concentrons sur ces trois théories car elles portent directement sur les relations interpersonnelles (contrairement à l'éthique de la vertu) et elles attribuent chacune une valeur différente à la vie humaine, ce qui facilite la comparaison. Bien entendu, la liste des doctrines morales n'est pas exhaustive.

Les manières dont il est acceptable de traiter une personne dans chacune des doctrines morales susmentionnées sont les suivantes :

(1) L'égoïsme éthique

L'égoïsme éthique est une doctrine morale selon laquelle il est toujours *légitime* et même *obligatoire* pour un individu de poursuivre ses intérêts individuels. Dans une telle doctrine, il est ainsi toujours légitime de sacrifier les intérêts d'une personne afin de réaliser ses propres intérêts ; il est autorisé de tuer, de torturer, de maltraiter, d'exploiter autrui, etc., si cela permet d'améliorer son existence – ou celle de ses proches. Autrement dit, tout individu peut *légitimement* sacrifier les intérêts d'une autre personne pour satisfaire ses propres intérêts. La théorie morale de l'égoïsme éthique rend ainsi légitime l'exploitation de l'homme par l'homme pour la satisfaction d'intérêts personnels.¹⁰⁷

¹⁰⁷ Bien entendu, les partisans de l'égoïsme éthique essaient de montrer que, dans les circonstances du monde tel que nous le connaissons, il est contraire à l'intérêt personnel bien compris de tuer, insulter ou exploiter autrui. Selon eux, l'égoïsme éthique requiert de poursuivre son « intérêt bien compris », ou intérêt objectif, plutôt que de faire ce qui nous plaît. Or, selon eux, poursuivre son intérêt bien compris nécessite d'adopter des principes d'action incompatibles avec le fait de porter atteinte à autrui. Nous trouvons ce type d'argument notamment chez Ayn Rand : « L'éthique Objectiviste soutient que l'acteur doit toujours être le bénéficiaire de ses actions et qu'un homme doit toujours agir en vue de son propre intérêt personnel rationnel. Mais son droit d'agir de cette manière [...] n'est applicable que dans le contexte d'un code de principes moraux rationnels, objectifs, démontrés et validés. Il ne s'agit pas d'une autorisation à « faire ce qui lui plaît. » [“The Objectivist ethics holds that the actor must always be the beneficiary of his action and that man must act for his own rational self-interest. But his right to do so [...] is applicable only in the context of a rational, objectively demonstrated and validated code of moral principles which define and determine his actual self-interest. It is not a license “to do as he pleases” [...] (RAND, 1967, p. x). Pour notre propos, cette réponse importe peu. En effet, puisque la valeur des personnes est déterminée par ce qu'il est moralement acceptable de leur faire

Une personne qui adhère à la doctrine de l'égoïsme éthique considère, en ce sens, qu'il est toujours moralement *acceptable* que sa propre vie soit sacrifiée pour satisfaire les intérêts de quelqu'un d'autre. Dans les termes de Thomas Nagel :

L'égoïsme [éthique] autorise les autres, non seulement à ignorer mes intérêts, mais il les autorise à m'utiliser d'une manière qui servira leurs propres intérêts.¹⁰⁸

Une personne qui adhère à la doctrine morale de l'égoïsme éthique n'attribue donc à sa propre existence qu'une valeur *relative*. Elle pense que sa vie possède de la valeur *de son propre point de vue*, mais qu'elle n'en possède pas pour autrui – puisqu'elle considère comme moralement acceptable d'être sacrifiée par autrui pour servir leurs intérêts. Une personne adhérant à l'égoïsme éthique n'est donc pas égocentrique, puisqu'elle ne croit pas que sa vie soit importante pour les autres.

(2) L'utilitarisme

L'utilitarisme est la doctrine morale selon laquelle il faut maximiser la quantité totale de bien-être (ou de bonheur) dans le monde. Pour chaque état du monde possible, les quantités de bien-être des différentes personnes sont additionnées les unes aux autres et l'état du monde qui contient la plus grande quantité totale de bien-être est considéré comme étant l'état du monde le meilleur du point de vue moral.

Dans l'utilitarisme, contrairement à l'égoïsme éthique, la vie de chaque personne compte du point de vue impersonnel. En effet, toutes choses égales par ailleurs, lorsque le bien-être d'une personne est positivement affecté, cela augmente la quantité totale de bien-être dans le monde et rend de ce fait le monde meilleur du point de vue moral. La quantité de bien-être contenue dans une existence individuelle est donc importante du point de vue impersonnel, en tant qu'elle augmente ou diminue la quantité totale de bien-être. Autrement dit, le bien-être de chaque personne compte en tant que composant de la valeur impersonnelle du bonheur général de la société.

Dans la doctrine utilitariste, la vie des personnes reçoit une valeur supérieure à celle qu'elle reçoit dans la doctrine de l'égoïsme éthique. En effet, la vie des personnes – et ce qui

subir, et non par le traitement qu'elles subissent effectivement, une morale de l'égoïsme accorde dans tous les cas une valeur très faible à l'existence humaine.

¹⁰⁸ "Egoism permits others not merely to disregard my interests, but to use me in any way that will serve their aims." (NAGEL, 2008, p. 111).

s’y produit – compte du point de vue impersonnel et constitue de ce fait un objet légitime de préoccupation pour tout le monde. De ce fait, un partisan de l’utilitarisme, contrairement au partisan de l’égoïsme éthique, reconnaît à sa propre existence une valeur d’un point de vue *impersonnel*, en tant qu’elle contribue à l’augmentation du bien-être général. Un utilitariste pense donc que sa propre vie *compte pour autrui*, en tant qu’elle participe à l’augmentation du bien-être de la société. Un adhérent à l’utilitarisme est donc plus égocentrique qu’un adhérent à l’égoïsme éthique.

La reconnaissance de valeurs impersonnelles [comme le bien-être général] représente une avancée par rapport à l’égoïsme pur en termes de la valeur que nous attribuons à notre propre existence, puisque ce qui nous arrive devient un objet de préoccupation pour tout le monde.¹⁰⁹

Toutefois, la valeur que l’utilitariste attribue à sa propre existence reste modérée. Le problème provient du fait que, dans la perspective utilitariste, une vie peut toujours, comme dans l’égoïsme éthique, légitimement être sacrifiée. La seule différence est qu’au lieu qu’elle puisse être légitimement sacrifiée pour satisfaire des intérêts privés, elle peut seulement être sacrifiée dans le but de réaliser la valeur impersonnelle du bien-être général.

[U]ne morale [...] comme l’utilitarisme [...] implique que n’importe qui a le droit de me sacrifier ou de m’utiliser au service de ces valeurs impersonnelles, c’est-à-dire au service du bien-être *général*.¹¹⁰

Par exemple, si la mise à mort d’une personne est le seul moyen d’en sauver cinq autres, si la torture d’une personne est le seul moyen d’éviter l’assassinat d’un président dont les actions sont bénéfiques pour la société dans son ensemble, si l’exploitation des femmes dans la prostitution augmente le bien-être général – en satisfaisant les désirs sexuels des hommes et en diminuant la pression sur les épouses concernant leurs devoirs conjugaux, si l’exploitation des travailleurs permet d’améliorer la prospérité économique d’une société, si la condamnation d’innocents permet de diminuer le taux total de

¹⁰⁹ “The admission of [...] impersonal values [such as the general welfare] is an advance over pure egoism in the evaluation of oneself, for it makes what happens to one a matter of impersonal concern to anyone.” (NAGEL, 2008, p. 111).

¹¹⁰ [A] morality [...] such as utilitarianism, implies that anyone may sacrifice me or use me in the service of those values, that is, in the service of the general welfare. (NAGEL, 2008, p. 111).

criminalité, etc., alors les sacrifices des vies individuelles doivent être consentis au nom de la valeur impersonnelle du bien-être général.

Un individu utilitariste considère ainsi qu'il est moralement acceptable que sa vie soit sacrifiée au nom d'une valeur impersonnelle générale. Il croit, certes, que sa vie possède une valeur *pour autrui*, dans le sens où elle contribue à une valeur impersonnelle qui est un objet approprié de préoccupation pour les autres, mais il croit néanmoins que sa propre vie peut légitimement être sacrifiée si cela permet d'augmenter le bien-être général. Il ne croit pas que sa vie possède de la valeur indépendamment de sa contribution à la valeur impersonnelle du bien-être général.¹¹¹ Il ne croit pas que sa vie possède une valeur *intrinsèque*. De ce fait, le partisan de l'utilitarisme est *modérément* égocentrique : il est plus égocentrique que le partisan de l'égoïsme éthique, puisqu'il croit que sa vie possède une valeur impersonnelle. Cependant, comme nous allons le voir, il est moins égocentrique qu'un partisan de l'éthique déontologique, puisque l'utilitariste ne croit pas que sa vie possède une valeur intrinsèque.

(3) Le déontologisme

L'éthique déontologique définit un certain nombre de règles morales que les individus doivent respecter en toutes circonstances ; par exemple, les individus ont le devoir moral de ne pas tuer autrui, de ne pas porter atteinte à la propriété d'autrui, de ne pas violer, de ne pas mentir, de ne pas porter atteinte à l'intégrité physique ou mentale d'une autre personne, etc. Ils doivent respecter ces règles indépendamment des circonstances ; il est toujours vrai, selon l'éthique déontologique, qu'il est moralement interdit de tuer autrui, de le torturer ou d'abuser de lui.

Selon la morale déontologique, il n'est jamais légitime de sacrifier les personnes. C'est pourquoi celles-ci doivent être protégées par des *droits* fondamentaux interdisant que l'on puisse les sacrifier pour satisfaire des intérêts privés ou collectifs. Chaque vie humaine possède donc non seulement une valeur d'un point de vue impersonnel, mais elle possède,

¹¹¹ On peut considérer que cette incapacité de rendre compte du fait que, en cas de viol, par exemple, c'est une *personne* qui a subi un tort, plutôt que le bien-être général, constitue un défaut la doctrine utilitariste. En effet, dans la pensée morale ordinaire, nous considérons que, lorsqu'une femme, par exemple, subit un viol, c'est *cette femme particulière* qui a subi un tort. Nous pensons que l'agresseur a fait du mal à *la femme* particulière, et non pas qu'il a porté atteinte à la valeur impersonnelle du bien-être général. Or, il est possible que l'utilitarisme ne puisse pas rendre compte de ce phénomène. Sur ce sujet, voir l'article de Mark Nelson (2015).

en outre, une valeur *intrinsèque*. La vie d'une personne a de la valeur en elle-même, indépendamment de sa contribution au bien-être d'autrui et indépendamment de sa contribution à une valeur impersonnelle comme le bien-être général.

Comme le fait remarquer Nagel, l'attribution de droits fondamentaux confère à chaque personne un statut moral *d'inviolabilité* ; une personne est considérée comme *inviolable* précisément lorsque certaines manières de la traiter sont considérées comme *toujours* inacceptables, indépendamment des conséquences de cette interdiction sur le bien-être général ou sur des valeurs collectives.

Bien entendu, avoir le statut moral d'inviolabilité ne signifie pas que *dans les faits*, les droits fondamentaux de la personne ne *sont* jamais bafoués. Il est en effet tout à fait possible que, malgré l'existence d'un droit fondamental à ne pas être torturé, *dans les faits* la torture existe néanmoins. Avoir un statut moral d'inviolabilité signifie plutôt que, à chaque fois qu'un droit fondamental est violé, peu importe que cela ait servi les intérêts d'autrui ou que cela ait permis de réaliser une valeur impersonnelle générale, une faute morale a été commise – et mérite d'être sanctionnée.

Etre inviolable n'est pas un état, comme être heureux, ou être libre – tout comme la propriété de pouvoir légitimement être abusé n'est pas un état, comme être malheureux ou être opprimé. Etre inviolable ne signifie pas que l'on ne nous portera pas atteinte. Il s'agit d'un statut moral: il signifie qu'il n'est pas légitime de traiter une personne d'une certaine manière – qu'un tel traitement n'est pas admissible, et que, s'il se produit néanmoins, alors la personne a été illégitimement lésée.¹¹²

Selon la morale déontologique, il est toujours vrai qu'une personne torturée, violée, maltraitée subit un tort moral, peu importe les circonstances.

Un partisan d'une morale déontologique considère donc que sa propre vie mérite l'attention des autres personnes *pour elle-même*, indépendamment de sa contribution à l'augmentation du bien-être général. Il croit qu'un *mal intrinsèque* se produit si sa vie est

¹¹² "Being inviolable is not a condition, like being happy, or free—just as being violable is not a condition, like being unhappy or oppressed. To be inviolable does not mean that one will not be violated. It is a moral status: It means that one may not be violated in certain ways—such treatment is inadmissible, and if it occurs, the person has been wronged. So someone's having or lacking this status is not equivalent to anything's happening or not happening to him." (NAGEL, 1970, p. 107)

sacrifiée et que l'existence de ce mal intrinsèque est l'objet approprié de la préoccupation *des autres personnes*. Un partisan de la morale déontologique est donc hautement égocentrique, puisque, de son point de vue, ce qui se produit dans sa vie importe aux autres personnes, indépendamment de la contribution de sa vie à une valeur impersonnelle générale.

La morale déontologique semble bien rendre compte de la manière dont nous évaluons nos existences ; nous pensons que notre vie importe pour elle-même, qu'elle a de la valeur en soi. C'est pour cette raison que nous ressentons comme une injustice la négligence ou l'indifférence dont peuvent faire preuve les autres personnes à l'égard de ce qui se produit dans nos vies ; nous pensons mériter leur attention du fait que notre vie est importante en elle-même, et qu'elle est, de ce fait, importante pour ces personnes également, qu'elle est un objet approprié de leur attention. Nous considérons comme moralement inacceptable que notre vie soit sacrifiée, même si cela permet d'augmenter le bonheur général, car nous considérons que notre vie est *en soi* un objet légitime de préoccupation et d'attention. Nous avons donc une conception hautement égocentrique de l'importance de notre propre existence et devrions, en ce sens, adhérer à une morale déontologique, puisque celle-ci représente adéquatement la valeur que nous pensons avoir.

3.3.3. L'ÉGOCENTRISME ET LA THÉORIE ÉCONOMIQUE

Un individu hautement égocentrique considère que sa vie possède une valeur intrinsèque et, de ce fait, mérite l'attention des autres personnes. La question est désormais de savoir si la théorie économique représente adéquatement cette conception *pessimiste-réaliste*¹¹³ de la nature humaine. Dans ce contexte, il convient de partager l'approche *positive* et l'approche *normative* de la science économique.

Dans son approche positive, la théorie économique adopte une certaine conception des motivations humaines et établit des planifications et prévisions sur la base de cette conception. Nous avons vu que la théorie économique admet, sur ce point, l'hypothèse de

¹¹³ Cette conception est considérée comme *pessimiste* dans la mesure où la tendance à se prendre pour le centre du monde, et à évaluer le monde à la lumière de ce qui se produit dans sa propre vie, est considérée comme un défaut de caractère. Bien entendu, dès lors qu'il a été montré que la conception morale qui répond le mieux à cette conception de l'importance de sa propre vie est l'éthique déontologique, il semble difficile de soutenir encore qu'il s'agit d'une conception pessimiste de la nature humaine. Il s'agit toutefois précisément de la force de l'argument : plus un individu accorde de l'importance à sa propre vie, plus il est égocentrique, plus il convient pour lui de préférer vivre dans un monde régi par une éthique stricte.

l'égoïsme universel, selon laquelle les individus n'accordent pas de valeur au bien-être d'autrui, sauf lorsque satisfaire les intérêts des autres permet de satisfaire leurs propres intérêts.

Dans son approche normative, la théorie économique définit quels sont les objectifs sociaux qui doivent être poursuivis par le gouvernement et les institutions publiques et offre des critères d'évaluation des différentes politiques sociales et économiques envisagées. Bien qu'il existe d'importantes discussions au sujet du choix des critères normatifs, les critères les plus utilisés sont dérivés de l'approche utilitariste dont la théorie économique est l'héritière historique. En effet, les critères de maximisation des fonctions d'utilité ou d'optimalité de Pareto sont des critères qui se préoccupent principalement de la quantité totale de bien-être ou de la satisfaction globale des préférences, plutôt que de la protection individuelle des personnes.

La question que nous nous posons dans ce contexte est la suivante : quelle serait les modifications à apporter à la théorie économique si celle-ci devait représenter les êtres humains comme hautement égocentriques, c'est-à-dire comme accordant une valeur intrinsèque à leur propre existence ?

L'approche positive admet l'hypothèse de l'égoïsme universel, qui concerne la valeur attribuée à *l'existence des autres personnes*, mais la théorie ne se prononce pas explicitement sur *l'égocentrisme* des individus, qui concerne la valeur attribuée à *sa propre existence*. Autrement dit, l'individu égoïste de la théorie économique ne semble pas avoir d'opinion explicite sur la valeur de sa propre existence, sur la manière dont il croit qu'il est acceptable ou inacceptable d'être traité par les autres personnes. Toutefois, plusieurs éléments suggèrent implicitement que l'individu représenté dans la théorie n'est pas hautement égocentrique. En effet, la description de l'individu comme égoïste, c'est-à-dire comme indifférent au bien-être des autres personnes, ne semble compatible qu'avec *deux* attitudes à l'égard de la morale : (a) l'absence de morale et (b) l'éthique de l'égoïsme.

(a) L'absence de morale

Celui qui n'accorde pas de valeur au bien-être des autres peut simplement *ne pas avoir* de croyances morales. Il ne considère ni qu'il est légitime de sacrifier autrui pour satisfaire ses intérêts privés, ni qu'il est illégitime de le faire. Un tel individu sacrifie effectivement la

vie d'autrui lorsque cela sert ses intérêts, mais il ne croit pas, *en plus*, qu'il a le *droit*, qu'il est *moralement légitime* de le faire. Notre individu n'a tout simplement pas de pensées morales.

Le problème provient alors du fait que notre individu n'a en fait *pas du tout de conception* de la valeur de sa propre vie. Il n'existe pas, selon lui, d'attitudes que les autres personnes *n'ont pas le droit* d'adopter à son égard. Il n'existe pas de traitements qu'il est moralement inacceptable de lui faire subir. Il n'a pas non plus d'idée de *mériter* certains égards ou certaines considérations de la part des autres personnes. Cela implique que, du point de vue de cette personne, sa vie n'a simplement pas de valeur ; il n'existe aucun traitement qu'il est acceptable ou inacceptable de lui faire subir. Son égoïsme est nul.

(b) L'éthique de l'égoïsme

Selon la seconde conception de la valeur que l'individu de la théorie économique attribue à sa propre existence, celui-ci croit qu'il a *le droit* d'utiliser, de maltraiter, de sacrifier autrui pour satisfaire ses propres intérêts. Non seulement il le *fait*, comme l'admet la théorie économique, mais il pense qu'il a le *droit*. Notre individu entretient donc bien des pensées proprement morales. Toutefois, comme il semble évident, la théorie morale qui correspond à cette attitude est la doctrine de *l'égoïsme éthique*, selon laquelle chacun a le droit d'utiliser autrui pour satisfaire ses propres intérêts. Or, comme nous l'avons vu, adhérer à une morale égoïste revient à dévaluer sa propre existence, puisque l'on considère que les autres *ont le droit* de nous utiliser pour satisfaire leurs propres intérêts. Et cette dévaluation de notre propre existence se produit même si, dans les faits, nous avons les moyens de nous protéger contre les atteintes d'autrui. Comme nous l'avons vu, un individu qui considère ainsi qu'il est *légitime* de le sacrifier pour satisfaire des intérêts privés est *faiblement* égoцентриque. En conséquence, la théorie économique *positive* n'est compatible au final qu'avec un égoïsme *nul* ou un égoïsme *faible*. Elle ne semble pas compatible avec un égoïsme modéré ou fort.

3.3.4. L'OBJECTION ROYALISTE

On pourrait nous objecter ceci : l'individu *maximalement* égoцентриque non seulement pense que sa propre vie possède une valeur intrinsèque, et donc qu'il est inacceptable de la sacrifier, mais il pense aussi qu'il est acceptable de sacrifier la vie de *tous les autres* individus – contrairement à la sienne. Autrement dit, il pense que sa propre vie possède une valeur

intrinsèque, mais que la vie des autres personnes n'a pas une telle valeur. Un tel individu adopte un point de vue moral que nous pourrions qualifier de *royaliste* – i.e. point de vue selon lequel la vie du « roi » a une valeur intrinsèque, celle des autres personnes une valeur purement instrumentale.

Cependant, même un tel individu royaliste ne pose pas de difficulté à notre théorie. En effet, il existe plusieurs manières d'interpréter cette doctrine, et aucune de ces manières ne soulève de difficulté insurmontable pour nous.

Une première possibilité est que la doctrine morale royaliste soit une doctrine indexicale : notre royaliste considère que « moi » a le droit d'utiliser et sacrifier autrui pour ses propres intérêts, mais qu'il est moralement inacceptable que « moi » soit utilisé ou sacrifié pour satisfaire les intérêts d'autrui ou pour promouvoir l'intérêt général. Dans ce cas, il convient de faire remarquer plusieurs éléments : en premier lieu, si chacun adopte la même doctrine morale, attribuant une signification différente à l'indexical « moi », alors la vie de notre individu n'aura en fait d'importance que pour lui-même, comme dans la doctrine égoïste. En effet, les autres personnes ne reconnaissent pas de valeur à l'existence du royaliste et pensent qu'il est moralement acceptable de le sacrifier pour leurs propres intérêts. Dans ce cas, la doctrine royaliste se réduit à la doctrine de l'égoïsme éthique.

Une autre possibilité serait que notre royaliste considère bien qu'il a le droit de sacrifier la vie d'autrui, mais que les autres n'ont pas le droit de sacrifier la sienne, sans pour autant que sa doctrine morale soit une doctrine indexicale. Par exemple, il pense que « le roi » a le droit de sacrifier les existences des autres personnes pour ses propres intérêts, mais que les autres personnes n'ont pas le droit de sacrifier les intérêts du « roi ». Etant lui-même le roi en question, il croit en conséquence que qu'il a le droit de sacrifier autrui et que les autres n'ont pas le droit de le sacrifier. Mais sa théorie n'est pas indexicale, car la référence du « roi » ne change pas selon que le terme soit utilisé par une personne ou par une autre.

Dans ce cadre, il semble que la valeur de la vie des personnes – même celle du roi lui-même – soit déterminée par leur contribution à une valeur impersonnelle – ici « la royauté » – et qu'elles ne possèdent pas de valeur intrinsèque. La valeur que le roi attribue à son existence correspond en ce sens à la valeur modérée que l'utilitariste attribue à sa propre existence : celle-ci n'a de valeur que dans la mesure où elle contribue à une valeur

impersonnelle – la royauté ou le bien-être général. En effet, ce n'est que parce qu'il s'avère que notre individu occupe *lui-même* la fonction du roi qu'il croit que sa vie possède une valeur supérieure, mais ce n'est pas parce que cette vie est simplement *la sienne*. Autrement dit, s'il n'occupait pas la fonction du roi, il n'attribuerait pas de valeur intrinsèque à son existence. Cela prouve bien que, ce qui importe, c'est en fait la valeur de la « royauté ».

Pour finir, notons encore ceci : la valeur reconnue à notre vie est fonction de la manière dont il est considéré comme moralement acceptable ou inacceptable de traiter notre personne. De ce point de vue, la valeur attribuée à notre vie n'est pas différente dans une doctrine morale *royaliste* ou dans une doctrine morale *déontologique*. Dans les deux cas, il est moralement inacceptable d'être maltraitée, torturée, violée ou tuée par autrui. La valeur de notre vie est donc identique dans les deux doctrines. La valeur de notre vie n'est ainsi pas supérieure dans une doctrine *royaliste* que dans une doctrine déontologique. La seule différence se situe dans la valeur attribuée à la vie *des autres personnes*, mais non dans la valeur attribuée à notre propre vie. En résumé, l'égoïsme du royaliste et l'égoïsme du déontologiste sont identiques.

3.3.5. L'ÉCONOMIE NORMATIVE

En ce qui concerne la théorie économique *normative*, il est plus aisé de voir qu'elle est incompatible avec la conception hautement égoïste de l'être humain. En effet, la théorie préconise une conception *utilitariste* de la valeur des vies humaines, dans laquelle la valeur de la vie d'une personne est déterminée par sa contribution à la valeur du bien-être général.¹¹⁴ Or, nous l'avons vu, un individu hautement égoïste ne pourrait se satisfaire

¹¹⁴ Le critère d'optimalité de Pareto partage la caractéristique de l'utilitarisme d'évaluer la valeur de la vie des personnes en fonction de la quantité de bien-être que cette vie contient – mesuré par la satisfaction des préférences. Il ne partage toutefois pas la caractéristique *d'autoriser les sacrifices* de bien-être d'une personne en faveur d'une augmentation plus grande du bien-être d'une autre personne, puisque les comparaisons interpersonnelles d'utilité sont proscrites. Toutefois, cela ne signifie pas que la valeur morale des personnes soit plus élevée dans la doctrine parétienne que dans la doctrine utilitariste. Le problème est, tout d'abord, que ces sacrifices ne sont pas interdits *pour des raisons morales* – comme il est le cas dans la morale déontologique – mais pour des raisons épistémologiques – concernant l'impossibilité de connaître les quantités d'utilité ou concernant le caractère non scientifique des énoncés moraux permettant d'établir ces comparaisons. La valeur accordée à la vie des personnes n'est donc pas supérieure dans une morale parétienne que dans une morale utilitariste.

En second lieu, l'éthique parétienne peut, par exemple, considérer comme légitime que la situation d'une personne soit hautement améliorée par le travail d'une personne œuvrant dans des conditions misérables. A titre d'illustration, nous pouvons mentionner le cas très discuté du caractère économiquement avantageux de la délocalisation des industries polluantes des pays développés vers les pays en voie de développement. Cette délocalisation, qui satisfait au critère de Pareto, avait été défendue, dans les années

d'une morale utilitariste, car celle-ci ne reconnaît pas la valeur intrinsèque de sa propre existence. Un individu hautement égocentrique ne pourrait pas accepter de vivre dans un monde où, si les circonstances le demandent, il serait considéré comme moralement acceptable de sacrifier sa vie pour le bien-être général.

En conclusion, aussi bien dans sa partie descriptive que dans sa partie normative, la théorie économique adopte une conception de l'être humain ou de la valeur de la vie humaine qui ne correspond pas à la valeur que des individus *fortement égocentriques* attribueraient à leur propre existence. En ce sens, nous pouvons la considérer comme encore trop angélique.

Comme le remarque Nagel, notre résultat peut sembler paradoxal : plus une personne est *égocentrique*, plus elle considère que sa vie a une valeur élevée, plus cette personne aura tendance à adopter un point de vue moral déontologique. Autrement dit, plus elle est égocentrique, plus il est approprié pour elle d'adopter une morale impartiale et universelle.

Ce qui nous guide en direction de l'universalisation, c'est la difficulté que rencontre chaque personne pour se regarder elle-même comme n'ayant de la valeur que pour elle-même, mais pas en elle-même. Si les personnes ne sont pas des fins en soi – c'est-à-dire, si elles n'ont pas de valeur impersonnelle – elles appartiennent à un ordre bien inférieur de valeur. L'égoïsme [éthique] implique de se dévaluer soi-même, en même temps que l'on dévalue tous les autres.¹¹⁵

Malgré son apparence paradoxale, il s'agit d'un résultat intéressant ; il est envisageable que les dispositions morales des personnes s'expliquent par la force de leur égocentrisme. Si tel est le cas, alors le fait de s'accorder à soi-même une valeur élevée ne porte pas préjudice, mais stimule plutôt, la motivation morale.

1990, par le chef économiste de la Banque Mondiale de l'époque, Lawrence Summers. Le journal *The Economist* (Economist, 1992) avait révélé l'affaire. Or, il semble moralement inacceptable que, par contrainte économique, les pays pauvres doivent subir les nuisances de la consommation des pays riches.

¹¹⁵ "[W]hat drives us in the direction of universalizability is the difficulty each person has in regarding himself as having value only for himself, but not in himself. If people are not ends in themselves—that is, impersonally valuable—they have a much lower order of worth. Egoism amounts to a devaluation of oneself, along with everyone else." (NAGEL, 2008 p. 110).

3.4. LA NORME MORALE FONDAMENTALE DU MARCHÉ

Nous allons, dans cette section, présenter une dernière variante de l'objection de « naïveté » ou de « manque de réalisme » qui peut être formulée à l'encontre du programme de « mécanisation » de la vertu. Pour rappel, ce programme consiste à obtenir des membres de la société un comportement extérieurement vertueux sur la base de motivations exclusivement égoïstes, tels que, notamment, l'appât du gain ou le désir de richesse. Il s'agit donc, pour les partisans de ce programme, d'envisager l'organisation de la société en faisant « l'économie de la vertu », c'est-à-dire en se dispensant de faire appel aux dispositions morales des individus. Dans cette perspective, un projet de société ou un programme politique qui dépendrait, pour réussir, de la motivation morale des individus est considéré comme utopiste et irréalisable, du fait qu'il repose sur une conception idéalisée, angélique et naïve de la nature humaine.

La critique de ce programme de « mécanisation de la vertu » que nous allons étudier maintenant prend la forme suivante : il n'est pas possible de se dispenser entièrement des dispositions morales des individus. En effet, même l'institution du marché, censée incarner l'institution amoralisée par excellence, repose en fait sur l'acceptation, par les acteurs économiques, d'une norme morale fondamentale. Pour que le marché mène à des résultats collectivement avantageux, il est nécessaire que les acteurs économiques, et notamment les entreprises, soient motivés à respecter cette norme fondamentale au nom de l'intérêt général, même lorsque cela implique de renoncer à un profit personnel. Il est donc nécessaire que les acteurs économiques possèdent une motivation morale.

Cette norme morale fondamentale, que nous expliquerons ci-après, se formule ainsi :

La norme morale	« Les acteurs du marché ont l'obligation morale de
fondamentale du marché	ne pas profiter des défaillances du marché ».

Comme nous le verrons, si les acteurs économiques ne sont pas prêts à respecter cette norme fondamentale du marché, alors il n'est plus le cas que l'institution du marché participe à l'amélioration générale des conditions matérielles de la société. Autrement dit, si les acteurs économiques n'acceptent pas de renoncer à certains de leurs intérêts, ou à la maximisation de certains de leurs intérêts, lorsque la satisfaction de ces intérêts contredit

les « règles du jeu » du marché, alors le marché ne permet pas d'amener une société à la prospérité matérielle.

Nous concluons ainsi que l'institution du marché ne peut réaliser le bien commun que sous condition que les acteurs disposent d'une *motivation morale*, i.e. d'une motivation à renoncer à leurs intérêts privés lorsque leur satisfaction est contraire à l'intérêt commun. En ce sens, même le marché ne peut pas fonctionner en faisant « l'économie de la vertu ».

3.4.1. LES DÉFAILLANCES DU MARCHÉ

Les actions égoïstes dans le cadre institutionnel d'un marché ne mènent à la prospérité économique collective que si certaines conditions idéales sont satisfaites. Depuis longtemps déjà, les économistes néo-classiques ont établi lesdites conditions ; pour que le marché guide les intérêts particuliers en direction de la prospérité générale – notamment en diminuant les prix à la consommation – il faudrait que les conditions idéales suivantes soient satisfaites :

- 1) Concurrence parfaite¹¹⁶
- 2) Absence d'externalités¹¹⁷
- 3) Information complète¹¹⁸
- 4) Coûts de transaction nuls¹¹⁹
- 5) Les biens publics¹²⁰

¹¹⁶ La condition de concurrence parfaite se décompose en plusieurs sous-conditions : (i) atomisticité du marché : il existe une multitude d'acteurs dont aucun ne détient la capacité d'influencer à lui seul les prix, (ii) homogénéité des produits : les produits sur un marché doivent être interchangeables, seul le prix permet la comparaison, (iii) ouverture du marché aux concurrents, (iv) mobilité complète des facteurs de production.

¹¹⁷ Les coûts et les gains d'une transaction sont assumés entièrement par les parties concernées par la transaction. Si une transaction génère des coûts pour une personne externe, il s'agit alors d'une externalité négative – la pollution, qui impose des coûts au reste de la société est un exemple paradigmatique. Si la transaction génère des gains pour une tierce personne, il s'agit alors d'une externalité positive – l'exemple emblématique est celui de l'apiculteur qui profite de l'activité de son voisin arboriculteur sans payer pour le « service » rendu par son voisin. En général, le but de l'action publique est de décourager les externalités négatives par le biais de taxes et d'encourager les externalités positives par le biais de subventions.

¹¹⁸ Les acteurs qui prennent part à la transaction doivent bénéficier d'une information égale et complète.

¹¹⁹ Se renseigner, rédiger un contrat et négocier sont des activités coûteuses. Le marché donne lieu à un résultat efficient uniquement lorsque de tels coûts n'existent pas.

¹²⁰ Les biens publics se définissent par le fait que la consommation par une personne de ce bien ne diminue pas les possibilités de consommation des autres personnes (non-rivalité) et par le fait qu'aucune personne ne peut être exclue de la consommation (non-excluabilité). L'éclairage public ou la défense nationale sont des biens publics paradigmatiques. Le marché ne permet pas de les produire.

Le problème provient du fait que, dans le monde réel, ces conditions ne sont pas toujours réalisées. Par exemple, il n'est pas toujours le cas que les acteurs économiques bénéficient d'une information complète (et égale) sur les biens échangés ; un vendeur de voiture d'occasions dispose d'informations sur les défauts du véhicule dont ne bénéficie pas l'acheteur ; un futur employé connaît ses propres traits de caractères (travailleur ou fainéant, loyal ou déloyal, discipliné ou négligeant), alors que son (potentiel) futur employeur n'a pas accès à cette information ; une entreprise dispose d'informations sur ses produits ou sur les conditions de travail dont ne disposent pas ses acheteurs et ses employés.

Il en va de même pour les autres conditions, à savoir la concurrence, les externalités, les coûts de transaction ou les biens publics ; ces conditions ne sont pas toujours réalisées dans la réalité. Or, à chaque fois que l'une de ces conditions n'est pas satisfaite dans la réalité, il devient possible, pour les acteurs économiques, de tirer avantage de cette défaillance pour faire un profit personnel contraire à l'intérêt général. Ainsi, dans chaque situation où le marché est « défaillant », l'intérêt personnel des acteurs s'oppose à l'intérêt général. Par exemple, en cas d'asymétrie d'information, un producteur peut choisir de mentir au sujet des qualités ou défauts de ses produits pour tirer un bénéfice personnel au détriment du consommateur. Ou encore, lorsque la condition d'atomicité n'est pas pleinement satisfaite, un dirigeant d'entreprise peut s'arranger avec ses « concurrents » sur les prix dans le but d'augmenter son profit personnel, au détriment de l'intérêt des consommateurs.

Du constat que la réalité ne correspond pas toujours aux conditions idéales, dans quel cas l'intérêt personnel des acteurs diverge de l'intérêt général, il devient possible de tirer la conclusion normative suivante : lorsque l'occasion de tirer profit d'une « défaillance du marché » se présente, « les acteurs du marché ont l'obligation *de ne pas* profiter de ces défaillances non corrigées (NÉRON, 2010, p. 12) ». La nécessité que les acteurs économiques respectent cette « norme morale fondamentale » est reconnue depuis longtemps par certains économistes, tels que Kenneth Arrow, développeur de la théorie néo-classique de la théorie de l'équilibre général.

Le succès du marché ne doit pas amener à l'opinion souvent exprimée que l'égoïsme pur est économiquement optimal. Quand il y a des avantages de taille,

de sorte qu'il existe un pouvoir de contrôler le marché, quand les individus peuvent s'infliger des dommages les uns aux autres ou faire le bien des autres d'une manière qui n'est pas pénalisée ou récompensée par le marché, ou, par-dessus tout, lorsque les acteurs économiques disposent d'une information privée, alors il y a une obligation morale d'utiliser ces rôles avantageux pour le bénéfices des autres.¹²¹

Lorsque leurs intérêts privés contredisent l'intérêt général, les acteurs économiques doivent favoriser l'intérêt général. Cela implique, par exemple, de ne pas dissimuler d'informations, même si l'opportunité d'en tirer profit se présente :

[Les] entreprises de production connaissent mieux en général la qualité et la sécurité de leurs produits que les acheteurs. Il y a clairement une obligation de révéler cette vérité, même aux dépens des profits, parce que le marché ne parvient pas de lui-même à présenter les faits lorsque les acheteurs ne sont pas informés [...]. De façon similaire, l'entreprise connaît mieux que ses travailleurs, par l'expérience, les conditions de sécurité de ses sites de production. En conséquence, les conditions qui permettent au marché de fonctionner sont violées, et l'obligation morale doit prendre sa place.¹²²

Quiconque refuse de respecter ces règles agit de façon immorale, d'après les critères normatifs de l'économie néo-classique elle-même. Le marché repose ainsi sur une norme morale fondamentale *interne* à l'institution ; les acteurs économiques ont l'obligation morale « de ne pas tirer profit des défaillances du marché ». Autrement dit, ils ont l'obligation de renoncer à tout « profit facile ».¹²³ Et puisque l'État ne peut pas garantir le

¹²¹ "[T]he successes of the market should not lead to the view so frequently drawn that pure selfishness is economically optimal. When there are advantages to size so that power to control the market exists, when individuals can inflict harm on each other or even do good for each other in ways that are not penalized or rewarded in the market, or, above all, when economic agents have private knowledge, then there are moral obligations to use these advantageous roles for the benefit of others." (ARROW, 1992, p. 22).

¹²² "[A] producing firm generally knows the quality and safety of its product more than most buyers. There is clearly an obligation to reveal this truth, even at the expense of profits, for the market will do very poorly in sorting out the facts then the buyers are uninformed. [...] Similarly, the firm knows, by experience, the safety conditions in its plants more than workers can. Hence, the conditions that the market works well are violated, and moral obligation should take its place." (ARROW, 1992, p. 20).

¹²³ « Sous l'angle éthique, [...] s'adressant directement au comportement individuel, le libéral dira ceci: le bon fonctionnement du marché dépend d'une éthique des entrepreneurs et plus généralement des agents économiques, d'une éthique [...] dans laquelle on ne recherche spontanément que les avantages licites de l'activité économique. [Ainsi] l'unité motivationnelle se rompt chez l'agent économique. Tantôt c'est l'optimisation de ses rétributions dans l'interaction marchande qui prévaut, tantôt c'est un autre motif tiré

respect de cette norme par le biais des lois et incitations¹²⁴, le marché dépend du fait que les acteurs économiques disposent d'une motivation spécifiquement morale. En l'absence d'une telle motivation *morale* – i.e. d'un souci désintéressé du bien commun ou général – le marché ne permet pas de réaliser la prospérité économique générale.

En résumé, si les conditions d'un marché idéal étaient réalisées dans la réalité, nous n'aurions pas besoin de faire appel aux motivations morales des acteurs économiques. Or, les conditions idéales du marché ne sont pas satisfaites dans la réalité. Il faut donc que les acteurs économiques soient capables de renoncer *intentionnellement* à leur intérêt privé au nom de l'intérêt public. En d'autres termes, le marché ne peut pas fonctionner en « faisant l'économie de la vertu ». Le bien général ne peut pas être réalisé comme la conséquence non intentionnelle de la poursuite des intérêts privés.

3.4.2. L'IMMORALITÉ DE CERTAINES PRATIQUES DE LOBBYING DES ENTREPRISES

Dans la section précédente, nous avons vu que le marché ne permet de réaliser un état bénéfique pour l'ensemble de la société que si les acteurs économiques acceptent de renoncer à leur intérêt personnel lorsque cet intérêt s'oppose à l'intérêt général, c'est-à-dire lorsque cet intérêt personnel contredit la norme fondamentale du marché qui consiste à « ne pas tirer avantage des défaillances du marché ». Comme le montre Joseph Heath, cette norme fondamentale du marché peut être décomposée en un ensemble de règles spécifiques permettant de guider et d'évaluer les décisions des individus sur le marché ou dans la sphère politique (HEATH, 2004, p. 84):

- 1) Minimiser les externalités négatives
- 2) Ne faire concurrence que par le biais du prix et de la qualité
- 3) Réduire les asymétries d'information entre les entreprises et les consommateurs

d'une moralité traditionnelle qui prescrit le renoncement au "profit facile" du cartel, des opérations d'initiés, etc. » (SCHULTHESS, 1998, p. 255).

¹²⁴Contrairement à ce que semble supposer par exemple Jean Tirole, selon lequel il est uniquement du rôle de l'État de supprimer les défaillances du marché : « [C]omme on entend le montrer dans ce livre, bénéficier des vertus du marché requiert souvent de s'écarter du laisser-faire. De fait, les économistes ont consacré une grande partie de leurs recherches à l'identification des défaillances du marché et à leur correction par la politique publique : droit à la concurrence, régulation par les autorités sectorielles et prudentielles, taxation des externalités environnementales ou de congestion, politique monétaire et stabilisation financière, mécanismes de fournitures des biens tutélaires comme l'éducation et la santé, redistribution, etc. » (TIROLE, 2016, p. 56)

- 4) Ne pas exploiter la diffusion de la propriété – i.e. le fait que les actionnaires sont trop nombreux pour pouvoir individuellement exercer un contrôle sur la direction
- 5) Eviter de placer des barrières d'entrée
- 6) Ne pas utiliser les subventions croisées pour éliminer les concurrents
- 7) Considérer le prix comme étant fixé de manière exogène

Ces règles permettent d'évaluer et de guider le comportement des acteurs économiques et des acteurs politiques. Toutefois, l'évaluation du comportement des entreprises en fonction de la norme fondamentale du marché ne se limite pas à la seule sphère économique. L'existence d'une norme morale inhérente au marché concurrentiel a des répercussions non seulement sur l'évaluation du comportement des entreprises à l'intérieur du marché, mais aussi sur l'évaluation de leur comportement dans la sphère politique. En effet, dans ce contexte, les entreprises exercent une forte influence par le biais des activités de lobbying, qui leur permettent d'obtenir des avantages à moindre frais.

L'action politique [des entreprises] peut s'avérer autrement plus rentable. En investissant dans l'action politique, les producteurs peuvent obtenir une protection contre la concurrence et augmenter leur revenu substantiellement, sans devoir consentir des efforts soutenus pour être concurrentiels, sans risquer des fonds dans la modernisation de leur équipement ou dans l'amélioration de leur produit (RAYNAULD & STRINGER, 2003, p. 65).

En investissant dans l'activité de lobbying politique, les entreprises ont l'occasion d'obtenir une réglementation qui leur est avantageuse, au détriment de l'intérêt collectif. En effet, les lobbies économistes s'affairent souvent à *empêcher* les interventions de l'État qui visent à supprimer une défaillance du marché. Pire encore, les lobbies font également pression sur l'État pour établir des lois qui *créent* des défaillances de marché – jusqu'alors inexistantes. Or, nous avons vu que *du point de vue économique* lui-même, de telles pratiques sont moralement contestables.

[II] faut [...] noter que par leur implication politique, les firmes peuvent aussi influencer les tentatives de corriger ces défaillances des marchés, et traditionnellement, le monde des affaires n'hésite pas à s'opposer à des régulations visant à le faire. Ce qui est moralement troublant avec l'activité des firmes sur la scène politique est donc la façon dont elles peuvent profiter de

défaillances des marchés *en créant politiquement* celles-ci ou en bloquant les tentatives par le gouvernement de les corriger. La moralité implicite du marché nous renseigne donc sur la conduite des firmes sur la scène politique. Elle suggère que les acteurs du marché ont l'obligation de *ne pas* créer des défaillances des marchés par leurs activités politiques (ou de ne pas s'opposer aux régulations qui ont pour objectifs de corriger de telles défaillances) (NÉRON, 2010, p. 13).

Concrètement, par exemple, les entreprises qui font du lobbying pour empêcher que soient introduites des taxes sur les énergies fossiles, taxes dont le but est que les parties de la transaction internalisent les coûts imposés à la société (externalités), violent la norme fondamentale du marché. Elles violent une norme morale *interne* du marché.¹²⁵

Nous pouvons ainsi constater que le marché ne représente pas une institution « dépourvue de morale » où tous les coups seraient permis dans le but de satisfaire l'intérêt personnel. Au contraire, les acteurs économiques sont soumis à un ensemble de règles plutôt contraignantes « internes » au marché. Ces règles concernent aussi bien l'activité économique que l'activité politique de l'entreprise et portent sur un ensemble plus large d'exigences que les seules exigences légales.

3.4.3. UNE ÉTHIQUE DES AFFAIRES CONTRAIGNANTE ET CONFORME À LA CRITIQUE MORALE ORDINAIRE DU MARCHÉ

Nous pourrions croire que l'approche de l'éthique d'entreprise basée sur les « défaillances du marché », ou « éthique des défaillances du marché », que nous avons présentée ici, est (trop) peu exigeante. En effet, l'approche développée ne semble pas permettre de formuler une critique radicale du capitalisme actuel, alors même que nous observons au quotidien ses effets dévastateurs, alors même que nous voyons que le monde est affligé par la cupidité sans borne des dirigeants et propriétaires d'entreprises. En ce sens, la critique du marché développée dans ces lignes paraît bien trop conciliante.

Cependant, ce reproche d'un manque de « radicalité » n'est pas approprié. En effet, le monde dans lequel nous vivons ne correspond pas du tout au monde moral idéal du point de vue économique. D'un côté, dans leurs actions, les acteurs économiques, et notamment les

¹²⁵ Dans la même direction, nous pouvons arguer, avec Wim Dubbink (2015), que les entreprises ont le devoir moral de promouvoir les normes conformes à la norme fondamentale du marché. Autrement dit, elles ont le devoir moral d'œuvrer en faveur de la mise en place d'un cadre législatif moralement acceptable.

entreprises, ne respectent pas la « norme fondamentale du marché » ; de l'autre côté, et cela est plus grave encore, les lois elles-mêmes ne sont pas conformes à la norme morale fondamentale du marché. Comme le soutient Joseph Heath, le monde économique et légal devrait être radicalement transformé pour réaliser le monde économique idéal développé par l' « éthique des défaillances du marché ».

Je pense qu'il est clair [...] que, plutôt que d'être moralement laxiste, l'approche par les défaillances du marché est plutôt restrictive. En fait, dans le monde réel, toute firme qui commencerait de manière unilatérale à respecter ces contraintes [du marché], serait rapidement éliminée du marché. Par exemple, l'exigence que les firmes rivalisent uniquement par le prix et la qualité exclut l'usage de publicité non informative comme procédé pour obtenir des parts du marché. La publicité, comme forme non productive de compétition, impose des pertes significatives à l'économie.¹²⁶

Loin d'être une approche « laxiste » de l'éthique d'entreprise, le respect complet de la « norme fondamentale du marché » entraînerait une transformation radicale des pratiques commerciales que nous connaissons aujourd'hui.

Nous pourrions croire aussi que l'évaluation du comportement des entreprises selon le critère des défaillances du marché est une forme d'évaluation très éloignée des critiques morales traditionnelles ou communes du marché. En ce sens, l'éthique d'entreprise développée dans ces lignes ne permettrait pas de rendre compte ou de saisir les contours du malaise que de nombreuses personnes éprouvent à l'égard du marché. Néanmoins, selon les partisans de l'éthique des défaillances du marché, la critique est infondée. Selon eux, la majorité – voire la totalité – des critiques du marché concerne en fait – de manière non explicite – des situations de défaillances du marché. Nous trouvons cette idée par exemple chez le philosophe Joseph Heath :

[II] n'y a pas de raison de penser qu'une éthique des affaires focalisée sur la maximisation du profit ne puisse pas rendre compte des obligations qui ont été

¹²⁶ "I think it is clear [...] that, rather than being morally lax, the market failure approach is actually quite restrictive. In fact, in the real world, any firm that began to unilaterally respect these [market] constraints would be quickly eliminated from the marketplace. For instance, the requirement that firms compete only through price and quality excludes the use of non-informative advertising as a way of building market share. Advertising, as a form of non-productive competition, imposes significant deadweight losses on the economy." (HEATH, 2004, p. 84)

traditionnellement décrites sous l'appellation de « responsabilité sociale ». Ce qui irrite si souvent les individus au sujet des entreprises – et qui fait que la poursuite du profit a une mauvaise réputation – est l'exploitation d'une ou l'autre forme d'imperfection du marché. Les individus n'ont en général aucun problème avec les entreprises qui font de l'argent en fournissant de bons services, des biens de qualité, qui pratiquent des prix bas, etc. A mon avis, si toutes les entreprises internalisaient complètement l'ensemble des coûts, et facturaient à leurs clients le prix total que la production du bien a imposé à la société, il deviendrait impossible de défendre une conception plus large de la "responsabilité sociale" de l'entreprise, au sujet, par exemple, de l'environnement.¹²⁷

Certains économistes également considèrent que la majorité des critiques adressées à l'institution du marché sont, en réalité, des critiques basées sur des situations de défaillances du marché. Par exemple, Jean Tirole considère que le refus prétendument moral de monnayer l'amitié est un cas d'asymétrie d'information, le refus moral de vendre des enfants, de la prostitution ou du lancer de nains sont des cas d'externalités négatives et la régulation de la consommation de drogue des cas d'internalités (TIROLE, 2016, pp. 59-72). Ainsi, selon lui, la majorité des critiques qui demandent de réguler ou d'interdire certains marchés vise en fait à résoudre des cas de défaillance du marché. Ces critiques ne font donc pas appel à des critères moraux *externes* à ceux du marché.

[La] régulation ou l'interdiction d'un marché peut être une réponse à un problème d'information [...], à une externalité [...], ou encore à une internalité (qui renvoie au comportement d'un individu qui va contre son propre intérêt). Dans les trois cas, la régulation ou l'interdiction du marché répond à une simple défaillance du marché. En l'occurrence, invoquer l'éthique apporte peu à l'analyse ou à la conclusion [...] (TIROLE, 2016, pp. 58-59).

¹²⁷ "[T]here is no reason to think that a business ethics focused on profit-maximization cannot deal with the obligations that have traditionally described under the heading of "social responsibility". What so often annoys people about corporations – and what gives profit-seeking a bad name – is the exploitation of one or another form of market imperfection. People generally have no problem with companies that make money by providing good service, quality goods, low prices, and so forth. In my opinion, if all companies fully internalized all costs, and charged consumers the full price that production of their good imposed upon society, it would be impossible to make the case for any further "social responsibility" with respect to, for example, the environment." (HEATH, 2004, pp. 82-83)

Toute autre critique morale relève, selon Tirole, des préférences individuelles et ne mérite pas considération.¹²⁸

Autrement dit, malgré les apparences, les critiques morales du marché n'emploient pas un critère moral différent de celui employé par les économistes pour évaluer les pratiques commerciales. En ce sens, l'« éthique des défaillances du marché » exprimerait parfaitement le malaise ressenti par la critique morale ordinaire de cette institution.¹²⁹

3.4.4. LE CARACTÈRE INADAPTÉ DE LA CRITIQUE MORALE ORDINAIRE DES DIRIGEANTS D'ENTREPRISES

Dans la présente sous-section, nous allons défendre l'idée que la manière ordinaire ou « classique » d'aborder l'éthique d'entreprise, à savoir la doctrine des « parties prenantes », est inadaptée et qu'il faut lui préférer l'« éthique des défaillances du marché » que nous avons présentée ci-dessus.

La critique classique à l'encontre du marché consiste dans le fait que cette institution repose sur la concurrence entre les individus, sur la capacité d'une personne à évincer ses concurrents pour obtenir un maximum d'avantages personnels – au détriment de ses concurrents, cela va de soi. Or, la critique continue, une telle institution encourage le développement de dispositions de caractère contraires aux dispositions valorisées par la morale ordinaire, telles que la bienveillance, le partage, l'entraide, la coopération, etc.

Dans le cas du marché, [un élément] qui pose problème est que le fonctionnement de l'institution semble *cautionner* ou *excuser*, au nom des résultats économiques d'efficience et de prospérité, toute une gamme de comportements étant autrement considérés moralement problématiques. Elle semble en effet accorder une permission morale aux comportements égoïstes visant la promotion des intérêts uniquement privés (NÉRON, 2016, p. 195).

De ce fait, puisque le marché encourage un esprit de concurrence fondamentalement opposé au développement des dispositions vertueuses de bienveillance et de coopération, le

¹²⁸ « En l'absence d'une identification claire des externalités, l'affirmation des préférences individuelles des uns peut rapidement prendre le pas sur la liberté des autres. » (TIROLE, 2016, p. 72)

¹²⁹ Il nous semblerait possible de critiquer Jean Tirole de la manière suivante : il fait un usage extrêmement large de la notion d'externalité. Cet usage nous semble dépasser l'interprétation purement économique du terme et nous amène donc à sortir du cadre d'analyse purement économique.

marché est considéré comme une institution *dangereuse* pour la société.¹³⁰ Afin de contrecarrer l'effet néfaste de l'esprit de concurrence, il convient alors, selon la critique, d'inculquer aux dirigeants d'entreprise les valeurs de la morale ordinaire, de les sensibiliser à la coopération, au souci de l'autre et au partage, de les éloigner de la seule préoccupation du profit pour développer leur sensibilité morale. Selon cette critique, basée sur les valeurs morales « ordinaires », le rôle de l'éthique et des philosophes moraux consiste à aider les dirigeants d'entreprise à prendre de « bonnes décisions », non seulement profitables, mais aussi soucieuses des intérêts des autres personnes affectées. Le but est de créer des « dirigeants éthiques » (NÉRON, 2010, p. 16).

Cette critique du marché a donné naissance à l'approche de l'éthique des affaires la plus connue, à savoir la doctrine « des parties prenantes » (*stakeholders theory*). Le but de cette approche est d'amener les dirigeants d'entreprise à élargir leur perspective, afin de veiller non seulement à satisfaire les intérêts des actionnaires (*shareholders*) mais de toutes les parties affectées (employés, consommateurs, fournisseurs, communautés locales, gouvernement) (NÉRON, 2010, p. 5). Le manager devrait ainsi considérer l'entreprise comme jouant un rôle non seulement économique, mais aussi social – on parle alors de responsabilité sociale des entreprises (RSE). Il s'agit, dans cette approche d'évaluer et de guider l'action des dirigeants d'entreprise à la lumière des critères moraux utilisés dans la pratique morale ordinaire.¹³¹

Le défaut de cette approche est triple : en premier lieu, l'approche des parties prenantes suppose que l'institution du marché est elle-même complètement dépourvue de toute exigence morale. La morale est traitée comme un ensemble de règles *externes* à l'activité économique, qu'il s'agit d'une manière ou d'une autre d'introduire dans les prises de décisions des dirigeants d'entreprise. Ce faisant, la critique ignore complètement le fait

¹³⁰ Il nous semble utile de mentionner que cette caractérisation des pratiques commerciales comme étant *opposées* et *dangereuses* pour les pratiques morales revient régulièrement dans le discours « ordinaire » sur l'immoralité du marché. A titre d'exemple : « Le fond de l'affaire est que la compétition à laquelle pousse le marché encourage l'affirmation de soi, ce qui nécessite quasiment de la malhonnêteté, de l'inégalité dans la répartition des richesses, une consommation excessive, le gaspillage et l'inefficacité » (HINDE, 2013, p. 40). L'extrait est traduit du livre de l'éthologue et psychologue Robert Hinde (HINDE, 2011).

¹³¹ Cette manière de représenter l'éthique des « parties prenantes » est celle explicitement acceptée par ses défenseurs : « Les exigences des consommateurs, des investisseurs, des salariés et plus globalement des citoyens font apparaître une demande sociale croissante en faveur d'une plus grande intégration de l'éthique dans la vie des entreprises » (MERCIER, 2014, p. 3). Ces exigences prennent une forme institutionnelles par le biais d'un « ensemble de principes et règles qui dirigent et limitent les actions des dirigeants. Le questionnement se focalise sur le rôle du dirigeant [...] » (MERCIER, 2014, p. 23).

que le marché n'est pas une institution dépouillée de toute morale, mais, au contraire, que cette institution repose sur une norme morale propre.

En deuxième lieu, cette conception de l'éthique des affaires attribue au dirigeant d'entreprise une responsabilité morale démesurée. En effet, la réussite ou l'échec moral de l'entreprise dans sa globalité semble dépendre uniquement de la moralité de son dirigeant.

[Le modèle des parties prenantes] [...] met une emphase exagérée sur le caractère moral et vertueux de certains agents, dans ce cas-ci les hauts dirigeants d'entreprises. Le présupposé de base d'un tel modèle est que si les hauts dirigeants d'une organisation savent cultiver certaines vertus et développer leur caractère moral, c'est l'organisation dans son ensemble qui se conformera à ces plus hautes attentes morales. En somme, le dirigeant éthique conduit à l'entreprise éthique (NÉRON, 2010, p. 6).

Cette exagération du rôle éthique du dirigeant amène en outre à réduire l'entreprise à son équipe de direction, ce qui est également problématique (NÉRON, 2010, p. 7).¹³²

En troisième lieu, la critique se montre aveugle au fait que les traits de caractère valorisés dans la morale ordinaire sont incompatibles avec les traits de caractères utiles ou nécessaires à l'activité commerciale dans le cadre institutionnel du marché. En effet, les dispositions qui sont valorisées dans les relations interpersonnelles ordinaires ne sont pas des dispositions qui sont adaptées à tous les contextes institutionnels. Elles ne sont notamment pas adaptées au contexte institutionnel du marché.

[P]lusieurs des vertus que nous célébrons dans les relations interpersonnelles peuvent avoir des effets pervers en contexte institutionnel. La loyauté envers les amis et la famille est parfaitement louable dans nos relations individuelles, mais elle est une source importante de corruption et de favoritisme dans un contexte institutionnel. De même, gentillesse, douceur et générosité peuvent être parfaitement louables dans nos relations quotidiennes, mais parfois quelque peu inappropriées dans certaines relations institutionnelles, comme lorsqu'un dirigeant d'entreprise doit congédier des employés ou lorsqu'un policier doit

¹³² L'attribution au dirigeant d'entreprise du rôle de « leadership éthique » (*ethical leadership*) est récurrente dans la littérature sur l'éthique d'entreprise. On retrouve de nombreux articles dans le *Journal of Business Ethics* sur cette thématique.

appréhender un suspect. En somme, la vertu interpersonnelle ne peut pas être simplement « transposée » en contexte institutionnel sans une prise en compte attentive des exigences fonctionnelles de l'institution (NÉRON, 2010, p. 7).

Dans l'activité économique, comme pour d'autres activités, la concurrence et l'esprit compétitif – par opposition à la coopération et l'entraide – sont encouragés en vertu du bienfait qu'ils procurent à des personnes externes à la compétition.

[La] compétition est justifiée lorsqu'elle tend à produire des bénéfices pour des acteurs « externes » à cette compétition, à tout le moins des acteurs qui ne sont pas *directement* impliqués dans la compétition [...] (NÉRON, 2016, p. 202).

Dans le marché, la concurrence entre les entreprises – dans les limites imposées par la norme fondamentale du marché – apporte des bénéfices aux consommateurs, qui voient notamment les prix diminuer. Dans d'autres contextes, notamment dans le sport, la concurrence et l'esprit de compétition sont aussi justifiés par les bienfaits qu'ils procurent à des parties externes; dans la compétition sportive, les parties externes sont les spectateurs et le bienfait est une forme de plaisir esthétique.

Dans le cas d'un sport d'équipe comme le football, chacune des équipes tente de remporter la victoire en utilisant diverses stratégies, mais l'objectif de l'institution est de procurer un certain plaisir esthétique aux spectateurs qui, eux, sont indirectement impliqués dans le jeu et donc externes à l'activité compétitive (NÉRON, 2016, pp. 201-202).

Que deviendrait le plaisir des spectateurs de compétitions sportives si les concurrents commençaient à se soucier du sort de leur adversaire ? Que penserait-on d'un joueur de foot qui, attristé par le retard de points de l'équipe adverse, décide de l'aider à marquer un but, par exemple en marquant un auto-goal ? Que penserait-on d'un coureur qui ralentirait pour attendre son concurrent et lui céder la priorité sur la ligne d'arrivée, dans le but de lui faire plaisir ? Nous penserions certainement que cette personne agit d'une manière bizarre, qu'elle n'a pas compris le but de la compétition. Plus encore, une telle personne priverait les spectateurs de la joie et du plaisir de la compétition sportive ; elle priverait les tierces personnes du bienfait que la compétition a précisément pour objectif de réaliser. C'est

pourquoi les actions qui visent à battre son concurrent, dans les institutions compétitives, ne sont pas uniquement *facultatives*, mais elles sont *obligatoires*.

De la même manière, un entrepreneur qui, par pure bienveillance à l'égard de son concurrent, qu'il ne souhaite pas mettre en difficulté, décide de ne pas abaisser les prix et de coopérer avec lui pour maintenir des prix élevés, priverait les consommateurs des bienfaits que la concurrence est précisément censées procurer. La bienveillance, la coopération, le partage, etc., ne sont pas appropriés dans des contextes où la compétition entre des individus est avantageuses pour des personnes externes à la compétition.

Bien entendu, tout comme la compétition sportive, l'esprit de compétition n'est encouragé et légitime qu'à l'intérieur des limites fixées par les « règles du jeu ». Un coureur a le droit de courir plus vite que son concurrent et de passer la ligne d'arrivée en premier, mais cela ne l'autorise cependant pas à bousculer son concurrent, à l'empoisonner ou à le payer pour remporter la victoire. De la même manière, si la concurrence est bénéfique sur le marché, elle doit néanmoins être contenue à l'intérieur des « règles du jeu ». Ainsi, l'esprit compétitif n'est pas en soi condamnable, bien qu'il ne corresponde pas aux traits de caractères valorisés dans la morale commune. Toutefois, l'esprit compétitif ne doit pas amener à outrepasser les règles de la compétition équitable.

En conclusion, nous voyons que, sur le marché, mais aussi dans d'autres contextes¹³³, la compétition est à la fois *restreinte* mais néanmoins *obligatoire*. Le marché, comme le sport, n'amène des bénéfices sociaux que si la compétition y est certes modérée, ou limitée, mais il faut néanmoins qu'elle soit bien présente.

[La] canalisation de comportements non coopératifs peut évidemment avoir des effets pervers. [...] Elle peut notamment mener à une compétition qui est trop agressive et conduire à l'adoption de toutes sortes de tactiques déloyales ; espionnage industriel, publicité mensongère ou négative, etc. Mais la compétition entre firmes n'est pas moralement problématique que lorsqu'elle est « sauvage », mais aussi lorsqu'elle est insuffisante, lorsqu'elle prend des allures de collusions et de cartels (NÉRON, 2016, p. 204).

¹³³ Par exemple la « représentation électorale, les régimes parlementaires et divers systèmes politiques. » (NÉRON, 2016, p. 206).

De l'analyse qui précède, nous tirons la conclusion suivante : l'approche adaptée pour évaluer les comportements des dirigeants d'entreprise sur le marché est celle de la morale « interne » du marché plutôt que celle de la morale ordinaire « externe » (bienveillance, coopération, ...) au marché, employée notamment dans l'éthique des « parties prenantes ». Nous suivons ainsi, encore une fois, Pierre-Yves Néron, qui lui-même se réfère aux idées de John Boatright (BOATRIGHT, 1999).¹³⁴

L'objectif de nos réflexions philosophiques sur le commerce ne devrait pas être d'imaginer la figure du « dirigeant éthique », mais plutôt de penser les modalités de la construction de « marchés éthiques » (NÉRON, 2010, p. 16).

Comme nous l'avons vu, l'approche de l'éthique des affaires qui se focalise plutôt sur les critères moraux *internes* à l'activité économique – i.e. l'éthique des « défaillances du marché », dont nous avons vu que la portée critique est déjà très large – a plus de chances (i) de ne pas soumettre le comportement du dirigeant d'entreprise à des exigences paradoxales, (ii) de ne pas exagérer le rôle moral du dirigeant et (iii) de rendre explicite la morale implicite du marché.

3.4.5. LES LIMITES DE L'ÉTHIQUE DES DÉFAILLANCES DU MARCHÉ

Nous souhaitons encore apporter quelques précisions ; il ne faut pas considérer l'analyse précédente comme offrant une *caution morale*¹³⁵ sans limite aux acteurs du marché. S'il est vrai que les critères de la morale ordinaire ne s'appliquent que difficilement à l'évaluation des comportements des dirigeants d'entreprise, cela ne signifie pas que, pour ces dirigeants, tout soit permis au nom de la concurrence. En effet, la *caution morale* n'est accordée aux acteurs économiques que de façon *limitée* et *conditionnelle* (NÉRON, 2010, p. 11). Comme nous l'avons vu, elle est *limitée* dans le sens que la recherche du profit et la

¹³⁴ Cette approche se distingue, en outre, des approches visant à démontrer que les vertus morales sont *bénéfiques*, voire *nécessaire* à la réussite économique. Chez les économistes, cette approche est soutenue par exemple, sous des formes différentes, par (MCCLOSKEY, 2006; BRUNI & SUGDEN, 2013). Le philosophe Julian Nida-Rümelin soutient que la réussite commerciale n'est possible que si l'on possède les vertus telles que la fiabilité, la force de volonté, la capacité de jugement, le respect, la loyauté et l'attention (NIDA-RÜMELIN, 2011). Dans les textes non académiques, de telles défenses des vertus, en tant qu'elles sont propices à l'activité économique, se trouvent par exemple dans un volume du magazine économique suisse *Swissquote* de juillet 2012 intitulé « Les Bénéfices de l'éthique » dont l'objectif est de montrer que « l'éthique paie » (BÜRKI, 2012).

¹³⁵ Pour rappel, la caution morale, ou immunité morale, consiste à accorder à des personnes, dans un contexte institutionnel, le droit d'accomplir des actions qui, hors de ce contexte institutionnel, seraient considérées comme moralement inacceptables (NÉRON, 2010, p. 11).

compétition cessent d'être légitimes lorsqu'elles requièrent de violer la norme fondamentale du marché. En outre, cette caution morale est *conditionnelle*, dans le sens qu'elle dépend de la réalisation d'un objectif collectif, à savoir l'efficacité économique. Si l'objectif est abandonné, alors la caution morale est abandonnée également ; la compétition n'est alors plus acceptable. Or, au regard de cette condition, nous allons voir que le critère d'efficacité économique peut entrer en conflit avec d'autres objectifs sociaux. Autrement dit, la recherche de l'efficacité économique ne représente pas le seul objectif social légitime et cette recherche empêche même parfois de réaliser d'autres objectifs sociaux légitimes.

Quelles sont les différentes façons dont l'objectif social que le marché doit réaliser, à savoir l'efficacité économique, peut entrer en conflit avec d'autres objectifs sociaux ? Le conflit le plus connu est celui de la *distribution équitable* des richesses ; en effet, la réalisation de l'objectif d'efficacité économique est compatible avec une distribution très inégalitaire des richesses, elle-même contraire aux objectifs de justice sociale. Néanmoins, bien que le conflit entre efficacité économique et justice distributive représente le conflit le plus connu entre différents objectifs sociaux, il ne s'agit pas du seul conflit possible.¹³⁶

Pour comprendre l'apparition de conflits entre différents objectifs sociaux, il convient de décrire la condition de l'être humain à l'état de nature, c'est-à-dire en l'absence d'institutions sociales assignant des droits et devoirs à chacun.¹³⁷ A l'état de nature, les êtres humains sont confrontés à un ensemble de problèmes collectifs, c'est-à-dire de situations où la poursuite des intérêts privés mène à une situation désavantageuse pour tout le monde ; pour résoudre ces problèmes, il leur est nécessaire d'établir un ensemble de règles contraignantes que tous les individus ont l'obligation de suivre. Les institutions sociales ne sont alors rien de plus que cet ensemble de règles visant à résoudre des problèmes d'action collective.

[A]fin de parvenir à un meilleur résultat [...] les individus ont raison d'accepter une certaine forme de contraintes sur leur comportement. Une institution sociale peut

¹³⁶ Joseph Heath va même plus loin : selon lui, en se focalisant exclusivement sur le problème de la justice distributive, la critique du marché a eu pour conséquence de camoufler le fait que l'État a pour fonction de réaliser d'autres objectifs sociaux que la seule justice distributive, et que ces objectifs aussi sont parfois contraires à l'efficacité économique.

¹³⁷ Pour cette partie, nous nous baserons largement sur l'analyse de Joseph Heath dans son texte « Institutions et bénéfices coopératifs » (HEATH, 2012).

alors être définie comme un ensemble de normes codifiant ces contraintes (HEATH, 2012, p. 15).

Le marché est l'une de ces institutions ; le respect des règles du marché permet de réaliser des gains collectifs d'efficience. Cependant, cet objectif d'efficience peut entrer en conflit avec d'autres objectifs collectifs poursuivis par d'autres institutions.

Nous pouvons distinguer cinq types de gains coopératifs, c'est-à-dire de gains qui ne peuvent être obtenus que par la coopération : (1) les économies d'échelle, (2) les gains de complémentarité ou d'efficience, (3) les gains de gestion des risques, (4) les gains de l'engagement préalable, (5) les gains de diminution des coûts d'apprentissage de l'information. Pour réaliser chacun de ces gains spécifiques, des institutions sociales différentes doivent être mises en place.

(1) *Les économies d'échelle*

Les économies d'échelle caractérisent des situations dans lesquelles « l'addition d'un individu à la réalisation d'une tâche engendre un bénéfice plus grand que la contribution de la première personne » (HEATH, 2012, p. 19). En d'autres termes, il est parfois avantageux collectivement d'effectuer une tâche à plusieurs. Les entreprises, dans le secteur privé, et l'octroi d'un monopole, dans le secteur public, sont des solutions institutionnelles dont le but est précisément de garantir l'augmentation de la productivité par les économies d'échelles.

(2) *Les gains issus de la complémentarité des besoins et des compétences (ou gains d'efficience)*

Le fait que les êtres humains possèdent des besoins et compétences complémentaires offre également des occasions de gains collectifs (HEATH, 2012, p. 20). En effet, la complémentarité des besoins a pour conséquences que deux personnes ayant des besoins différents peuvent procéder à un échange et améliorer de cette manière leur situation respective. De son côté, la complémentarité des compétences permet que les personnes les plus habiles pour produire un bien soient désignées pour le produire pour tout le monde. Dans le secteur privé, c'est l'institution du marché – par la propriété privée et la division du travail – qui a pour fonction précisément de permettre d'obtenir ces gains liés à la complémentarité des besoins et compétences. Dans le secteur public, c'est l'État, en

garantissant l'accès aux biens publics, qui a pour rôle de permettre de réaliser certains gains de complémentarité.

(3) *Les gains collectifs de gestion des risques*

La condition humaine est telle que nous ne parvenons pas « à déterminer avec précision ce que l'avenir nous réserve » (HEATH, 2012, p. 22). De ce fait, nous vivons avec un sentiment subjectif d'incertitude et d'insécurité. Or, la gestion collective des risques permet « aux individus de réduire la subjectivité de cette incertitude » (HEATH, 2012, p. 22). En s'associant pour se prémunir contre certains dangers probables de la vie humaine, les associés acquièrent la maîtrise de l'incertitude et de l'insécurité. Dans le secteur privé, les assurances ont pour fonction de maîtriser les risques associés à la condition naturelle de l'homme. Pour le secteur public, l'institution de la sécurité sociale par l'État a pour fonction de diminuer l'insécurité créée par le système économique et social.

(4) *Les gains collectifs de l'engagement préalable*

Ces gains découlent d'une autre caractéristique de la nature humaine : les préférences des êtres humains présentent – de manière parfaitement prévisible – certaines formes d'instabilité. En effet, les préférences humaines subissent dans le temps des « renversements » tout à fait prévisibles et nuisibles à l'individu sur le long-terme. Or, il est possible, individuellement et, surtout, collectivement, de se prémunir contre ces renversements de préférences. La solution consiste à formuler des *engagements préalables*. Par exemple, nous faisons appel à un coach personnel pour nous assurer de tenir nos engagements sportifs ou nous demandons une tutelle pour gérer nos finances. Sur le plan collectif, le financement de la retraite par une ponction obligatoire sur les salaires constituent une manière de protéger les intérêts des individus contre l'instabilité temporelle de leurs préférences.

(5) *Les gains de la diminution des coûts d'apprentissage de l'information* représentent le dernier avantage de la coopération.

Ces gains consistent dans le fait que le langage, et la communication en général, permettent d'acquérir et de transmettre des informations à un coût réduit. En effet, l'acquisition de certaines informations par un individu isolé serait bien souvent impossible ou extrêmement coûteuse. Grâce au langage et à la communication, l'accès à certaines informations pour un individu isolé devient possible et nettement moins coûteuse. Bien

entendu, une communication avantageuse pour tous requiert de respecter un ensemble de contraintes, notamment celle de *dire la vérité* (HEATH, 2012, p. 28).¹³⁸ Les institutions sociales qui ont pour fonction d'obtenir et de transmettre l'information sont les médias et l'édition dans le secteur privé et la recherche académique dans le secteur public.

Or, il est problématique que les gains que permettent d'obtenir ces différentes institutions sont parfois exclusifs ; l'obtention d'un gain coopératif d'un certain type exclut l'obtention d'un gain coopératif d'un autre type. Par exemple, un gain d'efficience peut empêcher un gain de gestion du risque ou empêcher un gain de complémentarité des besoins et compétences. Comme le soutient Joseph Heath :

A cause des incitatifs qu'elles créent, les institutions dont nous avons besoin pour surmonter un type de problème d'action collective empêchent souvent la résolution d'un autre problème – dans certains cas, elle peut même l'exacerber. (HEATH, 2012, p. 38)

Par exemple, la recherche des gains de complémentarité ou d'efficience – gains de type (2) – par la mise en place de l'institution du marché concurrentiel peut entrer en conflit avec la recherche des gains collectifs de gestion du risque – gains de type (3).

Le système de la propriété privée, qui résout la tragédie des lots communs, a aussi comme effet la « démutualisation » d'un certain type de risque. Si le pâturage [d'un paysan] est touché par une nielle localisée, un système de propriété commune divisera automatiquement la perte également entre tous les paysans. Le lot commun est une forme d'assurance pour les paysans. Lorsqu'il est divisé en propriété privées, il peut en résulter de meilleures pratiques de gestion du territoire, mais chacun se trouve ainsi exposé à un bien plus grand risque. C'est précisément le mécanisme d'« internalisation » des pertes qui, d'un côté, encourage les comportements plus responsables, de l'autre, engendre une bien plus grande instabilité (HEATH, 2012, pp. 40-41).

¹³⁸ Nous pourrions certainement compléter la liste des contraintes sur les intérêts privés qui rendent possibles la communication à l'aide des théories de l'éthique de la communication de Jürgen Habermas (HABERMAS, 1999) et Karl-Otto Apel (1996-1998).

Cette analyse permet d'évaluer l'économie du marché sous un angle nouveau ; certes, si le marché a permis d'obtenir des gains concernant l'augmentation de l'efficacité et de la satisfaction des besoins, il a par contre largement augmenté l'insécurité.

Ce qu'a accompli l'institution de l'économie du marché, selon cette perspective, est la démutualisation massive de la société. Cela a entraîné une augmentation de la production, une meilleure utilisation des ressources matérielles et une meilleure distribution du produit final, mais cela a aussi exposé les travailleurs à des niveaux de risques jamais vus (chômage, invalidité, maladie, etc.), contre lesquels les institutions de la société rurale traditionnelle offraient certaines protections. [...] Il n'est pas clair à quel point les conditions de vie épouvantables des travailleurs, documentées par de nombreux observateurs, étaient dues aux conséquences distributives du marché lui-même, plutôt qu'à l'effet de démutualisation. Il n'est pas non plus clair si l'histoire de la lutte des classes devrait être considérée comme un conflit à propos des parts de la distribution ou plutôt comme un conflit à propos des types de bénéfices coopératifs qui devraient être prioritaires. Le développement de l'État providence peut certainement être vu comme un processus à travers lequel l'État en est venu à récupérer nombre de fonctions assurantielles dont s'acquittaient les institutions rurales précapitalistes, venant ainsi combler le vide qui fut créé par l'avènement du marché (HEATH, 2012, p. 42).

Des conséquences néfastes du marché peuvent également être observées sur les possibilités de bénéficier des gains de l'engagement préalable – gains du type (4). En effet, le marché ayant pour objectif de satisfaire les préférences *actuelles* des consommateurs, il est incompatible avec les stratégies visant à les prémunir contre les renversements de leurs préférences dans le temps. Cet effet néfaste de la priorité accordée à l'efficacité se constate, selon Joseph Heath, surtout dans l'absence ou la fragilité des systèmes de retraite dans les pays hautement partisans du libre marché – comme les États-Unis.

L'épargne est peut-être le secteur où la tension entre les gains de l'échange et les stratégies d'engagement préalable est la plus apparente. Dans une parfaite économie de marché, il n'y aurait aucune imposition d'épargne, étant donné que les préférences des individus face à la consommation présente et face à la consommation futures sont différentes. Un régime d'épargne privées volontaires

procure des avantages grâce à la complémentarité entre ceux qui ont un plus grand intérêt à consommer maintenant et ceux qui ont un plus grand intérêt à consommer plus tard. Cependant, cela crée un problème persistant étant donné la contradiction dynamique des individus en regard de leurs préférences. La plupart des gens ont de la difficulté à épargner et un pourcentage significatif de la population n'épargnera jamais un sou (HEATH, 2012, p. 43).

Ainsi, si le marché permet d'obtenir certains gains coopératifs, il ne permet pas de réaliser *tous* les types de biens coopératifs. Bien au contraire, l'institution du marché *empêche* même, ou met en danger, l'obtention *d'autres types* de gains coopératifs.

De façon générale, l'économie de marché est un excellent facteur de promotion des bénéfices coopératifs du second type (les gains obtenus de l'échange et de la spécialisation), mais elle peut [...] devenir un obstacle lorsqu'il s'agit de fournir des bénéfices coopératifs des quatre autres types (HEATH, 2012, p. 44).

En ce sens, toute défense de l'institution du marché ne peut qu'être limitée ; si le marché permet bien de réaliser des gains coopératifs de complémentarité (ou d'efficience), il ne permet pas, et même empêche, les gains d'économie d'échelle, de gestion du risque, des engagements préalables et de diminution des coûts d'apprentissage de l'information.

L'existence de tels conflits entre les objectifs collectifs poursuivis par différentes institutions sociales n'apparaît pas dans la théorie économique. L'explication est simple ; la théorie économique présuppose *explicitement* que les autres problèmes d'action collective (gestion du risque, transmission de l'information, ...) n'existent tout simplement pas.

[N]ous devons toujours garder à l'esprit à quel point est partielle l'image de la société que nous donnent les modèles standard de l'économie du bien-être. La preuve du premier théorème fondamental de l'économie du bien-être, par exemple, commence en tenant pour acquis qu'il n'y a pas d'économies d'échelles, qu'il y a un marché d'assurances complet, que les agents ont des préférences stables et qu'il y a une symétrie d'information. En d'autres termes, tous les autres bénéfices coopératifs sont explicitement exclus de toute considération (HEATH, 2004, p. 45).

En ignorant complètement qu'il existe d'autres *types* de gains possibles que les seuls gains monétaires, l'analyse économique amène à la conception fautive selon laquelle le seul gain collectif pertinent est l'efficacité économique. De ce fait, l'analyse économique amène à considérer que le seul rôle des institutions sociales – à l'exception du marché – consiste toujours à *redistribuer* la richesse au nom de l'équité et de la justice distributive. Or, cette conception est fautive ; un grand nombre d'institutions sociales n'ont aucunement pour but de réaliser un objectif de justice distributive. Elles ont pour but de réaliser des gains collectifs de nature essentiellement différents des gains économiques, à savoir les économies d'échelle, la gestion du risque, la protection contre l'instabilité de nos préférences ou encore la diminution des coûts d'apprentissage de l'information. Il est par exemple trompeur de considérer la sécurité sociale comme une méthode de *redistribution* des richesses, alors que cette institution a pour fonction de *gérer les risques*. Or, il s'agit de deux objectifs complètement différents. L'un est économique, l'autre concerne le sentiment d'insécurité.

La question qui se pose n'est pas celle de savoir en quoi consiste la distribution [...]. Il s'agit plutôt de savoir quel type de problèmes d'action collective on devrait résoudre, aux dépens d'autres types de problèmes d'action collective. [...] Il existe énormément de confusion à ce propos. Le « filet de sécurité sociale », par exemple, est souvent décrit comme un système de redistribution, plutôt qu'un système assurantiel géré par l'État (HEATH, 2012, p. 47).

Au lieu de considérer l'ensemble des institutions sociales comme étant des institutions visant à *corriger* les effets distributifs du marché, il faut les voir comme des institutions dont la fonction est de réaliser des objectifs collectifs *différents, sui generis* et parfois opposés les uns aux autres. Ces objectifs sont parfois incompatibles avec l'objectif d'efficacité poursuivi par le marché, mais ils sont tout aussi légitimes.

Ainsi, une analyse exhaustive des problèmes d'action collective devrait nous amener à reconnaître que le marché représente une institution visant à résoudre *un type particulier* de problème d'action collective et que la résolution de ce problème peut s'opposer à la résolution d'autres problèmes d'action collective tout aussi importants.¹³⁹

¹³⁹ L'économiste Kenneth Arrow (1950) a également mis en évidence le fait que les objectifs d'efficacité économique peuvent entrer en conflit avec d'autres objectifs sociaux tout aussi valables. Son fameux « théorème d'impossibilité » prouve qu'il est impossible pour un choix social de satisfaire simultanément tout

En résumé, nous avons vu qu'il n'est pas légitime de critiquer le marché en vertu du seul fait qu'il autorise la poursuite des intérêts personnels. En effet, nous avons vu que les acteurs économiques bénéficient d'une *caution morale* sur le marché, c'est-à-dire qu'ils sont autorisés à adopter certains comportements qui, hors du marché, seraient considérés comme moralement inacceptables. Nous avons vu cependant que cette caution morale est à la fois conditionnelle et limitée ; les acteurs économiques ne sont autorisés à poursuivre leur intérêt privé au détriment des intérêts de leurs concurrents que tant et aussi longtemps qu'ils ne tirent pas avantage de défaillances du marché et que le marché est bénéfique à la société dans son ensemble. Or, de ce dernier point de vue, nous avons vu que le marché ne doit pas être considéré comme systématiquement bénéfique, même lorsqu'il permet d'obtenir des gains d'efficience. En effet, il est possible que l'objectif poursuivi par l'institution du marché, la prospérité économique, contredise d'autres objectifs sociaux. Dans ce cadre, accorder systématiquement la priorité à la prospérité économique, au détriment des gains d'échelle, de la gestion du risque et des gains de l'engagement préalable ou de la transmission de l'information ne se justifie aucunement.¹⁴⁰

En conclusion, notre analyse précédente ne consistait pas à soutenir que la seule critique légitime du marché est celle des défaillances du marché. Elle consistait plutôt à soutenir que, contrairement à l'approche des parties prenantes, l'évaluation morale ne doit

un ensemble de critères intuitivement désirables. Ainsi, pour tout choix social portant sur au moins trois options, il est impossible que le résultat satisfasse simultanément les conditions désirables suivantes ; (U) universalité ou souveraineté du citoyen (pas de restriction sur les préférences individuelles), (D) non-dictature (le résultat n'est pas déterminé par les préférences d'une seule personne), (P) efficience de Pareto (il n'est pas possible d'améliorer la situation d'une personne sans détériorer la situation de quelqu'un d'autre), (I) indépendance des options (le fait qu'un individu préfère une option A à une option B ne dépend pas d'une option C). Or, il est possible d'interpréter ce théorème comme prouvant que le critère d'efficience n'est pas toujours désirable – puisqu'il peut entrer en conflit avec d'autres objectifs désirables.

¹⁴⁰ Il est aussi possible d'argumenter que la *caution morale* accordée aux acteurs du marché au nom de la prospérité économique et de l'amélioration du niveau général de bien-être n'est légitime que dans des pays pauvres. Dans les pays riches, au contraire, la violation des normes de la morale ordinaire, à la faveur de l'augmentation de la richesse générale, n'est plus légitime. « Ainsi, dans une société qui souffre de la pauvreté, les considérations de bien-être vont presque avec certitude justifier la relaxation de la morale ordinaire afin d'augmenter la production (sous condition que l'augmentation soit distribuée aux personnes qui en ont besoin). Mais si la production totale d'une société est plus que suffisante pour satisfaire les besoins de la population, et l'augmentation de la production ne sert, admettons, qu'à empêcher l'expression de demandes en faveur d'une juste distribution, les gains en production seront nettement moins justifiés (s'ils le sont encore) par la valeur morale du bien-être. » [“Thus a society plagued by poverty considerations of welfare will almost certainly justify relaxing common morality so as to increase output (at least if the increases are distributed to those in need). But if the total output of a given society is more than adequate to meet the needs of the population, and increasing output serves, say, mainly to forestall demands for a just distribution, gains in output will be supported much more weakly (if at all) by the moral value of welfare.” (MCMAHON, 1981, p. 272).]

pas porter principalement sur les traits de personnalités du dirigeant d'entreprise, mais principalement sur la moralité du marché dans sa globalité.

4. CONCLUSION

Nous avons vu que l'une des principales raisons pour les économistes de favoriser une conception égoïste de l'être humain consiste à vouloir adopter une conception « réaliste » et « non utopiste » de la nature humaine. Or, nous avons vu que le projet de fonder la société sur le seul intérêt personnel pose deux types de difficultés :

- (1) En premier lieu, ce projet risque d'avoir quatre conséquences sociales nuisibles.
- (2) En second lieu, nous avons vu quatre raisons de penser que l'hypothèse de l'égoïsme universel ne satisfait pas, elle non plus, l'exigence de « réalisme politique ».

Pour conclure ce chapitre, nous soutiendrons que le projet de « faire l'économie de la vertu » repose sur une conception inappropriée du « réalisme politique ». Nous proposerons également une possible explication du succès des doctrines politiques fondées sur l'intérêt personnel.

4.1. UNE FAUSSE CONCEPTION DU RÉALISME POLITIQUE

La conception du réalisme politique qui constitue la motivation principale de l'hypothèse de l'intérêt personnel universel repose sur l'idée que, pour être réaliste, une théorie politique doit reposer sur une conception *pessimiste* de la nature humaine. Il s'agit d'admettre ou de concéder que, parmi les plus puissantes des motivations humaines, figurent des motivations moralement problématiques.

Toutefois, cette conception du réalisme politique est trop exigeante. Il n'est pas nécessaire, pour qu'une représentation de la nature humaine soit *réaliste*, qu'elle dépeigne le *pire* de la nature humaine. En d'autres termes, il n'est pas correct que l'exigence d'éviter tout *angélisme* au sujet de la nature humaine, requiert de considérer la nature humaine sous un jour *pessimiste*. Même s'il est vrai qu'il existe, parmi les motivations humaines, des passions antisociales et égoïstes très puissantes, cela ne signifie pas que nous devons considérer ces motivations comme étant représentatives de la véritable nature de l'homme.

Comme le fait remarquer Bernard Williams, le fait que l'être humain soit capable, dans des circonstances hostiles, des pires atrocités, ne doit pas nous amener à considérer que ces circonstances révèlent la « véritable » nature de l'homme.

Si le test pour savoir comment sont *vraiment* les êtres humains s'effectue dans des conditions de grand stress, de privation ou de pénurie (le test que Hobbes, dans sa conception de l'état de nature, a imposé), nous pouvons alors demander: en quoi ce test est-il approprié? Des conditions de stress extrême ou de privation ne sont pas des situations typiques pour observer le comportement type d'un quelconque animal ou pour observer d'autres caractéristiques de l'être humain. Si quelqu'un affirme que, si vous voulez voir ce que les hommes sont vraiment, alors il faut les observer après qu'ils aient passé trois semaines sur un bateau de sauvetage, il n'est pas clair qu'il s'agisse d'une meilleure recommandation pour ce qui concerne leur motivation que pour ce qui concerne leur condition physique. (WILLIAMS, 1994, p. 8)

De la même manière que nous ne considérerons pas que l'état ou la physiologie d'un être humain dans une situation de privation durable soit caractéristique de la nature physique et physiologique de l'être humain, il est erroné de croire que, parce que l'homme serait susceptible, dans des circonstances extrêmes, de faire des actes moralement condamnables, nous pouvons conclure que l'homme est par nature moralement mauvais.

4.2. LA CRAINTE DU COMMUNISME

Le projet de « faire l'économie de la vertu » dans l'organisation de la vie sociale a rencontré un vif succès parce que ses partisans considèrent qu'il n'existe qu'un seul système politique alternatif et que celui-ci vise naïvement à *supprimer complètement* les tendances égoïstes de la nature humaine. Dans un tel système politique et économique, où le désir d'améliorer son propre sort est considérée comme moralement condamnable, il est attendu des membres de la société que toutes leurs actions soient dirigées vers la réalisation du bien de la société dans son ensemble. Or, ce système politique, cherchant à *supprimer* toute aspiration individualiste, n'est autre que le système *communiste* ou socialiste que la majorité des économistes rejette avec force.

Les théoriciens du socialisme soviétique avaient primitivement supposé que leur système pouvait se passer des stimulants qui se font valoir dans une économie libre et dont la plupart tiennent à l'égoïsme de l'individu. Ils avaient cru pouvoir les remplacer par l'enthousiasme et la conscience socialistes. Cette conception s'est bientôt heurtée aux données réelles de la nature humaine qui est loin d'être guidée par une abnégation sublime (KENDE, 1959, p. 348).¹⁴¹

Etant donné que le communisme, pour les libéraux et, parmi eux, nombre d'économistes, sert de *repoussoir théorique*, il n'est pas étonnant que, pour s'opposer à la *suppression* de l'égoïsme et de l'individu dans le communisme, ils aient fait l'apologie et la promotion de l'égoïsme dans un système de marché concurrentiel.

Cette défense d'une société d'*égoïstes* par opposition à la société *sacrificielle* du communisme se retrouve clairement exprimée chez Ayn Rand, qui a vécu et fui dans sa jeunesse les conditions de vie dramatiques du régime soviétique.

Le capitalisme est basé sur les droits individuels – non sur le sacrifice de l'individu pour le « bien public » de la collectivité. Le capitalisme et l'altruisme sont incompatibles. Il faut choisir l'un ou l'autre. ¹⁴²

On retrouve aussi le même argument chez Jean Tirole, qui considère que tout système politique reposant sur l'hypothèse que les individus sont capables de renoncer à leur intérêt privé pour promouvoir le bien public mènerait nécessairement à un socialisme dangereux pour les individus.

¹⁴¹ L'une des critiques adressées au système communiste pourrait aujourd'hui, à notre avis, être adressée au système de l'entreprise par actionnariat. En effet, dans le socialisme, il a été nécessaire d'introduire un système individualisé de récompenses afin de motiver les personnes à réaliser les objectifs sociaux. Cela a eu pour conséquence, selon la critique, que les « managers » du socialisme se sont concentrés sur la réalisation des objectifs quantitatifs du plan centralisé de production afin d'obtenir leurs primes personnelles, même si cela menait finalement à des résultats socialement inférieurs ou si cela exigeait de recourir à des méthodes douteuses : « [L]a consigne de « satisfaire les plans » peut dans bien des cas aller à l'encontre de la rationalité économique (gaspillage des fonds d'investissement), soit de la meilleure satisfaction des besoins (production de biens superflus). » (KENDE, 1959, p. 358). Or, il nous semble que nous retrouvons aujourd'hui un problème similaire au sein de l'organisation de l'entreprise, lorsque ceux qui prennent des décisions ne sont pas en même temps les propriétaires de l'entreprise, et n'assument ainsi pas les risques liés à l'activité économique : ils ont tendance à vouloir réaliser les objectifs quantitatifs établis dans la planification, afin d'obtenir leurs primes personnelles, même si, au final, la réalisation de ces objectifs exigent des actions contraires à l'intérêt de l'entreprise. Ce problème trouve même un nom dans la littérature économique : le problème du principal-agent.

¹⁴² "Capitalism is based on individual rights – not on the sacrifice of the individual for the "public good" of the collective. Capitalism and altruism are incompatible. It's one or the other." (RAND, 1967, p. §19).

Pour restreindre les choix et vous interdire de « botter en touche » par le biais d'une réponse chimérique, je reformulerai légèrement la question : « Dans quelle organisation de la société aimeriez-vous vivre ? » La question pertinente n'est en effet pas de savoir dans quelle société idéale nous aimerions vivre, par exemple une société dans laquelle les citoyens, les travailleurs, les dirigeants du monde économique, les responsables politiques, les pays privilégieraient spontanément l'intérêt général au détriment de leur intérêt personnel. [...] Le manque de prise en considération d'incitations et de comportements pourtant fort prévisibles [...] a mené par le passé à des formes d'organisation de la société totalitaires et appauvrissantes (TIROLE, 2016, p. 15).

Lorsque l'on associe respectivement l'égoïsme et l'altruisme à deux systèmes politiques et économiques concurrents, à savoir le marché concurrentiel comme système fonctionnant sur la base des égoïsmes individuels, et le communisme comme système économique basé sur l'oubli complet de soi, alors le projet de « mécanisation de la vertu » devient séduisant ; afin d'éviter le désastre communiste, il faut privilégier un système politique fondé sur l'égoïsme. Toutefois, cet argument ignore la possibilité de systèmes politiques intermédiaires, qui ne s'appuient ni exclusivement sur l'égoïsme, ni exclusivement sur l'altruisme.¹⁴³

¹⁴³ Nous avons étudié l'une des raisons pour lesquelles le marché ne devrait pas, selon les économistes, être « moralisé ». Cette raison consistait dans le fait qu'il est dangereux de faire reposer le système économique sur la supposition que les acteurs ont de bonnes intentions.

Une autre raison qui est invoquée pour refuser de « moraliser » le marché consiste dans la crainte d'une ingérence des acteurs politiques dans la sphère économique ou des acteurs économiques dans la politique. En effet, les relations entre ces deux sphères peuvent s'avérer problématiques et dangereuses pour la société : les risques de corruption, d'abus de pouvoir, de lobbying ou d'atteinte à la liberté, laissent penser qu'il vaut mieux que chaque domaine soit autonome. Cela signifie qu'il faut préserver les décisions des politiciens de l'influence des milieux économiques et préserver les interactions économiques de l'influence des politiciens. Toutefois, même s'il est vrai que l'interaction entre la politique et l'économie peut être dangereuse, cela n'implique pas que la *moralisation* du marché soit dangereuse. En effet, il existe une différence fondamentale entre, d'un côté, la *politisation* du marché, qui implique l'intervention de l'État dans les relations d'affaires, et, de l'autre côté, la *moralisation* du marché, qui consiste uniquement à reconnaître que les acteurs économiques sont soumis à certains devoirs moraux. « La moralisation n'implique pas la politisation, ni analytiquement, ni sociologiquement. La politisation implique une interférence du gouvernement dans les prises de décisions sur le marché ou l'imposition de changements dans la structure de pouvoir de l'entreprise. La moralisation soutient seulement que les entreprises, bien que libres, sur le plan légal, de prendre leurs propres décisions sur le marché, sont néanmoins soumises à des principes moraux positifs. » [“Moralization does not imply politicization, analytically nor sociologically. Politicization implies governmental interference with decision-making on the market or making compulsory changes in the political power structure of the corporation. Moralization holds that businesses, although legally free to make their own decisions on the market, should also be bound by positive moral principles.” (DUBBINK & VAN DE VEN, 2012, p. 232)].

4.3. UNE LISTE NON EXHAUSTIVE DES POSSIBILITÉS

Il peut sembler que la situation dialectique dans laquelle nous nous trouvons est bloquée entre deux conceptions irréalistes de la société. D'un côté, la conception sacrificielle de la morale communiste, qui exige des individus de renoncer complètement à leur individualité au nom de l'intérêt général apparaît comme irréalisable – voire cruelle – puisqu'elle repose sur une conception utopiste de la nature humaine. De l'autre côté, une société qui reposerait sur la poursuite illimitée des intérêts personnels ne semble pas offrir de meilleures perspectives, pour les raisons que nous avons vues dans les chapitres précédents. Heureusement, le sacrifice total de soi et la poursuite sans limite de l'intérêt personnel ne constituent pas les deux seules conceptions de la nature humaine disponibles pour élaborer une doctrine politique.

Pour saisir la possibilité d'une troisième option, il est nécessaire de comprendre que la notion de *sacrifice de soi* peut recevoir plusieurs définitions. En effet, il existe *deux conceptions* du sacrifice de soi. La notion *forte* est celle qui est nécessaire à la société communiste. Elle peut être définie comme suit :

Selon la première conception, forte, le sacrifice implique de *supprimer, d'éliminer* ses intérêts personnels pour favoriser le bien commun [...] (HAMEL, 2015, p. 112).

Le sacrifice des intérêts personnels est ici *absolu*. Lorsque les partisans de la « mécanisation de la vertu » soutiennent qu'il n'est pas possible de se passer du motif de l'intérêt personnel dans l'organisation de la société, ils rejettent la conception *forte* du sacrifice de soi. Ils considèrent qu'une société dans laquelle l'intérêt personnel est *éliminé* n'est pas possible.

Il existe toutefois une *seconde* conception du sacrifice de soi qui n'est pas aussi exigeante que la première. Selon cette seconde conception, la vie en société ne requiert pas un sacrifice *absolu* des intérêts privés, mais un sacrifice *partiel* de ceux-ci. Ce sacrifice partiel des intérêts privés se divise en deux catégories : d'un côté, la satisfaction de certains intérêts particuliers reste *absolument* interdite car elle est toujours incompatible avec l'intérêt public. De l'autre côté, la satisfaction des autres intérêts est *autorisée*, mais uniquement dans les limites de sa compatibilité avec l'intérêt général.

Selon cette seconde conception, modérée, du sacrifice, celui-ci implique soit de ne pas satisfaire ceux de nos intérêts personnels qui seraient incompatibles avec le bien commun – par exemple, resquiller quand on sait que les chances d’être sanctionné sont improbables ou nulles –, soit de ne pas maximiser la satisfaction de ceux de nos intérêts dont la satisfaction *maximale* (mais pas la satisfaction tout court) entrerait en conflit avec les exigences du bien commun [...] – par exemple, satisfaire son désir de gagner une élection (parfaitement légitime en lui-même) en remplissant les urnes de bulletins d’électeurs fictifs (HAMEL, 2015, p. 113).

En conséquence, une conception *réaliste* de la vie en société ne devrait ni supposer qu’il est possible *d’éliminer complètement* la recherche de l’intérêt personnel, ni supposer qu’il est possible de fonder la vie sociale exclusivement sur l’intérêt personnel, sans faire appel aux dispositions morales des membres de la société. En ce sens, l’exigence de *réalisme politique* ne signifie pas qu’il est utopiste d’attendre des individus qu’ils sacrifient (modérément) leurs intérêts personnels, mais seulement qu’il est utopiste d’espérer pouvoir *éliminer* le désir de richesse personnelle ou d’aisance matérielle (HAMEL, 2015, pp. 109-110).

Si on considère que la poursuite du bien commun n’est pas *incompatible* avec la poursuite de l’intérêt privé, mais qu’elle lui impose des *limites*, alors l’idée qu’un individu puisse *intentionnellement* accorder de l’importance au bien commun en *refusant d’accomplir* les actions incompatibles avec le bien commun semble tout à fait plausible. L’individu ne doit pas intentionnellement renoncer à tout intérêt privé, mais seulement *limiter* la satisfaction de ses intérêts privés à ce qui est compatible avec le bien commun.

Bien entendu, l’ensemble du système suppose que les citoyens soient *informés* et puissent *accepter* les limites imposées à la poursuite de leurs intérêts privés au nom de certains objectifs moraux. Cela suppose donc que *les exigences morales* elles-mêmes ne requièrent pas une élimination totale des intérêts privés. Selon Christopher Hamel, il est *évident* que cela n’est pas le cas ; il est *évident* que les exigences morales ne peuvent pas être entièrement incompatibles avec la poursuite de l’intérêt personnel. Selon lui, au contraire, les exigences morales sont *compatibles* avec la satisfaction d’une large extension d’intérêts personnels.

Pour confirmer cette compatibilité, il faudrait montrer que le souci du bien commun – le motif civique –, qui définit la vertu civique d’esprit public, n’exige pas l’élimination ou la suppression de tous les intérêts personnels. [...] Or, [...] personne ne défend l’idée selon laquelle le bien commun n’est constitué *d’aucun* intérêt particulier (HAMEL, 2015, p. 110).

Selon Hamel, *aucun auteur* ne soutient que les individus doivent *entièrement* sacrifier leur intérêt privé au nom de l’intérêt collectif. Tous les philosophes admettent donc, selon lui, une forme de compatibilité entre la satisfaction des intérêts privés et celle de l’intérêt collectif. Pour quelles raisons admettent-ils *a priori* une telle compatibilité ? Nous verrons et critiquerons, par la suite, les raisons pour lesquelles ces auteurs – dont Hamel – supposent que les normes morales et politiques ne s’opposent jamais complètement à la satisfaction des intérêts privés.

TROISIÈME PARTIE

L'erreur moraliste :
L'argument des métaéthiciens

CONTRE L'ALTRUISME

1. UNE SOCIÉTÉ D'INDIVIDUS PUREMENT ALTRUISTES N'EST PAS DÉSIRABLE

Dans la précédente partie, nous avons soutenu la thèse selon laquelle toute société devait reposer sur la motivation spécifiquement morale des personnes pour pouvoir fonctionner de manière stable et durable. Pour parvenir à cette conclusion, nous avons vu plusieurs raisons de considérer qu'une société fondée exclusivement sur l'intérêt personnel n'est pas désirable. De la même manière, nous avons affirmé que le système alternatif, basé exclusivement sur *l'altruisme*, n'est pas non plus désirable. Pour défendre cette dernière thèse, nous avons accepté l'argument selon lequel il est *naïf* et utopiste de vouloir empêcher ou interdire *toute expression* de l'intérêt personnel. De ce fait, un système fondé sur l'élimination totale de l'intérêt personnel n'est pas non plus réalisable.

Il ne s'agit toutefois pas du seul argument qui puisse être avancé pour appuyer la thèse selon laquelle une société composée d'individus purement altruistes n'est pas désirable. Après avoir présenté d'autres arguments, nous tenterons d'expliquer et de résorber la confusion très répandue entre motivation morale et altruisme. Nous pensons que la confusion entre motivation morale et altruisme explique la croyance erronée que l'altruisme constitue la seule alternative possible à l'égoïsme. En supprimant la confusion, nous espérons donc corriger la croyance.

1.1. L'IMPOSSIBILITÉ DE L'ALTRUISME

Dans cette section, nous verrons que les dispositions altruistes ne suffisent pas pour garantir le bon fonctionnement de la société. Pour le montrer, nous allons avancer trois arguments : (1) l'altruisme n'est possible que s'il existe des motivations égoïstes ; (2) un altruiste est incapable d'arbitrer entre les intérêts conflictuels des personnes dont il se soucie ; (3) les motivations altruistes peuvent motiver les personnes à commettre des actions immorales.

Le premier argument est le suivant : pour qu'il soit possible que des personnes altruistes puissent satisfaire leur désir de faire le bien d'autrui, il est nécessaire qu'il existe

des personnes égoïstes qui acceptent que leur propre bien, plutôt que celui des autres, soit réalisé (par autrui). En effet, si tout le monde était altruiste, il ne serait plus possible de réaliser le bien de quiconque, car chacun refuserait que son propre bien soit réalisé plutôt que celui de quelqu'un d'autre. Autrement dit, il faut qu'il existe des personnes qui se soucient de leur propre bien pour qu'il puisse y avoir des personnes altruistes.

John Rawls (2009 , p. 209) a avancé un argument de ce type.¹⁴⁴ Son argument affecte uniquement l'altruisme fondé sur une conception *spécifique* du bien des personnes, à savoir la conception selon laquelle le bien des personnes se mesure en fonction de la satisfaction de leurs préférences. Etant donné, toutefois, qu'il s'agit d'une conception répandue du bien personnel, admise notamment par les économistes, l'argument de Rawls possède une certaine portée et mérite notre attention.

L'argument de Rawls se présente comme suit : admettons que le bien d'une personne soit défini par la satisfaction de ses préférences. Imaginons ensuite une société d'altruistes parfaits dans laquelle chacun désire exclusivement faire le bien d'autrui, c'est-à-dire satisfaire les préférences des autres. Or, dans une telle société, il est impossible de satisfaire les préférences de quiconque. En effet, si un individu A préfère satisfaire les préférences d'un autre individu B, mais que B lui-même préfère satisfaire les préférences d'une tierce personne C, alors A devrait satisfaire les préférences de C pour satisfaire les préférences de B. Cependant, si C elle-même a des préférences qui portent sur les préférences de quelqu'un d'autre, alors la régression dans la satisfaction des préférences continue – possiblement à l'infini – et il n'est jamais possible de satisfaire les préférences de quiconque.

La difficulté provient du fait que les individus parfaitement altruistes ont uniquement des préférences de second ordre, c'est-à-dire des préférences qui elles-mêmes portent sur des préférences, mais ils ne possèdent pas de préférences de premier ordre, c'est-à-dire des préférences qui portent directement sur des objets du monde. En conséquence, chacun ne se soucie que de satisfaire les préférences des autres, mais personne n'ayant de préférence portant directement sur des objets, il devient impossible de satisfaire les préférences de quiconque.

¹⁴⁴ David Estlund (1990) reprend et développe cet argument.

Un altruiste parfait ne peut satisfaire son désir que si quelqu'un d'autre a des désirs indépendants ou « de premier ordre » [i.e. désirs portant sur des objets indépendants et non sur d'autres désirs, à la différence des désirs d'ordre plus élevé ou de « deuxième ordre », comme l'amour et l'altruisme qui, eux, portent sur des désirs déjà donnés.] Pour illustrer ce fait, supposons que, quand il faut prendre une décision, tous choisissent de faire ce que tous les autres veulent faire. Il est alors évident que rien ne sera décidé, en fait il n'y a rien à décider (RAWLS, 2009 , p. 219).

Pour que la régression dans la satisfaction des préférences prenne fin, il faudrait qu'au moins une personne ait des préférences ou désirs qui ne concernent pas les préférences ou désirs de quelqu'un d'autre, mais qui concernent directement des objets. Il faut qu'au moins une personne ait des préférences ou désirs « de premier ordre ».

En fait, il semble que la situation de l'altruiste décrit par Rawls soit encore plus problématique que ne le suppose le philosophe. En effet, un parfait altruiste du type envisagé par Rawls ne peut tout simplement pas *exister* dans le monde physique tel que nous le connaissons. En effet, un individu dont les désirs ne portent que sur les désirs d'autrui ne survivrait pas longtemps ; il ne désire ni boire, ni manger, ni avoir une bonne santé, ni avoir des amis ou un travail. Il semble alors évident que ses chances de survie à court terme sont nulles.

Un altruiste plus raisonnable que celui que nous décrit Rawls posséderait bien des désirs qui le concernent directement, tels que les désirs de se nourrir, d'avoir une bonne santé et un travail. En effet, il aurait conscience que, pour pouvoir satisfaire son désir de second ordre de faire le bien d'autrui, il a besoin d'être lui-même en bonne santé, autonome et capable de venir en aide aux autres personnes. De ce fait, notre altruiste aurait le désir d'être lui-même en bonne santé, autonome et capable de venir en aide aux autres personnes. De tels désirs, bien qu'ils concernent le *soi*, ne sont cependant pas des désirs égoïstes, puisqu'ils dérivent du désir altruiste de satisfaire les préférences des autres personnes.

Par contre, il reste vrai que cet individu altruiste, en bonne santé, autonome et capable de satisfaire les préférences des autres personnes, sera toujours incapable de faire le bien d'autrui, puisque les préférences des autres personnes concernent toutes, elles aussi, les

préférences d'autres personnes. De ce fait, la société d'individus purement altruistes sera une société dans laquelle chacun a satisfait ses besoins fondamentaux lui permettant d'être apte à satisfaire les préférences des autres personnes, mais personne ne pourra jamais exercer son altruisme en faveur d'autrui. Il n'est donc pas certain que, dans une telle société, où chacun se contente de satisfaire ses besoins fondamentaux, les individus se portent mieux que dans une société dans laquelle chacun poursuit la satisfaction de préférences égoïstes.

Bien entendu, nous rétorquera-t-on, l'avantage de la société altruiste est que personne n'acceptera jamais de *porter atteinte* aux intérêts d'autrui dans le but de satisfaire ses propres préférences – contrairement à ce qui serait le cas dans une société d'individus égoïstes. En ce sens, en ce qui concerne le risque de subir des nuisances des autres personnes, la situation est bien meilleure dans la société altruiste.

La seule remarque que nous pouvons faire à ce titre est que la comparaison entre la société altruiste et la société égoïste dépend largement des circonstances. Il n'est pas possible d'offrir une réponse *de principe* à la question de savoir si l'une est meilleure que l'autre. En effet, si la société composée d'égoïstes n'offre pas de garantie contre les nuisances, la société altruiste, de son côté, n'offre pas de perspective de *dépasser* un certain niveau de satisfaction de préférences.

Il nous faut encore noter que l'argument de Rawls dépend entièrement du fait que le bien-être des personnes est défini en termes subjectifs de satisfaction de préférences. La situation serait toute différente si le bien-être des personnes était défini en termes objectifs, selon une liste objective des composants d'une vie bonne et réussie. Par exemple, si une vie bonne consiste à avoir ses besoins fondamentaux satisfaits, des relations sociales solides, une fonction sociale chargée de sens, une bonne santé durable, des activités de loisirs, etc., alors il devient possible d'agir pour le bien d'autres personnes indépendamment de leurs préférences. Toutefois, dans ce cadre, des problèmes de paternalisme et d'atteinte à l'autonomie individuelle apparaissent immédiatement. Il n'est donc pas non plus évident qu'une société altruiste, où le bien des personnes est défini par une liste objective, soit une société meilleure qu'une société égoïste.

En conclusion, bien qu'il ne soit pas possible d'affirmer qu'une société altruiste est *moins bonne* qu'une société égoïste, il n'est pas non plus possible d'affirmer le contraire. Il

semble que la réponse dépende très largement des circonstances. En outre, comme nous le verrons dans les sections suivantes, une société fondée sur le seul altruisme rencontre d'autres difficultés.

1.2. L'IMPOSSIBILITÉ POUR L'ALTRUISTE DE RÉSOUDRE LES CONFLITS

Un second problème, avancé également par John Rawls, consiste dans le fait qu'une personne altruiste se soucie bien souvent de promouvoir le bien de plusieurs personnes, voire, lorsqu'il s'agit d'un altruiste parfait, elle se soucie de promouvoir le bien de tout le monde. Or, un tel altruiste, ne possédant aucun critère moral d'évaluation et de balance des intérêts, est incapable d'arbitrer les intérêts contraires des personnes dont il se soucie.

Très souvent, la façon dont on doit réaliser ce désir [de favoriser le bien d'autrui] est assez claire. La difficulté vient de ce que l'amour pour plusieurs personnes à la fois conduit à la confusion quand celles-ci émettent des revendications conflictuelles. [...] L'altruiste est désorienté tant que ses nombreuses affections sont en opposition dans les personnes qu'il aime. (RAWLS, 2009 , p. 220)

En l'absence d'une conception du bien moral ou de la justice, il est impossible, pour l'altruiste, de savoir comment résoudre les conflits d'intérêts qui peuvent apparaître entre les personnes auxquelles il accorde de la valeur. Or, les problèmes de conflit entre égoïsme et altruisme apparaissent précisément dans les situations où les intérêts de plusieurs personnes ne peuvent pas être satisfaits simultanément. Dans une société d'altruistes parfaits, comme dans une société composée exclusivement d'individus égoïstes, il n'existe pas de possibilité de résoudre les conflits d'intérêts.

1.3. LE MANQUE DE FIABILITÉ DE L'ALTRUISME

Un autre problème est que les dispositions altruistes, tout comme les dispositions égoïstes, peuvent, dans certaines circonstances, amener une personne à accomplir une action moralement mauvaise. Certes, la plupart du temps, nos dispositions altruistes nous poussent à faire des actions moralement louables. Par exemple, nos dispositions altruistes peuvent nous motiver à aider une grand-mère à transporter ses lourds achats chez elle.

Cependant, il peut arriver que nos motivations altruistes nous fassent accomplir des actions moralement mauvaises. L'exemple classique est celui d'un braqueur de banque qui peine à extraire du bâtiment le lourd magot volé dans le coffre-fort – de la même manière que la grand-mère peinait à transporter ses achats chez elle (HERMAN, 1981, p. 364). Voyant le bandit ainsi en difficulté, un individu altruiste, poussé par son désir de venir en aide à autrui, sera motivé à apporter son aide au malfrat. Or, dans ce second cas, contrairement au premier, la motivation altruiste pousse notre individu à accomplir une action moralement discutable.¹⁴⁵

Pour citer d'autres exemples, plus ordinaires et moins extravagants, prenons le cas suivant : l'amour pour nos proches, à l'instar de la bienveillance, nous pousse la plupart du temps à faire des bonnes actions. Il est en effet moralement approprié, en général, d'aider un ami ou un proche. Cet amour peut cependant également nous amener à accomplir des actions immorales. Par exemple, si un ami ou un proche nous demande de l'aider à dissimuler sa fortune au fisc, notre affection pour lui peut nous amener à l'aider et, ainsi, à commettre une faute morale – sous l'hypothèse, bien entendu, que l'honnêteté fiscale soit un devoir moral. Dans cet exemple, comme dans celui du braqueur de banque, nos dispositions altruistes nous poussent à faire des actions immorales.

Comme nous pouvons l'observer dans les exemples précédents, il dépend entièrement des circonstances entourant le choix que l'action satisfaisant les désirs altruistes soit simultanément l'action moralement appropriée. De ce constat, nous pouvons tirer la conclusion qu'il est dangereux de s'appuyer exclusivement sur les motivations altruistes pour faire fonctionner la société. En effet, même si ces motivations poussent le plus souvent les individus à faire des actions moralement bonnes, elles risquent également de les amener, en fonction des circonstances, à accomplir des actions immorales. C'est pourquoi il est risqué de faire dépendre de motivations *non morales* le bon fonctionnement de la société, que celles-ci soient égoïstes ou altruistes. En effet, si de telles motivations peuvent, selon les circonstances, amener des personnes à bien agir, tout changement de circonstances risque d'avoir pour conséquence que ces mêmes motivations donnent lieu à des actions immorales. Comme le dit Barbara Herman :

¹⁴⁵ Pour rappel, il suffit de *se soucier des intérêts de quelqu'un d'autre* pour être considéré comme altruiste. C'est le cas de notre individu qui cherche à aider le voleur. Il se soucie du bien du bandit.

Même si les institutions sociales étaient organisées pour garantir que le profit et l'honnêteté réalisés simultanément [par le biais de pénalités, de sanctions sociales, ...] l'accomplissement de l'action morale, ainsi motivé, ne serait toujours que l'effet de la « chance » ; c'est-à-dire qu'il dépend entièrement de circonstances externes et contingentes. [...] Des motifs non moraux peuvent bien amener à accomplir des actions morales, et ils peuvent même le faire avec n'importe quel degré désiré de régularité. Le problème est que les actions moralement correctes sont le produit d'un alignement chanceux des motivations et des circonstances.¹⁴⁶

A l'inverse, en faisant appel aux motivations spécifiquement morales, il ne dépend plus de circonstances factuelles et contingentes que les individus soient motivés à faire l'action qui réalise le bien de la collectivité.¹⁴⁷ Une société fondée sur la motivation morale de ses membres est plus viable et robuste, en ce qui concerne la réalisation du bien commun et de la justice, car sa préservation ne dépend pas entièrement des circonstances non morales qui entourent les choix individuels.¹⁴⁸

¹⁴⁶“Even if social institutions were arranged to guarantee that profit and honesty went together (through penalties, social sanctions, etc.), the performance of honest actions, so motivated, would still be no more than "fortunate": that is, dependent on external and contingent circumstances.[...] Nonmoral motives may well lead to dutiful actions, and may do this with any degree of regularity desired. The problem is that the dutiful actions are the product of a fortuitous alignment of motives and circumstances.” (HERMAN, 1981, p. 366).

¹⁴⁷ Le problème du caractère contingent de la convergence entre l'intérêt personnel et la vertu est formulé par Christopher Hamel dans les termes suivants : « On ne peut certes pas exclure que dans un certain nombre de cas, l'intérêt et la vertu se trouveront converger, mais une politique de la vertu civique qui se contente d'une telle convergence factuelle et contingente est vouée à l'échec : elle a par définition besoin de dispositions qui rendent les individus capables de ne pas chercher à satisfaire ceux de leurs intérêts qui mettraient en danger les institutions gardiennes de la liberté. » (HAMEL, 2015, p. 114). Nous proposons d'étendre cette analyse aux dispositions altruistes.

¹⁴⁸ Gilbert Harman développe un argument différent pour soutenir l'idée que la motivation morale est plus susceptible de contrer l'égoïsme que la bienveillance ou l'altruisme : « La bienveillance est une motivation trop faible pour être identifiée avec la motivation de faire ce qui est juste/correct. Les individus sont parfois motivés par la bienveillance à améliorer le sort des autres personnes, mais cette motivation est en général plutôt fragile lorsqu'on la compare à l'aversion que les individus éprouvent ordinairement à l'égard du meurtre, du tort fait à autrui, du vol, du mensonge, du non-respect des promesses et du remboursement d'une dette. La bienveillance généralisée est normalement bien plus faible que le motif de l'intérêt personnel. Mais les motivations morales que nous venons de mentionner – éviter de tuer autrui et ainsi de suite – sont souvent aussi fortes, voire plus fortes, que l'intérêt personnel. La bienveillance généralisée ne peut pas constituer le tout de la motivation morale. » [“Benevolence is too weak a motive to be identified with the motive to do what is right. People are sometimes motivated by benevolence to try to improve the lot of other people, but this motivation is normally quite feeble when compared to ordinary people's aversions to murder, injuring others, stealing, lying, and failing to keep their promises and pay their debts. Generalized benevolence is normally a much weaker motive than self-interest. But the moral motives just mentioned – to avoid killing others, and so forth – are often just as and sometimes stronger than self-interest. Generalized benevolence cannot be the whole story about moral motivation.” (HARMAN, 2000, p. 186).]

En résumé, il est nécessaire qu'une personne altruiste dispose d'une notion de justice pour qu'elle choisisse de satisfaire, parmi ses désirs altruistes, ceux qui participent au bien de la société. Pour qu'une personne altruiste préfère aider un étudiant fauché à un joueur de poker fauché, une vieille dame en difficulté plutôt qu'un voleur trop chargé, un frère qui rencontre des difficultés administratives plutôt qu'un frère qui essaie de frauder le fisc, il est nécessaire qu'elle dispose d'une notion de justice qui oriente sa motivation altruiste dans une direction bénéfique à la société. L'altruisme seul ne suffit pas.

2. LA CONFUSION ENTRE ALTRUISME ET MOTIVATION MORALE

Dans la section précédente, nous avons vu trois raisons de considérer qu'une société composée d'individus parfaitement altruistes n'est pas plus désirable qu'une société composée d'individus exclusivement égoïstes. Cette conclusion permet de rejeter la thèse (à notre avis largement partagée dans la gauche de l'échiquier politique) selon laquelle il faut remplacer la société commerçante et capitaliste actuelle par une société non commerçante, fondée sur l'altruisme.

Dans la section présente, nous tenterons d'expliquer l'origine de la croyance (fausse) que la seule solution alternative à une société basée sur l'égoïsme est une société fondée sur l'altruisme. Selon nous, cette croyance provient du fait que la motivation morale et l'altruisme se trouvent souvent (faussement) identifiés l'un à l'autre. En expliquant et critiquant cette tendance à considérer l'altruisme et la motivation morale comme étant identiques, nous découvrirons que la forme d'égoïsme qui doit être combattue dans une société n'est pas forcément la forme d'égoïsme qui consiste à se soucier de son propre bien plutôt que de se soucier de celui des autres.

2.1. POURQUOI CONFOND-ON ALTRUISME ET MOTIVATION MORALE ?

La confusion entre motivation morale et altruisme s'explique assez facilement. En premier lieu, il est courant d'opposer l'égoïsme à l'altruisme. En effet, alors que le premier consiste en un désir de réaliser son propre bien, le second consiste en un désir de faire le bien de quelqu'un d'autre. L'objet sur lequel portent ces deux désirs est ainsi de nature

opposée. L'un exprime un souci de soi et l'autre un souci d'autrui. Bien entendu, les désirs altruistes et égoïstes existent sous de multiples formes : les désirs de sauver la vie de quelqu'un d'autre, de soulager sa souffrance, ou de le rendre heureux constituent des exemples de désirs altruistes. Les désirs d'éprouver du plaisir, de devenir riche ou célèbre sont quant à eux des cas emblématiques de désirs égoïstes (DORIS, STICH , PHILLIPS , & WALMSLEY , 2015). Ainsi, selon cette première approche, égoïsme et altruisme représentent deux motivations opposées.

Thèse 1 : L'égoïsme est l'opposé de l'altruisme.

Il est également usuel d'opposer l'égoïsme à la motivation morale, dans le sens où l'égoïsme serait la source principale de l'échec moral. Autrement dit, si les individus échouent souvent à faire ce qui est bon/juste/vertueux, cela est dû principalement au fait qu'ils sont poussés par égoïsme à agir d'une autre manière. L'égoïsme représenterait ainsi la source du mal dans notre monde. Il est ce qui empêche l'homme d'accomplir le bien, voire ce qui pousse l'homme à faire le mal. Dans cette seconde perspective, l'égoïsme est traité comme étant l'opposé de la motivation morale.

Thèse 2 : L'égoïsme est l'opposé de la motivation morale.

Le fait que l'égoïsme soit opposé à la fois à l'altruisme et à la motivation morale amène comme conséquence fâcheuse que la motivation morale et l'altruisme se trouvent souvent identifiés l'un à l'autre.¹⁴⁹ Cette conséquence semble tout à fait logique. Si les contraires de l'égoïsme sont aussi bien l'altruisme que la motivation morale, alors altruisme et motivation morale doivent avoir le même contenu. Plus formellement, la confusion entre altruisme et motivation morale peut être expliquée par le raisonnement suivant :

¹⁴⁹ On trouve des exemples de cette identification de la morale et de l'altruisme dans le langage des politiciens – notamment de gauche – mais aussi dans la littérature scientifique. L'exemple suivant est cité dans l'article de la *Stanford Encyclopedia of Philosophy* sur la psychologie morale (DORIS, STICH , PHILLIPS , & WALMSLEY , 2015): « Le comportement moral est, à son niveau le plus général, un comportement altruiste, motivé par le désir de promouvoir non seulement son propre bien-être mais aussi le bien-être des autres. » [*"Moral behavior is, at the most general level, altruistic behavior, motivated by the desire to promote not only our own welfare but the welfare of others"* (RACHELS, 2000, p. 81)]. Nous retrouvons cette identification de la motivation morale et de l'altruisme dans l'article sur la motivation morale de la *Routledge Encyclopedia of Philosophy*: "Un agent véritablement moral – contrairement à un égoïste – est altruiste, dans le sens qu'il se préoccupe directement des intérêts d'autrui". [*"A genuinely moral agent – as opposed to an egoist – is altruistic, in the sense of having some immediate concern for the interests of others."* (WALLACE, 1998, p. 525).]

Thèse 1 : L'égoïsme est l'opposé de l'altruisme.

Thèse 2 : L'égoïsme est l'opposé de la motivation morale.

Conclusion : $\therefore$ La motivation morale est identique à l'altruisme.

Alors que les deux prémisses paraissent tout à fait acceptables, nous souhaiterions rejeter la conclusion, puisque nous voulons distinguer la motivation morale de l'altruisme. Le défi consiste donc à rendre compte des deux prémisses, tout en parvenant à rejeter la conclusion. Comment procéder pour à la fois (a) préserver l'opposition entre l'égoïsme et l'altruisme (Thèse 1), (b) préserver l'opposition entre l'égoïsme et la motivation morale (Thèse 2), tout en (c) se débarrassant de l'identification gênante entre la motivation morale et l'altruisme (Conclusion)?

2.2. DEUX FORMES D'ÉGOÏSME

Pour parvenir à préserver nos intuitions concernant l'opposition entre égoïsme et altruisme et entre égoïsme et motivation morale, sans conclure que l'altruisme et la motivation morale sont identiques, il est nécessaire de changer notre manière de concevoir la relation entre ces trois motivations. En effet, aussi longtemps que nous focalisons notre attention sur le *bénéficiaire* de l'action, il est impossible de distinguer la motivation morale de l'altruisme. Pour pouvoir les distinguer, il convient donc de ne pas définir la motivation morale sur le même niveau que l'égoïsme et l'altruisme, c'est-à-dire au niveau du bénéficiaire de l'action. Autrement dit, il faut dépasser l'opposition entre bien *pour soi* et bien *pour autrui*.

La solution consiste alors à montrer que le concept *d'égoïsme* auquel il est fait référence dans les thèses (1) et (2) n'est pas le même concept dans les deux cas. Si le premier concerne bien l'opposition entre *soi* et *autrui*, le second concept d'égoïsme concerne un autre niveau d'analyse. Ainsi, alors que l'égoïsme opposé à l'altruisme est un souci exclusif de soi, l'égoïsme opposé à la motivation morale consiste à vouloir « faire une exception pour soi ». Nous pouvons alors reformuler les thèses précédentes, pour rendre manifeste le fait que les thèses (1) et (2) ne nous forcent pas à identifier la motivation morale à l'altruisme.

- Thèse 1 : L'égoïsme, comme souci de son propre bien-être, est l'opposé de l'altruisme comme souci du bien-être d'autrui.
- Thèse 2 : L'égoïsme, comme désir de faire une exception pour soi, est l'opposé de la motivation morale, comme désir de ne pas faire une exception pour soi.

Comme il est désormais évident, les deux formes d'égoïsmes ne sont pas identiques. L'égoïsme opposé à l'altruisme consiste à désirer améliorer son propre bien-être. L'égoïsme opposé à la motivation morale consiste, de son côté, à désirer que les autres respectent certains principes tout en s'accordant la liberté de ne pas s'y soumettre soi-même. C'est à cette seconde forme d'égoïsme que nous allons nous intéresser par la suite. Il ne consiste plus simplement à poursuivre son propre bien-être individuel en restant indifférent au bien-être d'autrui – ou au détriment du bien-être d'autrui – mais à poursuivre son propre bien *en violation de principes* sans lesquels la vie commune – et la poursuite de nos intérêts propres – ne serait plus possible. Cette forme d'égoïsme consiste à poursuivre des gains personnels au détriment de principes qui rendent possibles la vie commune. Ce que nous retiendrons pour l'instant est que cet égoïsme opposé à la morale, i.e. le désir de faire une exception pour soi, n'est pas identique à l'égoïsme opposé à l'altruisme – i.e. l'égoïsme comme *souci exclusif* de son propre bien-être.

DÉDRAMATISER LE CONFLIT ENTRE MORALE ET INTÉRÊT PERSONNEL

1. DEUX ARGUMENTS

Jusqu'à présent, nous avons vu qu'une société basée exclusivement sur l'intérêt personnel n'est pas réalisable ; il est nécessaire, pour qu'une société fonctionne, de faire appel à la vertu des acteurs du marché et des citoyens en général. Nous avons vu également que la disposition morale dont dépend la société n'est pas une forme d'altruisme, mais bien le souci moral de faire ce qui est bien, de faire ce qui est juste ou d'être une personne vertueuse.

Une question qui se pose désormais consiste à déterminer la gravité de notre situation présente ; à quelle *fréquence* est-il nécessaire de faire appel aux dispositions morales des citoyens ? A quelle fréquence doivent-ils accepter de sacrifier (au moins une partie de) leurs intérêts personnels ? Quelle est également l'*étendue* des sacrifices demandés ? Peut-il arriver que la réalisation du bien moral et de la justice exige un sacrifice *complet* des intérêts privés ? Ou existe-t-il des *limites* aux sacrifices qui peuvent être exigés d'une personne ?

Dans ce contexte, il est intéressant de constater l'attrait qu'exerce l'exigence méthodologique de « faire l'économie de la vertu », même sur celles et ceux qui considèrent qu'il est impossible de se dispenser complètement de la motivation morale des citoyens dans l'organisation de la société. En effet, même les plus fervents critiques du projet de « mécanisation de la vertu » admettent implicitement ou même explicitement que les sacrifices demandés ne doivent pas être *trop importants* ou *trop fréquents* pour que le modèle envisagé de société soit réalisable. Ils concèdent donc que, bien qu'il soit nécessaire d'attendre des sacrifices de l'intérêt personnel de la part des citoyens, ces sacrifices doivent néanmoins rester rares et limités. Sinon, le modèle de société proposé est à nouveau considéré comme irréaliste et utopiste.

Ainsi, par exemple, même Christopher Hamel, qui critique fermement le programme de « mécanisation de la vertu », se soucie dans son texte de nuancer sa conclusion en montrant que les sacrifices demandés aux citoyens, au nom du bien commun et de la préservation des institutions communes, seront *partiels* et *peu fréquents*. Selon le philosophe, sa propre vision d'une société juste, à savoir une société républicaine faisant appel à la vertu d'esprit public des citoyens, est compatible avec la satisfaction d'un grand nombre d'intérêts privés. De ce fait, sa théorie n'est pas irréalisable et utopiste. Si Hamel admet – en passant – qu'il peut exister des situations extrêmes dans lesquelles les sacrifices demandés sont nettement plus importants, il soutient que, dans le cours normal des événements, les exigences de sacrifice des intérêts privés seront rares et limitées. Il ne se pourrait notamment pas que, « dans le cours normal des événements », les citoyens aient le devoir moral de sacrifier leur vie elle-même.

Il ne faut pas perdre de vue que le sacrifice *de soi* (de sa vie) est une possibilité – peut-être une nécessité – dans certaines situations extrêmes (de guerre, par exemple) ou dans des scénarios moraux extrêmes [...]. Mais dans des circonstances politiques normales, il est clair que s'il est [...] invraisemblable que le bien commun soit par principe exclusif de *tous* [les] intérêts particuliers, il est a fortiori inconcevable que le coût de la vertu civique soit la vie même de l'individu (HAMEL, 2015, p. 112).

Ainsi, selon l'auteur, dans les circonstances normales de la vie en société, les sacrifices demandés aux citoyens et acteurs économiques sont seulement limités. Par cet argument, Hamel *dédramatise* sa conclusion selon laquelle une société a besoin de s'appuyer sur la motivation morale – ou vertu civique – de ses citoyens pour fonctionner. Il rend sa conclusion moins « utopiste » qu'elle ne peut paraître de prime abord.

L'argument de Christopher Hamel représente une version de deux types différents d'arguments de « dédramatisation » du recours à la motivation morale dans l'organisation de la société. Le premier type d'argument, dont l'argument de Hamel est un exemple, consiste à montrer que, dans les circonstances réelles et imparfaites du monde actuel, nous pouvons raisonnablement nous attendre à constater une très forte convergence entre l'intérêt personnel et la vertu. Autrement dit, dans de très nombreux cas, l'action moralement requise sera aussi l'action qui sert les intérêts privés. En conséquence, *la*

fréquence à laquelle les individus devront sacrifier leurs intérêts personnels au nom de la morale ou de la justice sera particulièrement faible.

La seconde version de l'argument concerne moins la *fréquence* des occasions dans lesquelles un individu doit sacrifier ses intérêts propres que la *force* des exigences morales auxquelles un individu peut être soumis. Ainsi, selon cette seconde forme d'argument, seules sont acceptables – pour des raisons que nous verrons – les théories morales qui ne sont pas *trop exigeantes* à l'égard des êtres humains, c'est-à-dire les théories morales qui *ne demandent* pas des sacrifices trop grands aux individus et qui sont compatibles avec une vie humaine heureuse. Cette seconde version de l'argument de « dédramatisation » du recours à la motivation morale consiste donc à montrer que, peu importe la *fréquence* à laquelle des sacrifices sont exigés, ces sacrifices ne pourront de toute façon jamais être trop importants.

Nous nommerons le premier argument « l'argument de dédramatisation de l'égoïsme » et le second « l'argument humaniste », dont la fonction est de dédramatiser la force des exigences morales. Nous discuterons à tour de rôle de ces deux tentatives visant à *nuancer* et *amoindrir* les conséquences de la thèse selon laquelle une société ne peut pas fonctionner sans faire appel à la motivation morale des citoyens. Le premier argument sera l'objet du présent chapitre, alors que le second, plus fondamental, sera l'objet du chapitre suivant.

2. L'ARGUMENT DE « DÉDRAMATISATION DE L'ÉGOÏSME »

Nous avons vu précédemment que l'égoïsme seul ne permet pas de générer des actions conformes à la morale avec suffisamment de fiabilité et de régularité pour garantir le bon fonctionnement de la société. Il est nécessaire, pour chaque situation dans laquelle la (maximisation de la) satisfaction de l'intérêt privé est incompatible avec la réalisation du bien commun ou avec la préservation des institutions sociales, que les individus renoncent *intentionnellement* à leurs intérêts privés au nom du bien commun. Cette conclusion peut sembler négative, dans la mesure où elle montre que la société dépend, pour fonctionner, de la présence de dispositions morales chez ses membres, dispositions dont il est aisé de penser qu'elles sont faibles, mal réparties au sein de la population, voire qu'elles dépendent de capacités émotionnelles non soumises au contrôle de la raison. Autrement dit, en

comparaison avec la motivation puissante, rationnelle et largement partagée de l'intérêt personnel, la motivation morale apparaît comme une ressource motivationnelle très fragile pour garantir le fonctionnement de la société.

Pour les opposants au programme de « mécanisation de la vertu », il est donc important de nuancer le rôle de la motivation morale dans la société, en montrant que la morale n'exige pas un *renoncement total* au bonheur, à la richesse, au plaisir et au bien-être personnel. Pour y parvenir, une première méthode consiste à dédramatiser la *quantité* de situations de conflit entre l'égoïsme et la morale. Nous allons ainsi étudier des arguments qui soutiennent que les exigences liées à la prudence, à la satisfaction de l'intérêt personnel, et les exigences liées à la morale *se recoupent* la plupart du temps ; dans la majorité des cas, les actions moralement bonnes sont aussi les actions qui satisfont l'intérêt personnel. Selon cet argument, les exigences prudentielles et morales convergent même à tel point, surtout dans certaines circonstances institutionnelles, que l'égoïsme cesse de constituer un véritable *défi* pour la réalisation d'une société morale.

Le défi que pose l'égoïsme humain pour la vie en société a été formulé en philosophie sous l'appellation du *défi de l'amoraliste*.¹⁵⁰ Bien que nous aurions pu également l'intituler le *défi de l'égoïsme*, nous utiliserons l'appellation « amoraliste » pour correspondre à la formulation classique de la discussion.¹⁵¹ L'égoïste et l'amoraliste partagent la caractéristique essentielle de *ne pas se préoccuper de morale*, sauf si cela sert leurs intérêts personnels. De ce fait, pour motiver un amoraliste ou un égoïste à adopter un comportement moralement adéquat, il faut lui montrer qu'il s'agit de la meilleure stratégie pour promouvoir ses propres intérêts.

L'amoraliste est un individu qui admet être soumis à un certain nombre d'exigences morales, mais qui reste parfaitement indifférent à ces exigences.¹⁵² Autrement dit,

¹⁵⁰ Bernard Williams décrit ce défi dans les termes suivants: « Lorsque l'amoraliste met en doute les considérations morales, et qu'il suggère qu'il n'existe pas de raison de suivre les exigences morales, que pouvons-nous lui répondre ? » [“When the amoralist calls ethical considerations in doubt, and suggests that there is no reason to follow the requirements of morality, what can we say to him?” (WILLIAMS, 2006, p. 22)]

¹⁵¹ Bernard Williams aussi identifie l'égoïsme et l'amoralisme, comme le montre le passage suivant : « L'égoïsme [est] la position d'un être sans rapport avec la moralité, qui rejette [...] toutes considérations morales, ou ne s'y intéresse pas, ou leur résiste, et que seul son propre intérêt sollicite » (WILLIAMS, 1994, p. 157).

¹⁵² Pour bien comprendre la nature de ce « défi de l'amoraliste », il convient de distinguer la figure de l'amoraliste d'autres personnages ou figures logiques discutées dans la littérature philosophique (BAYERTZ, 2004, pp. 20-32). En premier lieu, contrairement au sceptique, l'amoraliste ne doute pas de la possibilité de

l'amoraliste ne voit pas pourquoi il devrait se soucier de la morale. Il n'est pas du tout motivé à agir par la considération du caractère louable ou condamnable d'une action. Il estime qu'il n'a pas de *raison* de se plier aux exigences morales. Cette figure (logique) de l'amoraliste pose un défi pour toute théorie de la société qui dépend, pour fonctionner, du fait que les individus accordent de l'importance aux considérations morales. Si l'écart entre la poursuite de l'intérêt personnel de l'amoraliste et le respect des exigences morales dont *il ne se soucie pas* est trop important, il faut s'attendre à ce que la vie en société devienne difficile. Il faut donc pouvoir montrer que l'écart entre la poursuite de l'intérêt personnel et le respect des exigences morales n'est pas trop important.

Nous allons étudier trois arguments en ce sens : (1) l'argument des coûts de la posture de l'amoraliste, (2) l'argument du passage en douceur de l'amoralisme à la morale et (3) l'entrée rapide dans la sphère de la moralité. Le premier argument consiste à montrer que, dans le monde réel, il est extrêmement coûteux d'adopter une posture amoraliste ; cela demande du temps, de l'énergie, et des prises de risques qui ne sont pas nécessairement raisonnables. Le second argument consiste à soutenir que, dans le monde réel, il n'existe pas d'individus *totale*ment égoïstes. Or, en admettant qu'il existe chez tout individu une quantité (même limitée) et circonscrite d'altruisme, il devient plus facile de comprendre comment la motivation morale est possible, comment il est possible que les individus réels accordent de l'importance également au bien de personnes éloignées. Le troisième argument a pour fonction de rendre moins *séduisante* la posture de l'amoraliste. En effet, en général, la posture de l'amoraliste est « attractive » parce qu'elle représente une posture de rejet de la morale ordinaire, considérée comme trompeuse et nuisible, au nom d'un autre idéal moral – i.e. la morale de la puissance individuelle, de la domination du puissant, etc. Or, le dernier argument montre que cette posture morale *supérieure* n'est pas disponible pour l'amoraliste. Cela le rend, de ce fait, beaucoup moins séduisant.

connaître ce que la morale exige de lui. S'il est réticent à agir moralement, cela ne provient pas de doutes concernant la possibilité de la connaissance morale. En second lieu, l'attitude réfractaire de l'amoraliste face à la morale ne provient pas d'une crainte d'être exploité par les autres. L'amoraliste ne craint pas que d'autres profitent de son honnêteté ou même qu'il soit ridiculisé pour son comportement moral [inutile]. Le défi que lance l'amoraliste vis-à-vis de la morale est valable même si toutes les autres personnes agissent moralement. Au dernier lieu, l'amoraliste se distingue également de l'immoraliste. Alors que l'amoraliste ne prend pas position concernant la morale, l'immoraliste se donne comme principe de systématiquement rejeter la morale, même dans les situations où elle concorderait avec son intérêt. L'immoraliste agit contre la morale par principe, par volonté de désobéissance à la morale. De son côté, l'amoraliste ne possède aucune aversion de principe à l'égard de la morale, mais se contentera de suivre les règles morales lorsque celles-ci participent à la promotion de son propre intérêt et de leur désobéir dans le cas contraire

2.1. LES COÛTS DE LA POSTURE DE L'AMORALISTE

Selon le premier argument, la posture de l'amoraliste est extrêmement coûteuse *dans la réalité* – ce que ne montre pas l'approche purement logique du problème.¹⁵³

En effet, dans le monde réel, adopter l'attitude d'un amoraliste peut même exiger de se mentir à soi-même, de se dissimuler à soi-même ses tendances égoïstes ou immorales, à un point tel que l'amoraliste ne se distingue plus de l'homme vertueux, même dans ses motivations. En d'autres termes, il est tellement souvent avantageux, dans le monde réel, pour l'amoraliste – en termes de coûts et bénéfices attendus – de préférer l'action morale à l'action égoïste, qu'il devient avantageux pour lui de développer des *dispositions morales* stables. L'amoraliste aurait, en quelque sorte, avantage à se transformer en un individu moral.

John Rawls, par exemple, développe un argument dans ce sens (RAWLS, 2009). Il distingue cinq manières différentes dont les circonstances *réelles* dans lesquelles se trouve l'amoraliste permettent de faire converger la morale et l'intérêt personnel : (i) les coûts psychologiques, (ii) l'opposition à nos dispositions affectives, (iii) l'impossibilité de maîtriser les effets de notre action, (iv) le risque de désintégration sociale et (v) notre disposition naturelle au bien. Ces cinq arguments représentent autant de raisons de considérer la position amoraliste comme intenable dans les conditions de vie réelles.

2.1.1. LES COÛTS PSYCHOLOGIQUES

En premier lieu, Rawls déploie l'argument bien connu selon lequel adopter une posture amoraliste, qui implique de toujours être à l'affût des occasions de profiter d'un « ticket gratuit », lorsque le risque d'être démasqué est suffisamment faible, génère un important coût psychologique.

Soit un individu qui, par hypothèse, [...] considère la possibilité de faire semblant d'avoir certains sentiments moraux, tout en étant prêt à user du « ticket gratuit » chaque fois qu'il a la possibilité d'améliorer ainsi ses propres intérêts. [...] Le fait que le mensonge et l'hypocrisie soient de mauvaises actions ne devrait pas, je

¹⁵³ Comme l'affirme Jean-Claude Wolf, « la différence entre l'égoïsme et l'universalisme – aussi intéressante qu'elle puisse être en théorie – est sans conséquence pour la pratique. » [„[D]er Unterschied zwischen Egoismus und Universalismus – so interessant er für die Theorie sein mag – bleibt praktisch folgenlos“ (WOLF, 2002, p. 95).]

pense, le troubler ; mais il aura à tenir compte du coût psychologique des précautions à prendre et du rôle à jouer, ainsi que du manque de spontanéité et de naturel qui en résulte (RAWLS, 2009 , p. 611).

Etre un parasite social exige une vigilance pour ne pas être repéré ainsi qu'une forme d'hypocrisie à l'égard d'autrui, puisque nous devons publiquement nous déclarer comme favorable à la morale pour ne pas subir l'exclusion d'un système avantageux, tout en maintenant une attitude secrètement parasitaire. Or, cette vigilance et cette hypocrisie peuvent s'avérer psychologiquement fatigantes.

Dans des termes similaires, Jean-Claude Wolf et Kurt Bayertz soutiennent qu'une stratégie de parasitage *dans la réalité* exigerait une forme de « schizophrénie » très coûteuse, puisqu'il faudrait publiquement se déclarer attaché à la morale – puisque nous avons besoin que les autres se conforment aux règles – tout en réussissant à dissimuler notre propre égoïsme. Une telle séparation psychologique est difficile à entretenir sur le long terme.

Pour des raisons psychologiques, une stratégie amoraliste cohérente ne peut que difficilement être maintenue de façon durable, du fait qu'elle requiert une forme de double vie et implique une sorte de schizophrénie.¹⁵⁴

Si nous adoptons publiquement une stratégie égoïste, i.e. nous nous déclarons publiquement comme étant égoïstes, alors il faut s'attendre à ce que les autres personnes adoptent également une attitude égoïste à notre égard, ce qui est coûteux pour nous.¹⁵⁵ Si nous choisissons alors plutôt d'adopter une stratégie égoïste « raffinée » ou « secrète », i.e. nous nous déclarons publiquement comme soucieux de morale, tout en agissant secrètement comme un égoïste, alors nous serons confrontés à une forme psychologiquement coûteuse d'incertitude et d'hypocrisie.

¹⁵⁴ „Ein konsequenter strategischer Amoralismus dürfte aus psychologischen Gründen nur schwer dauerhaft durchzuhalten sein, da er eine Art von Doppelleben erfordert und eine Art von Schizophrenie impliziert“ (BAYERTZ, 2004, p. 254).

¹⁵⁵ Cette discussion rappelle l'argument de Hobbes dans le chapitre 15 du *Leviathan* (HOBBS, 1991 (1651)) à l'encontre du « fou » qui déclare *publiquement* ou *montre explicitement* qu'il ne se soucie pas de morale. Il existe sur ce point une discussion pour savoir si Hobbes attaque l'amoraliste *explicite* ou l'amoraliste *secret* (HOEKSTRA, 1997).

Un égoïste peu subtil se heurte en permanence à des concurrents et à des critiques. Il ne peut pas empêcher que les autres se considèrent eux-mêmes comme particulièrement importants et précieux et qu'ils revendiquent le même droit à la maximisation de leurs propres avantages. [...] A l'inverse, l'égoïste secret et subtil mise sur la tromperie constante des personnes qu'il peut le plus efficacement utiliser comme ressource pour la maximisation de ses propres avantages, aussi longtemps qu'elles ne le remarquent pas. Les deux variantes [de l'égoïsme] s'accompagnent d'une insécurité durable, à savoir l'insécurité de constamment devoir réfléchir si l'on peut ou non vaincre ou tromper les autres. [...] L'égoïste se trouverait dans une situation plus stable s'il misait sur une morale générale partagée, qui garantit les avantages de la coopération. [Mais] il fait toujours face au dilemme, soit de menacer ou de tromper le plus fréquemment et intelligemment possible, dans le but de vaincre ou de parasiter sans être vu, soit de vivre le plus agréablement possible et d'adapter son caractère et ses décisions personnelles à la morale du sens commun.¹⁵⁶

En ce sens, suivre les règles de la vie commune sans, à chaque occasion, calculer l'intérêt de « parasiter » la société permet d'éviter les coûts psychologiques liés au calcul, à l'incertitude et au manque de naturel de la posture amoraliste. Il semble donc clairement avantageux, pour l'amoraliste, de devenir un homme vertueux.

2.1.2. L'OPPOSITION À NOS DISPOSITIONS AFFECTIVES

Une seconde manière d'après laquelle la posture amoraliste génère dans la réalité des coûts pour la personne provient du fait que nous recherchons naturellement, dans nos vies, à établir des relations affectives avec autrui et que, sans raison contraire, nous préférons entretenir des liens positifs plutôt que des liens négatifs avec nos semblables. Pour préserver ces liens affectifs, auxquels il est attaché, l'amoraliste doit ne pas déclarer publiquement sa

¹⁵⁶ „Der rüde Egoist prallt permanent mit Konkurrenten und Kritiken zusammen. Er kann nicht verhindern, dass sich andere ebenfalls als besonders wichtig und wertvoll betrachten und auf das gleiche Recht zu Maximierung des eigenen Vorteils pochen. [...] Das heimliche und raffinierte Egoist dagegen setzt auf eine permanente Täuschung all derer, die man am wirksamsten als Ressourcen für die Maximierung des eigenen Vorteils benutzen kann, so lange sie es nicht merken. Beide Varianten [...] sind von einer dauernden Unsicherheit begleitet, nämlich von der Unsicherheit, ob man anderen wirklich immer und derart überlegen ist, dass man sie ständig besiegen oder überlisten kann. [...] Der Egoist wird dagegen stabiler, wenn er auf eine gemeinsame geteilte Moral setzt, welche Kooperationsvorteile gewährleistet. [Aber] er steht wiederum vor dem Dilemma, entweder möglichst häufig und geschickt zu drohen oder zu betrügen, um zu siegen oder um unbemerkt zu parasitieren, oder möglichst bequem zu leben und sich der Moral des common-sense auch in seinen inneren Entschlüssen und seinem Charakter anzupassen“ (WOLF, 2002, p. 101).

position. Par exemple, si l'amoraliste accorde de la valeur à l'amitié, il doit, au minimum, renoncer à son égoïsme dans ses relations avec ses proches.

L'égoïste fait face au dilemme suivant: soit perdre ses amis, soit accepter d'affaiblir son égoïsme à l'égard de certains amis.¹⁵⁷

Cet argument peut être encore étendu : selon Rawls, lorsqu'une société dispose des institutions adéquates, nous permettant d'entretenir des relations positives avec nos concitoyens, cela convient bien à notre nature et nous n'avons pas de raisons de nous y opposer. Nous développons même une forme d'attachement à ces institutions qui est similaire à l'attachement que nous entretenons à l'égard de nos amis.

[I]l existe une relation entre l'action juste et les attitudes naturelles (§74). [...] En supposant que l'on ait besoin de ces liens affectifs, l'attitude rationnelle à prendre sera probablement d'agir justement à l'égard de ceux auxquels nous sommes liés par l'affection et la sympathie et de respecter les modes de vie auxquels nous sommes dévoués. Mais dans une société bien ordonnée, ces relations s'étendent loin et comprennent les liens avec des formes institutionnelles (RAWLS, 2009 , p. 611).

Autrement dit, dans une structure institutionnelle adéquate, le respect des règles correspond mieux, selon Rawls, aux dispositions affectives de l'être humain que ne le fait la posture amoraliste.

2.1.3. L'IMPOSSIBILITÉ DE MAÎTRISER LES EFFETS DE NOTRE ACTION

En troisième lieu, Rawls fait remarquer que nous ne pouvons pas décider, dans une société complexe, quelles seront les personnes affectées par nos actions parasites. Or, sauf à supposer, à nouveau, que notre amoraliste réel ne possède aucune relation sociale à laquelle il accorde de la valeur – ce qui rendrait sa position sérieusement moins attractive – un amoraliste prend le risque que ses actions parasites affectent négativement la vie de ses proches.

¹⁵⁷ „Der Egoist steht vor dem Dilemma, entweder Freunde zu verlieren oder um einiger Freunde willen seinen Egoismus abschwächen zu müssen.“ (WOLF, 2002, p. 91).

Nous ne pouvons pas choisir d'une manière générale qui aura à souffrir de notre injustice (RAWLS, 2009 , p. 612).

Par exemple, une personne qui décide de pratiquer l'évasion fiscale ne peut pas choisir quelles seront les prestations étatiques qui seront affectées par son action, ni choisir si une baisse générale du financement public affectera sa propre vie ou celle de ses proches, négativement ou positivement. Il est donc risqué, même en termes d'intérêts personnels, d'adopter une stratégie d'action parasitaire.

2.1.4. LE RISQUE DE DÉSINTÉGRATION SOCIALE

La quatrième manière dont l'amoraliste ne bénéficie pas de sa propre action parasitaire provient du fait que la vie en société présente pour lui de grands avantages. Une mise en danger du bon fonctionnement de la société, issue par exemple d'une généralisation de l'action parasitaire, lui serait extrêmement coûteuse. Un certain nombre de possibilités, qui ne sont réalisables que par la vie en société, se fermerait à lui.

[En outre,] la participation à la vie d'une société bien ordonnée est elle-même un grand bien (§79). [...] [N]ous dépendons de la coopération des autres non seulement pour notre bien-être et les moyens d'y accéder, mais aussi pour amener à maturité nos possibilités latentes. Et, si nous avons un certain succès, chacun reçoit quelque chose de la richesse et de la diversité de l'activité collective. Mais, pour participer pleinement à cette vie, nous devons accepter les principes qui commandent sa conception, ce qui veut dire que nous devons respecter notre sens de la justice (RAWLS, 2009 , p. 612).

Une manière raisonnable de garantir l'accessibilité aux avantages individuels liés à la vie en société consiste alors à simplement accepter les règles qui régissent la vie commune et à ne pas adopter une attitude parasitaire risquant de la mettre en péril.

2.1.5. NOTRE DISPOSITION NATURELLE AU BIEN

Finalement, Rawls soutient que nous disposons d'une tendance naturelle au bien à laquelle nous devrions résister pour être capables d'adopter l'attitude parasitaire de l'amoraliste. Cette tendance naturelle au bien provient de notre conception fondamentale

de nous-mêmes comme étant des agents libres et y renoncer impliquerait de renoncer à une part importante de notre propre identité.

Pour finir, il y a une dernière raison qui est liée à l'interprétation kantienne : agir justement est quelque chose que nous voulons en tant qu'êtres rationnels, libres et égaux (§40). Le désir d'agir justement et celui d'exprimer notre nature de personnes morales libres s'avèrent pratiquement le même désir (RAWLS, 2009 , pp. 612-613).

Or, une vie qui requiert de s'opposer à nos tendances morales naturelles est moins satisfaisante, du point de vue du bien-être individuel, qu'une vie qui réalise ces tendances. Une vie morale est donc meilleure qu'une vie amoraliste parasitaire en termes de bien-être individuel.

Ainsi, Rawls soutient qu'il est désavantageux d'adopter une posture amoraliste (i) en termes de coûts psychologiques, (ii) en termes de contradiction avec nos affections sociales naturelles, (iii) en termes de risques pour nos proches et pour nous-mêmes, (iv) en termes des risques de perdre les avantages liés à la vie sociale et (v) en termes d'opposition à notre tendance naturelle au bien.

2.2. CRITIQUE DE L'ARGUMENT DES COÛTS DE L'AMORALISME

Il est possible cependant d'émettre des réticences à l'encontre de cette première démarche de « dédramatisation par les coûts » de l'amoralisme. En effet, même s'il est vrai que la posture amoraliste, dans la réalité, génère un certain nombre de coûts, il est possible de soutenir que lesdits coûts ne sont pas aussi nombreux que ce que prétendent les arguments précédents. Ainsi, par exemple, Bernard Williams offre trois raisons de modérer la prétention (rawlsienne) de prouver que le parasitage social est désavantageux.

En premier lieu, les avantages ou désavantages liés à l'attitude égoïste – ou amoraliste – dépendent essentiellement des circonstances. Il est donc difficile de donner une portée vraiment *systematique* à l'argument visant à dédramatiser le problème posé par l'amoralisme.

Les déclarations au sujet de ce dont l'égoïste se prive ont toutes les chances d'avoir un caractère de plus en plus empirique (WILLIAMS, 1994, p. 167).

Par exemple, la convergence de la morale et de l'intérêt s'opère chez Rawls dans le cadre d'une « société bien ordonnée » et non pas dans n'importe quel cadre social. Or, présupposer *contre l'amoraliste* l'existence d'une « société bien ordonnée » revient à présupposer une partie au moins de ce qui doit être prouvé, à savoir que la vie en société est meilleure pour l'amoraliste.

En second lieu, l'argument au sujet des coûts de l'égoïsme ne porte pas sur les désirs que l'amoraliste possède *déjà* mais tente plutôt de faire apparaître chez lui de nouveaux désirs – en lui montrant, par exemple, que l'amitié est désirable. Ainsi,

[...] il y a peu de chances qu'on doive se contenter de lui faire remarquer qu'il a des désirs insatisfaits [...] – il s'agira plutôt de l'exciter ou de le convertir à une nouvelle idée de ce qui peut présenter de l'intérêt (WILLIAMS, 1994, p. 167).

En ce sens, l'argument ne porte pas sur ce que l'amoraliste *veut déjà* mais cherche à convaincre le sujet de *vouloir de nouvelles choses* – notamment l'amitié, le respect des autres, etc. Or, la force de l'argument contre l'amoralisme diminue dans ce contexte.

Finalement, le fait que l'égoïste doive subir certains coûts lorsqu'il décide de parasiter la société ne doit pas faire oublier que la morale également engendre certains coûts pour l'égoïste.

[S]'il est sans aucun doute vrai que l'égoïste [...] se priverait des petites satisfactions de la vie quotidienne, ceci ne devrait pas être exploité avec trop de complaisance polémique par les défenseurs de la morale et de l'altruisme, car pour de nombreux individus, acquérir et conserver les dispositions morales exigées d'eux entraîne la perte de bien des satisfactions, non seulement en général, [...] mais encore dans ces domaines de la relation interpersonnelle que nous avons considérés comme perdus pour l'égoïste tout particulièrement (WILLIAMS, 1994, pp. 167-168).

Par exemple, respecter l'exigence de « dire la vérité » à ses amis peut, en certaines circonstances, mettre en danger la relation amicale elle-même. Autrement dit, à trop souligner les coûts de l'amoralisme, nous risquons d'oublier les coûts liés à l'acquisition des dispositions morales – et ce sont pourtant précisément ces coûts qui ont réclamé une

procédure de dédramatisation. En conséquence, les coûts associés à la morale ne doivent disparaître de vue dans l'argument.

En résumé, même s'il est vrai qu'il ne faut pas dramatiser l'écart qui sépare l'intérêt personnel de la morale, il ne faut pas non plus le sous-estimer. L'amoralisme reste une position *intéressante* malgré tout.¹⁵⁸

2.3. LE PASSAGE EN DOUCEUR DE L'AMORALISME À LA MORALE

Nous avons vu un premier argument dont le but est de dédramatiser le danger que pose l'amoralisme ou l'égoïsme pour la vie en société et la réalisation d'une société morale. Cet argument consistait à identifier des coûts liés à l'adoption d'une posture amoraliste *dans la réalité*.

Le second argument de dédramatisation de l'amoralisme peut se décrire ainsi : dans la réalité, nous ne trouvons pas de *véritables* égoïstes. Les hommes *réels* disposent également de motivations non égoïstes – bien qu'elles soient plus faibles que l'intérêt personnel.

Les êtres humains réels sont rarement des Saints, qui cherchent à faire le bien et seulement le bien; ils ne sont aussi que rarement de véritables égoïstes, qui sont rigide­ment concentrés sur leurs propres intérêts. Si l'on fait abstraction des psychopathes, nous pouvons supposer que les êtres humains ont des motivations mixtes; certains pencheront plus du côté de l'altruisme, d'autres plutôt du côté de l'égoïste. L'amoraliste serait alors une personne qui possède de très fortes motivations égoïstes, et de très faibles motivations altruistes. Mais même cette personne ne se situerait pas complètement du côté de l'égoïsme, parce qu'un être sans aucun sentiment moral n'est pas un homme [...] Les êtres humains réels agissent selon une variété de motivations, dont des motivations qui portent sur le bien-être d'*autres* êtres humains.¹⁵⁹

¹⁵⁸ Cela est d'autant plus le cas dans les problèmes d'action collective actuels, et notamment au problème du changement climatique, où les effets indésirables du parasitage ne sont pas visibles, où il est même difficile de désigner un coupable, et où la nature de la faute morale elle-même n'est pas évidente (CHAUVIER, 2013).

¹⁵⁹ „Reale Menschen sind zwar selten Heilige, die Ratlos darauf sinnen, Gutes und nur Gutes zu tun; sie sind aber ebenso selten reine Egoisten, die starr auf ihr eigenes Interesse fixiert sind. Wenn wir von Psychopathen absehen, können wir realen Menschen also gemischte Handlungsmotive unterstellen; die einen mehr auf der Seite des Altruismus, die anderen mehr auf der Seite der Egoismus. Der Amoralist wäre dann eine

L'argument montre alors que la transformation d'un être humain *réel* en un être moral est moins problématique que la transformation d'un amoraliste *pur* en un être moral. Dans la réalité – par opposition à la théorie – les dispositions morales ne représentent rien de véritablement mystérieux : elles ne consistent qu'en *une extension* de motivations déjà existantes. Autrement dit, les êtres humains possèdent déjà les dispositions nécessaires à la vie morale. Il suffit d'étendre ces dispositions, afin qu'elles intègrent un ensemble plus vaste de personnes, pour obtenir le comportement moral adéquat. Il n'existe ainsi pas de *fossé infranchissable* entre l'individu réel et l'individu vertueux, mais une forme de *gradation* qui autorise tout à fait que la morale ne soit pas entièrement opposée aux désirs de l'individu réel – même si, chez ce dernier, les motivations égoïstes restent dominantes. Comme le montre Bernard Williams, le fait que quelqu'un soit disposé à *se préoccuper* du bien-être d'autrui, même s'il ne s'agit que de personnes proches auxquelles il est attaché, suffit pour lui attribuer déjà les germes d'une disposition morale.

Nous pouvons décrire [l'amoraliste] comme ayant certaines affections, comme se souciant occasionnellement de ce qui arrive à quelqu'un d'autre. Quelque stéréotype de films de gangster vient à l'esprit, celui d'un personnage impitoyable et plutôt glamour, qui se soucie de sa mère, de ses enfants et même de sa maîtresse [...]. Un tel personnage nous apporte, je pense, suffisamment, puisqu'il possède la notion de faire quelque chose *pour quelqu'un d'autre*, parce que cette personne a besoin de quelque chose. [...] [C]et homme est capable de penser en termes des intérêts des autres, et son incapacité à être un agent moral provient (partiellement) du fait qu'il est disposé à agir de cette manière seulement de manière intermittente et capricieuse. Mais il n'existe pas de fossé sans fond entre son état et les dispositions de base de la morale.¹⁶⁰

Person, die sehr starke egoistische und sehr schwache altruistische Motive hat. Aber auch sie wäre nicht ganz auf der egoistische Seite, denn ohne alles moralisches Gefühl ist kein Mensch [...] Reale Menschen handeln aus einer Vielzahl von Beweggründen, darunter auch aus solchen, die auf das Wohl anderer Menschen gerichtet sind.“ (BAYERTZ, 2004, pp. 203-204).

¹⁶⁰“One might picture [the amoralist] as having some affections, occasionally caring for what happens to somebody else. Some stereotypes from a gangster movie come to mind, of the ruthless and rather glamorous figure who cares about his mother, his child, even his mistress. [...] His gives us, I think, almost enough. For he has the notion of doing something for somebody, because that person needs something. [...] [T]his man is capable of thinking in terms of others' interests, and his failure to be a moral agent lies (partly) in the fact that he is only intermittently and capriciously disposed to do so. But there is no bottomless gulf between this state and the basic dispositions to morality.” (WILLIAMS, 1972, pp. 10-11).

Même s'il manque encore à un tel individu la capacité d'agir moralement de manière plus régulière et étendue, il possède néanmoins déjà, du seul fait qu'il est capable *de se soucier de quelqu'un d'autre* que lui-même, d'une disposition à la morale.¹⁶¹

Bien entendu, nous pourrions imaginer un amoraliste qui ne s'inquiète vraiment de personne d'autre que de lui-même. Cependant, comme le souligne Bernard Williams, un tel individu non seulement ne correspond pas aux individus réels, mais il correspond plus à un rebutant psychopathe qu'à un être dont la vie pourrait nous offrir quelque *tentation* de renoncer à la vie morale.

Est-ce qu'il se soucie de qui que ce soit? Est-ce qu'il existe une personne dont la souffrance et le désarroi l'affecterait ? Si nous répondons « non », il semble que nous ayons produit un psychopathe. S'il est un psychopathe, l'idée de discuter et d'argumenter avec lui est certainement idiote. [...] L'amoraliste semblait important parce qu'il semblait offrir une alternative ; sa vie, après tout, paraissait être intéressante. Le psychopathe peut être important pour la pensée morale ; mais son importance provient du fait qu'il nous épouvante, et que nous devons chercher une compréhension plus profonde des raisons pour lesquelles le psychopathe nous épouvante. Son importance ne provient pas du fait qu'il nous offre une forme de vie alternative attrayante.¹⁶²

Pour que la figure logique de l'amoraliste présente véritablement un *défi* pour la possibilité d'une société morale, il doit présenter quelques caractéristiques *attractives*. La figure d'un psychopathe, incapable d'entretenir des relations sociales *quelconques*, ne représente pas le cas de figure adéquat. Pour que l'amoralisme représente un véritable défi pour la morale, il faut donc attribuer à notre individu la capacité d'entretenir des relations sociales et de se soucier d'autrui. Ce qui revient à *réduire* le fossé censé le séparer de la sphère de la moralité.

¹⁶¹ Nous avons vu dans la première partie que, la motivation morale et l'altruisme, bien que différentes, partagent néanmoins la caractéristique d'être des motivations désintéressées.

¹⁶² "Does he care about anybody? Is there anybody whose sufferings and distress would affect him? If we say 'no' to this, it looks as though we have produced a psychopath. If he is a psychopath, the idea of arguing him into morality is surely idiotic. [...] The amoralist seemed important because he seemed to provide an alternative; his life, after all, seemed to have its attractions. The psychopath is, in a certain way, important to moral thought; but his importance lies in the fact that he appals us, and we must seek some deeper understanding of how and why he appals us. His important does not lie in his having an appeal as an alternative form of life." (WILLIAMS, 1972, pp. 9-10).

2.4. L'ENTRÉE RAPIDE DANS LA SPHÈRE DE LA MORALE

Une autre manière de démontrer que le fossé séparant l'amoraliste de la morale est plus étroit que ne le suggère la présentation purement logique de la situation consiste à pointer la rapidité avec laquelle l'amoraliste peut se retrouver piégé dans le discours moral que pourtant il rejette. En effet, comme nous le verrons, en rejetant complètement le discours moral, l'amoraliste est forcé de renoncer à certaines attitudes de la vie ordinaire. Or, comme il est coûteux et difficile de renoncer à ces attitudes, l'amoraliste se retrouve rapidement à faire usage du discours moral que pourtant il critique. Bernard Williams offre à nouveau une analyse intéressante des situations concernées.

2.4.1. L'IMPOSSIBILITÉ DE LA JUSTIFICATION MORALE DE LA STRATÉGIE ÉGOÏSTE

En premier lieu, l'amoraliste peut considérer qu'il est *acceptable* pour lui et ses congénères de ne poursuivre que leurs intérêts privés. Cependant, il ne peut affirmer cela que dans la mesure où il n'attribue pas une signification *morale* ou *évaluative* à son propos. Ainsi, si « c'est acceptable » veut dire « je ne porte pas de jugement à ce sujet », alors un amoraliste conséquent peut affirmer cette proposition. Par contre, si « c'est acceptable » signifie « il est *légitime* que chacun poursuive son intérêt propre », alors l'amoraliste ne peut pas l'affirmer ; il est déjà rentré dans la sphère de la moralité.

En un certain sens, il est possible pour un homme de penser qu'il est "acceptable" de se comporter de manière intéressée ou égoïste, sans entrer dans le territoire distinctivement moral de la pensée ; cela est possible si, approximativement, « c'est acceptable » signifie « je ne vais pas moraliser à ce sujet ». Il entrera par contre dans le territoire proprement moral si « c'est acceptable » signifie quelque chose comme « c'est permis », puisque cette affirmation implique que « les personnes ne doivent pas interférer avec la poursuite des intérêts privés des autres personnes », et il ne s'agit pas d'une pensée qu'un amoraliste peut avoir.¹⁶³

¹⁶³“In one way, it is possible for a man to think it ‘all right’ for everyone to behave self-interestedly, without his having got into any distinctively moral territory of thought at all; if, roughly, ‘it’s all right’ means ‘I am not going to moralize about it’. He will be in some moral territory if ‘all right’ means something like ‘permitted’, for that would carry implications such as ‘people ought not to interfere with other people’s pursuit of their own interest, and this is not a thought which, as an amoralist, he can have.” (WILLIAMS, 1972, p. 5).

L'amoraliste ne peut donc pas défendre un amoralisme *moral*. Il ne peut pas considérer que l'égoïsme représente une position *morale* légitime. Cette attitude lui est interdite du fait qu'il incarne précisément le *rejet* de toute morale, de toute attitude moralisatrice.

2.4.2. L'IMPOSSIBILITÉ DE LA RÉVOLTE MORALE

En second lieu, l'amoraliste n'a pas à sa disposition un type comportement qui pourtant joue un rôle primordial dans les relations interpersonnelles : la possibilité de se *plaindre* d'un mauvais comportement à son égard, d'éprouver du ressentiment ou de désapprouver un comportement. Ces attitudes étant des attitudes proprement *morales*, elles ne sont pas accessibles à notre amoraliste.

De façon similaire, s'il s'oppose (comme il le fera certainement) au fait que les autres le traitent de la manière dont ils le traitent, l'amoraliste sera parfaitement cohérent aussi longtemps que son opposition consistera uniquement à ne pas apprécier et à frapper en retour. Ce que l'amoraliste ne peut pas faire de façon cohérente, c'est d'éprouver du ressentiment, ou de *désapprouver* le traitement qu'il subit, car il s'agit d'attitudes internes au système moral.¹⁶⁴

Les seules attitudes dont dispose l'amoraliste pour répondre à un mauvais traitement à son égard consistent à ne pas apprécier – à trouver, par exemple, que cela est *désagréable* – et à essayer de se sortir de cette situation désagréable, possiblement en *attaquant* à son tour. Mais il ne peut en aucun cas considérer que la situation est *injuste* et éprouver à l'égard du traitement infligé un sentiment moral tel que le *ressentiment*.¹⁶⁵

¹⁶⁴ "Similarly, if he objects (as he no doubt will) to other people treating him as he treats them, this will be perfectly consistent so long as his objecting consists just in such things as his not liking and fighting back. What he cannot consistently do is resent it or disapprove of it, for these are attitudes within the moral system." (WILLIAMS, 1972, p. 5).

¹⁶⁵ Nous trouvons un argument similaire chez Samuel Scheffler: « Considérons [...] la variété d'émotion et d'attitudes humaines puissantes [...] dont ne pourrait pas faire l'expérience une personne qui ne dispose pas de telles croyances. La culpabilité, le remords, l'indignation, le ressentiment, la conscience et un sens de dette, semblent tous, par exemple, tomber dans cette catégorie. Ainsi, [...] il est possible de ressentir de la colère lorsque l'on est maltraité, sans avoir une quelconque croyance morale, mais on ne peut pas éprouver du ressentiment si l'on ne croit pas que le mauvais traitement était immoral ou injustifié ou injuste. Et quelqu'un peut se sentir bien disposé à l'égard d'un bienfaiteur, sans avoir de croyance morale, mais il ne peut pas se sentir redevable s'il ne croit pas que l'on doit quelque chose en échange à ceux qui nous ont bien traité. » ["Consider [...] the range of powerful human emotions and attitudes [...] that could not be experienced by someone who had no such beliefs. Guilt, remorse, indignation, resentment, consciousness and a sense of indebtedness, all seem to fall into this category, for example. Thus [...] one can feel angry at being ill-treated

2.4.3. L'IMPOSSIBLE SUPÉRIORITÉ MORALE DE L'AMORALISTE

Finalement, l'amoraliste ne doit pas se montrer trop *fier* de son caractère. Il ne peut pas particulièrement *louer* son propre caractère – par rapport à d'autres types de caractères – car il risque assez rapidement de devoir faire usage de notions morales. Dans les termes de Bernard Williams:

De nouveau, l'amoraliste doit résister, s'il est cohérent, à la tendance plus insidieuse de penser de lui-même qu'il possède un caractère plutôt splendide – en particulier, du fait d'être comparativement plus courageux que la multitude lâche. Parce qu'en adoptant de telles pensées, il risque constamment de sortir du monde de ses propres désirs et goûts pour entrer dans une région où certaines dispositions sont considérées comme particulièrement excellentes pour les êtres humains, ou comme utiles à la société, ou quelque chose de ce genre ; et alors que de telles pensées ne mènent pas immédiatement à des considérations morales, elles leur offrent un fondement solide, puisqu'elles invitent immédiatement des questions au sujet de ce qui est bon au sujet de ces dispositions, et il sera difficile pour l'amoraliste de continuer la discussion de ces questions sans très rapidement penser en termes d'intérêts généraux et en termes des besoins des autres êtres humains, ce qui l'amènerait à nouveau dans le monde moral duquel il tente de s'exclure lui-même.¹⁶⁶

Un amoraliste ne peut se considérer comme une figure humaine *supérieure* que s'il peut montrer en quoi son positionnement est particulièrement meilleur que celui des autres personnes qui, contrairement à lui, sont fidèles à la morale ordinaire. Or, d'après l'argument de Williams, une telle défense amènera rapidement l'amoraliste à justifier son attitude égoïste sur la base de son effet positif sur le bien-être d'autrui ou sur l'intérêt général. Il

without having any moral beliefs, but one cannot resent the ill treatment unless one believes that it was wrong or unjustified or unfair. And one can feel kindly disposed toward a benefactor without having any moral beliefs, but one cannot feel indebted unless one believes that one owes something to those who have treated one well." (SCHEFFLER, 1992, p. 68)].

¹⁶⁶ "Again, he must resist, if consistent, a more insidious tendency to think of himself as being in character really rather splendid – in particular, as being by comparison with the craven multitude notably courageous. For in entertaining such thought, he will run a constant danger of getting outside the world of his own desires and tastes into the region in which certain dispositions are regarded as excellent for human beings to have, or good to have in society, or such things; and while such thoughts need not lead directly to moral considerations, they give a substantial footing to them, since they immediately invite questions about what is so good about those dispositions, and it will be difficult for him to pursue those questions very far without thinking in terms of the general interests and needs of his fellow human beings, which would land him once more back into the world of moral thought from which he is excluding himself." (WILLIAMS, 1972, pp. 6-7).

tomberait donc rapidement dans la sphère de la moralité. En conséquence, l'amoraliste cohérent, qui doit éviter tout engagement moral, doit plutôt adopter une attitude modeste à l'égard de sa posture et ne pas la considérer comme supérieure à une autre.

En résumé, l'amoraliste (i) ne peut pas soutenir la thèse de l'amoralisme ou de l'égoïsme à l'aide d'arguments moraux, (ii) ne peut pas éprouver du ressentiment ou de l'indignation et (iii) ne peut pas se considérer comme une figure moralement supérieure à celle de l'homme vertueux. Toutes ces attitudes réclament une prise de position morale que l'amoraliste rejette. Nous voyons donc que l'amoralisme est une position assez exigeante. L'essentiel est que les attitudes auxquelles l'amoraliste doit renoncer sont justement les attitudes qui, de prime abord, tendent à rendre la posture amoraliste attractive.

3. CONCLUSION

Nous avons dit que, en réponse à la nécessité de s'appuyer sur la motivation morale des citoyens, deux stratégies pouvaient être déployées pour dédramatiser l'écart entre l'intérêt personnel et la morale. La première stratégie, qui a été l'objet du présent chapitre, consiste à montrer que, dans le monde réel, la stratégie de l'amoraliste est moins séduisante que ce que suggère la théorie. Nous avons présenté trois arguments allant dans ce sens : (1) les coûts de la posture amoraliste, (2) le passage en douceur de l'amoralisme à la morale et (3) l'entrée rapide dans la sphère de la morale.

Nous considérons cependant que ces arguments ignorent le fait que l'amoralisme reste une stratégie très séduisante dans les conditions de la société contemporaine.

En premier lieu, comme le souligne Kurt Bayertz (2004), il existe des situations dans lesquelles il n'est pas possible d'attribuer l'échec moral à des individus particuliers. Il n'est pas possible *d'identifier* les responsables. De ce fait, il n'est pas possible de *sanctionner* des individus *particuliers* pour les atteintes au bien commun. Or, il existe de nombreuses situations dans lesquelles cette condition *d'identification* n'est pas satisfaite. C'est le cas, notamment, dans les interactions qui mettent en relation un grand nombre de personnes. Dans ces circonstances, il devient difficile d'identifier la stratégie de chaque personne, et donc d'identifier et sanctionner les parasites. C'est le cas également dans les situations où l'attribution de la responsabilité morale à des individus particuliers n'est pas évidente, du

fait qu'il est difficile de déterminer la nature de la faute morale. Par exemple, l'attribution de la responsabilité dans le changement climatique pose des difficultés, non seulement à cause du nombre de personnes concernées et de l'effet de seuil, mais aussi à cause de la difficulté que cela pose d'identifier la *nature* de la faute morale commise (CHAUVIER, 2013).

En second lieu, il existe des situations dans lesquelles il n'est pas possible de *répondre* à comportement immoral. En effet, il est possible que la relation entre la personne subissant le dommage et celle causant le dommage ne soit pas réciproque. En conséquence, la victime des actes immoraux ne dispose pas des moyens de répondre et de sanctionner. Par exemple, cette condition de *réciprocité* n'est pas satisfaite en cas de forte inégalité. Il peut s'agir de relations où la *réciprocité* est rendue impossible par la *nature* des individus affectés : par exemple, il existe une situation d'inégalité existentielle entre un adulte et un nourrisson ou une personne handicapée, entre une personne de la génération présente et une personne des générations futures, ou encore entre des êtres humains et d'autres êtres sensibles. Il peut également s'agir de situations dans lesquelles des relations inégalitaires de *pouvoir* empêchent l'une des parties prenantes de véritablement *sanctionner* un comportement non coopératif de l'adversaire. Les relations entre employeurs et employés sont par exemple parfois structurées de telle manière que la capacité de l'employé de réagir aux manquements de son employeur est fortement réduite.

Pour finir, le programme de dédramatisation de l'amoralisme nous paraît sous-estimer la capacité et l'énergie que les individus sont susceptibles de déployer pour éviter d'accomplir leur devoir moral, voire pour commettre des actions immorales. Les stratégies de fraude ou d'optimisation fiscale, ou le cas des moteurs diesel trafiqués, ne constituent que des exemples parmi d'autres de mise en place de stratégies très élaborées pour éviter de respecter des normes morales. De plus, il est probable que les personnes qui ont la possibilité de mettre en œuvre des stratégies amoralistes efficaces ne soient pas des personnes particulièrement sujettes aux risques d'être elles-mêmes affectées négativement. En conséquence, le fait, par exemple, qu'elles ne puissent pas faire un usage légitime de notions *morales*, ni pour légitimer leur comportement, ni pour se défendre d'un mauvais traitement, les laissera certainement indifférentes ; leur statut économique et social les protègent des nuisances dont elles pourraient avoir envie de se plaindre.

En tous les cas, selon nous, les arguments visant à réduire l'écart entre la poursuite des intérêts privés et l'adoption d'une attitude morale reposent sur des éléments contingents qui limitent leur portée. Ces arguments sont, en effet, susceptibles d'être invalidés par tout changement de circonstances.

L'ERREUR MORALISTE

Nous avons traité, dans le chapitre précédent, un premier type d'arguments visant à dédramatiser le conflit entre l'intérêt personnel et la morale. Ces arguments visaient à montrer que la *fréquence* à laquelle les citoyens devraient renoncer à leurs intérêts privés était, dans la réalité, relativement faible.

Il existe une autre manière de soutenir que la conclusion du premier chapitre, selon laquelle nous ne pouvons pas nous dispenser de la motivation morale des personnes, n'est pas aussi radicale qu'il ne semble de prime abord. En effet, il existe une bonne raison de penser que les exigences morales, quelle que soit leur nature par ailleurs, ne peuvent pas s'opposer *complètement* aux aspirations humaines, à la recherche individuelle du bonheur ou du bien-être, à l'accomplissement d'une vie individuelle épanouie. Autrement dit, il existe une bonne raison de penser qu'une vie moralement vertueuse est compatible avec une vie individuelle heureuse et épanouie.

Cette raison est fournie par un argument métaéthique, formulé à l'encontre des doctrines morales qui exigent un type de vie incompatible avec les aspirations humaines fondamentales. Nous nommons cet argument « l'argument de l'erreur moraliste ». Il constitue l'objet d'étude du présent chapitre.

1. L'ERREUR MORALISTE

L'argument de « l'erreur moraliste » consiste à soutenir la thèse suivante : il existe des *limites a priori* à ce qui peut être exigé des êtres humains. Une norme qui exigerait trop de sacrifices, ou qui serait trop difficile à respecter pour un être humain, du point de vue de sa motivation, n'est simplement pas une norme morale *valide*. Selon cette position, les normes morales sont *nécessairement* conformes ou « appropriées » à la nature humaine, de sorte que *nécessairement* la vie morale est compatible avec une vie humaine réussie et épanouie.

Ce second argument de « dédramatisation » a pour but de rejeter l'ensemble des théories morales qui sont trop exigeantes sur le plan motivationnel, c'est-à-dire qui sont trop difficiles psychologiquement à mettre en œuvre, et celles qui exigent trop de sacrifices, c'est-à-dire celles qui sont incompatibles avec une vie humaine heureuse et épanouie. Cet

argument vise, en d'autres termes, à rejeter les théories trop « moralisatrices », qui ne tiennent pas compte de la compatibilité de la vie morale avec le bonheur et l'épanouissement des personnes.

Nous allons à présent étudier l'argument métaéthique qui doit servir à montrer qu'il existe des *limites* à ce qui peut être exigé des êtres humains du point de vue moral. Dans un premier temps, nous présenterons la structure de l'argument. Nous listerons ensuite quelques théories qui, si l'argument fonctionne, devraient être rejetées du fait qu'elles ne tiennent pas compte des motivations humaines. Nous répondrons ensuite à l'accusation de « sophisme naturaliste » à laquelle l'argument proposé est confronté. Puis nous verrons que l'argument se fonde sur le principe « devoir-implique-pouvoir ». Nous terminerons le chapitre en présentant un autre argument – l'argument « traditionnel » – en faveur de la prise en compte des motivations humaines dans la formulation des principes moraux, mais nous montrerons que cet argument traditionnel rencontre de nombreuses objections et n'offre ainsi pas un argument suffisamment solide pour notre propos. Nous lui préférons donc l'argument fondé sur le principe « devoir-implique-pouvoir », qui sera l'objet du chapitre suivant.

1.1. LA CONTRAINTE HUMANISTE

L'erreur moraliste consiste à commettre une erreur dans le choix des normes devant servir à guider la conduite des individus humains. Il ne s'agit plus, comme c'était le cas pour le *sophisme moraliste*, de croire naïvement que les êtres humains se plieront spontanément aux exigences morales même en l'absence d'incitations. L'erreur moraliste consiste plutôt à adopter des normes de conduite dont le *contenu* est inapproprié. Autrement dit, il s'agit d'une erreur *normative*, d'une erreur dans le choix de la doctrine morale. Le problème ne provient plus d'une conception erronée de la *mise en application* des normes, mais d'une conception erronée du *contenu* des normes elles-mêmes.

L'erreur moraliste consiste à adopter des normes de conduite qui ne sont pas adaptées à la nature humaine. Ces normes peuvent être inadaptées à la nature humaine de deux manières :

- Il peut s'agir de normes que les êtres humains ne sont pas en mesure d'accomplir à cause de limitations motivationnelles.

- Il peut s'agir de normes de conduites qui ne sont pas compatibles avec une vie qui vaut la peine d'être vécue pour un être humain.¹⁶⁷

Dans les deux cas, commettre une erreur moraliste consiste à adopter des normes morales qui ne sont pas « adaptées » à la nature humaine. Ces normes sont trop exigeantes, soit parce qu'elles s'opposent à des motivations humaines fondamentales, soit parce que la vie morale requiert de sacrifier tout ce qui, pour un être humain, rend la vie digne d'être vécue.

Afin d'éviter de commettre cette « erreur moraliste », une théorie normative devrait respecter la « contrainte humaniste » :

La contrainte humaniste Une théorie politique normative est défectueuse et donc fausse si elle impose des standards ou exigences qui ignorent la nature humaine (ESTLUND, 2011, p. 208).¹⁶⁸

Si une théorie morale contient des exigences incompatibles avec certaines motivations humaines fondamentales ou avec une vie humaine suffisamment épanouie, alors cette théorie morale ne respecte pas la « contrainte humaniste » et doit, en conséquence, être rejetée.¹⁶⁹

¹⁶⁷ Nous verrons plus loin que cette seconde manière pour une norme d'être inadaptée à la nature humaine – i.e. en exigeant trop de sacrifices – est un cas particulier de la première – i.e. celle consistant à ignorer les limitations motivationnelles humaines. Autrement dit, nous considérons la contrainte de ne pas exiger trop de sacrifices comme étant une sous-catégorie de la contrainte d'accessibilité motivationnelle : un être humain ne peut pas désirer vivre une vie qui ne vaut pas la peine, pour un être humain, d'être vécue.

¹⁶⁸ Selon Estlund, le principal problème avec les exigences qui ignorent la nature humaine est qu'elles ne seront, à cause de la nature humaine et des incapacités motivationnelles qui l'accompagnent, jamais satisfaites. Comme il ne s'agit pas cependant, comme nous le verrons, de la seule justification possible de cette contrainte, nous préférons conserver une formulation plus large.

¹⁶⁹ Samuel Scheffler soutient un principe similaire, qu'il qualifie d' « idéal d'humanité » : « [L]es normes morales doivent pouvoir être intégrées de manière cohérente et attractive dans la vie de l'agent individuel. » [“[M]oral norms should be capable of being integrated in a coherent and attractive way into the life of the individual agent.” (SCHEFFLER, 1992, p. 102)]. Nous trouvons aussi une telle exigence, formulée de manière plutôt vague, chez les partisans de la thèse selon laquelle la psychologie morale empirique doit posséder une influence déterminante sur la recherche normative en philosophie morale. La référence régulièrement citée est Owen Flanagan (1991), qui a formulé le principe de réalisme psychologique minimal : « Faites en sorte, en construisant une théorie morale ou en projetant un idéal moral, que le caractère, la méthode de décision, et l'action prescrites soient possible, ou perçues comme telles, pour des êtres comme nous. » Ce principe est repris, de manière toujours aussi vague, par Vanessa Nurock, par exemple : « Le Principe de réalisme psychologique minimal (PRPM) [sert] à limiter les conceptions normatives à celles qui pourraient être réalisées par un Homo Sapiens biologiquement normal et rester stable dans certains arrangements sociaux possibles. » (NUROCK, 2011, p. 76).

1.2. LES THÉORIES MORALES QUI ÉCHOUENT À RESPECTER LA CONTRAINTE HUMANISTE

La contrainte humaniste requiert que toute théorie morale soit compatible ou adaptée à la nature humaine – et surtout à ses limites motivationnelles. Avant d'étudier les arguments métaéthiques en faveur de cette thèse, considérons au préalable quelles sont les théories morales qui se trouveraient affectées par une telle contrainte. Autrement dit, quelles sont les doctrines morales existantes qui pourraient être invalidées par le fait qu'elles ne respectent pas la « contrainte humaniste » ? Étonnamment, cette contrainte met en danger un assez grand nombre de doctrines morales.¹⁷⁰ Elle peut être utilisée pour réfuter, par exemple :

(i) L'utilitarisme

L'utilitarisme peut être accusé, d'un côté, d'être une doctrine morale qui exige trop de sacrifices, puisqu'elle peut exiger d'un individu de sacrifier l'intégralité de ce qui compte pour lui au nom du bien général. Elle peut aussi être accusée, d'un autre côté, d'être une doctrine morale incompatible avec certaines tendances humaines fondamentales, notamment la tendance à accorder à son propre bien-être (et à celui de ses proches) une importance plus grande qu'au bien-être de personnes éloignées. L'utilitarisme est donc inadapté à la nature humaine dans les deux sens susmentionnés ; il ignore les limites motivationnelles de l'être humain et exige (potentiellement) trop de sacrifices.

(ii) L'ascétisme

L'ascétisme préconise de rechercher la douleur et d'éviter le plaisir. Cette doctrine morale n'est pas adaptée à la nature humaine, car l'humain recherche naturellement à éprouver du plaisir et éviter la douleur. En outre, il est difficile de concevoir que la vie d'ascète soit compatible avec une vie humaine heureuse. Comme l'affirme Bentham « le principe de l'ascétisme n'a jamais été, et ne pourra jamais être, poursuivi de façon cohérente par une créature vivante ».¹⁷¹ Bentham parvient à cette conclusion par le biais du raisonnement suivant :

¹⁷⁰ Selon nous, le fait que la contrainte humaniste puisse avoir un effet aussi dévastateur constitue une première raison de suspicion à son égard.

¹⁷¹ "The principle of asceticism never was, nor ever can be, consistently pursued by any living creature." (BENTHAM, 1996, p. 21).

- a. La morale ascétique exige des êtres humains de ne pas éprouver de plaisir et d'éprouver de la douleur.
- b. Tous les êtres humains désirent éprouver du plaisir et ne pas éprouver de douleur ; ils ne peuvent pas désirer éprouver de la douleur et ne pas éprouver de plaisir.
- c. DONC l'être humain n'a pas le devoir moral d'éprouver de la douleur et de ne pas éprouver de plaisir.
- d. DONC la morale ascétique est fausse.

L'argument de Bentham constitue un exemple parfait du type d'argument anti-moraliste que nous étudierons plus loin.¹⁷²

(iii) La morale kantienne

Une conséquence normative de la doctrine morale kantienne est qu'un individu devrait accepter de sacrifier sa vie pour faire son devoir moral, par exemple pour ne pas mentir. L'exemple paradigmatique est celui d'un souverain qui menace de guillotiner un individu si celui-ci ne fait pas une fausse déposition à l'encontre d'un membre honorable de la société. Dans cette situation, pour éviter de mentir, notre individu a le devoir moral de surmonter son amour de la vie, aussi fort soit-il, et de se sacrifier pour préserver l'homme honorable (WHITE, 2005, p. 6). Il est possible de considérer qu'une telle morale est à la fois inaccessible à la motivation – s'oppose au désir de *vivre* – et qu'elle est incompatible avec une vie humaine heureuse – elle exige trop de sacrifices.

(iv) Les théories politiques anti-égalitaristes

L'argument de « l'erreur moraliste » peut également servir à critiquer les théories politiques favorables aux inégalités. Selon cette critique, les êtres humains sont tous envieux par nature – ils sont incapables de ne pas désirer que les autres ne les surpassent pas. En conséquence, les êtres humains n'ont pas le devoir moral d'accepter les inégalités.

¹⁷²Comme nous l'a fait remarquer le professeur Jaffro, l'argument de Bentham semble admettre une équivalence entre désirer « ne pas éprouver de plaisir » et désirer « éprouver de la douleur ». Or, il est tout à fait possible de désirer éviter le plaisir, sans que cela n'implique de *désirer* la douleur. Même s'il est vrai que *dans les faits*, la seule vie compatible avec l'absence de plaisir est une vie de douleur, cela ne signifie pas encore qu'un ascète *désire* cette vie douloureuse. Il désire seulement l'absence de plaisir. Ce constat ne pose cependant pas de problème pour notre propos : il suffit que l'être humain ne puisse pas ne pas désirer éprouver du plaisir pour que la norme morale ascétique exigeant d'éviter le plaisir ne respecte pas la contrainte humaniste.

Une telle critique a été adressée notamment à la théorie de la justice de John Rawls (TOMLIN, 2008). En effet, le philosophe ne tient pas compte, selon la critique, du phénomène de l'envie dans *la formulation* des principes de justice.¹⁷³ Il ne tient pas compte du fait que les moins bien lotis éprouveront de l'envie à l'égard des mieux lotis et risquent de mettre en péril le projet social. Cette absence de considération du phénomène de l'envie est d'autant plus grave que, dans la *formulation* des principes de justice, John Rawls tient compte, par contre, du fait que les êtres humains sont motivés à travailler et à accroître leur production uniquement si cela leur permet d'augmenter leur profit personnel. C'est même *pour cette raison*, i.e. pour tenir compte des faiblesses motivationnelles humaines, que Rawls considère qu'il faut autoriser les inégalités.

[L]es attentes plus grandes permises aux entrepreneurs les encouragent à faire des choses qui augmentent les perspectives de la classe laborieuse (RAWLS, 2009 , p. 109).

Mais alors, lui dira-t-on, pourquoi ne pas également prendre en compte l'envie dans le choix des principes?¹⁷⁴

¹⁷³ Selon John Rawls, nous pouvons « définir l'envie comme la tendance à éprouver de l'hostilité à la vue du plus grand bien des autres, même si leur condition plus favorisée que la nôtre n'ôte rien à nos propres avantages. Nous envions ceux dont la situation est supérieure à la nôtre [...] et nous voulons les priver de leurs plus grands avantages même s'il est nécessaire pour cela que nous renoncions nous-mêmes à quelque chose » (RAWLS, 2009 , p. 574). Le philosophe tient compte de l'envie dans la mise en application de la conception de la justice ; il estime que le phénomène de l'envie pourra être contré par l'amélioration de la confiance en soi des pauvres et l'amélioration des perspectives (RAWLS, 2009 , p. 577), ainsi que par le fait qu'une frontière physique tend naturellement à se mettre en place entre les riches et les pauvres, c'est-à-dire par le fait que « dans une société bien ordonnée, la multiplicité des groupes ayant leur propre vie interne tend à réduire la visibilité – ou du moins la visibilité douloureuse – des différences entre les perspectives des gens » (RAWLS, 2009 , p. 579). Par contre, Rawls refuse que l'on tienne compte de l'envie *dans la formulation* des principes de justice eux-mêmes.

¹⁷⁴ Une solution consisterait éventuellement à dire que l'envie, bien que répandue, ne représente jamais une motivation moralement légitime et qu'elle ne mérite donc pas d'être considérée dans la formulation des principes : « [É]tant donné qu'on considère généralement l'envie comme quelque chose à éviter et à craindre [...], il semble souhaitable que, dans la mesure du possible, le choix des principes ne soit pas influencé par ce trait de caractère. » (RAWLS, 2009 , p. 573) Cependant, une critique similaire pourrait être avancée contre l'absence de motivation des mieux lotis à travailler à augmenter la prospérité générale si cela n'améliore pas suffisamment leur propre situation : « [C]eux qui réduisent l'usage productif de leurs talents parce qu'ils ne perçoivent pas une rémunération excédant suffisamment la simple compensation de leur effort, se comportent en preneurs d'otage. » (ARNSPERGER & VAN PARIJS, 2000, p. 106). Si le manque de motivation des mieux lotis est aussi moralement problématique, alors il faudrait le traiter de la même manière que l'envie. Cela impliquerait, soit de ne tenir compte ni de l'un ni de l'autre dans la formulation des principes, soit de tenir compte des deux. Or, Rawls tient compte de la motivation à travailler, mais ne tient pas compte de l'envie.

Une autre manière de justifier la différence de traitement entre l'envie et la motivation à travailler pourrait être que « des sentiments subjectifs volatiles et malléables comme l'envie, ayant une propension à se fixer sur des objets futiles ou au contraire trop intimes, ne peuvent sérieusement inspirer l'élaboration

(v) Les théories politiques qui prônent l'égalité des chances

L'argument de « l'erreur moraliste » peut aussi être utilisé pour critiquer les politiques prônant l'égalité des chances. L'argument prendrait la forme suivante : les êtres humains ont une tendance naturelle à évaluer leur propre situation par rapport au « groupe de référence » auquel ils appartiennent (BOUDON, 1977). Cette caractéristique a pour conséquence que les êtres humains n'éprouvent pas de déception lorsque, sachant qu'il leur est impossible de gravir les échelons sociaux, ils n'essaient tout simplement pas de le faire. C'est ce qui se passe dans les sociétés de castes ; chacun a une place déterminée dans la société, et seules les personnes appartenant aux classes supérieures se concurrencent pour obtenir les positions revenant aux membres de cette classe. Dans ces sociétés, peu de personnes éprouvent de la frustration, car peu de personnes ont l'espoir d'obtenir une position sociale plus avantageuse.

Par contre, les êtres humains éprouvent beaucoup de déception dans des sociétés socialement plus égalitaires. En effet, dans ces sociétés, beaucoup plus d'individus *essaient* de gravir les échelons sociaux. Cela implique que beaucoup plus d'individus *échouent* à atteindre leurs objectifs. Et ces *échecs* génèrent beaucoup de frustration, puisque le « groupe de référence » auquel ces individus se comparent est le groupe social supérieur auquel ils ont échoué à accéder (BOUDON, 1977). De ce fait, les individus sont globalement plus frustrés et moins heureux dans une société organisée selon le principe de l'égalité des chances que dans une société de castes.

En conséquence, il est faux que les êtres humains ont le devoir moral de favoriser l'égalité des chances, car vivre dans une telle société implique pour eux de renoncer à une part importante de leur bien-être. De ce fait, les théories morales qui prônent l'égalité des chances doivent être rejetées. Elles ne tiennent pas compte du fait que les êtres humains évaluent leur propre situation selon leur groupe de référence.

(vi) Les théories morales impartiales

d'institutions justes » (FLEURBAEY, 1994, p. 16). Toutefois, cette caractérisation de l'envie, comme nécessairement *inappropriée*, c'est-à-dire comme ne représentant pas adéquatement l'objet – i.e. elle porte sur des choses futiles et trop intimes – nous semble reposer sur le « sophisme moraliste » révélé par d'Arms et Jacobson (D'ARMS & JACOBSON, 2000). Nous reviendrons sur ce point dans le dernier chapitre. Dans tous les cas, si nous avons raison, la caractérisation de l'envie proposée par Fleurbaey ne peut servir à refuser de considérer l'envie dans la formulation des principes de justice.

Il est impossible pour un être humain d'accorder à un individu étranger la même importance qu'il accorde à sa sœur, à son père ou à un ami. Si nous acceptons la contrainte humaniste, alors nous devrions conclure qu'il n'est pas le cas que les êtres humains ont le devoir moral d'accorder la même valeur aux individus éloignés qu'aux individus proches. Les théories morales impartiales seraient donc déficientes, voire fausses.

De la même manière, il est impossible pour un être humain d'accorder à la vie d'une taupe à nez étoilé ou à celle d'un cochon des mers ou à celle d'un rat-taupe¹⁷⁵ – pourtant en voie de disparition – la même importance qu'à la vie d'un chat ou d'un panda ou qu'à la vie d'un enfant humain. Cette seconde incapacité s'explique, selon le chercheur Hal Herzog, par une tendance naturelle de l'homme à favoriser les espèces qui « éveillent nos instincts parentaux » plutôt que les espèces qui ressemblent « à des choses primaires repoussantes comme la morve ou la matière fécale » (EXPRESS, 2019). Les scientifiques ont relevé que cette tendance pose des difficultés pour la préservation de la biodiversité (SMALL, 2011). Si nous tenons compte de cette tendance humaine dans la formulation des principes moraux, alors nous devrions conclure que les êtres humains n'ont pas les mêmes devoirs à l'égard des rats-taupes qu'à l'égard des pandas – ou qu'il n'est pas immoral de favoriser les pandas par rapport aux rats-taupes – notamment dans les politiques de protection des espèces menacées.

(vii) Le platonisme

Le platonisme requiert de la part des parents de la communauté de déléguer l'éducation et le soin de leurs propres enfants à l'État. Or, cette exigence ne tient pas compte de l'incapacité, pour les parents, de se motiver à abandonner leurs propres enfants au soin d'autrui. Dans les termes de David Estlund, « les humains ne peuvent pas se motiver à abandonner leurs enfants, et, pour cette raison, [beaucoup considèrent que] la justice n'exige pas que les parents lèguent leur enfants à l'État ».¹⁷⁶

¹⁷⁵ Ces trois animaux partagent trois caractéristiques; ils sont des êtres sensibles, ils sont en voie de disparition et ils sont particulièrement hideux. Or, leur laideur a pour conséquence qu'ils sont moins bien représentés dans les zoos et la recherche scientifique s'intéresse peu à eux. Ils sont ainsi discriminés par rapport à des animaux en voie de disparition ayant des qualités esthétiques supérieures, comme le panda ou le tigre.

¹⁷⁶ "It would be supremely difficult for many parents to surrender their infant children to be raised by the state. [...] [H]umans cannot bring themselves to surrender their children and so, for that reason, [many consider that] justice does not require it." (ESTLUND, 2011, p. 221).

(viii) Les doctrines morales altruistes ou non égoïstes

Selon plusieurs auteurs, la défense de l'égoïsme *éthique* – et donc le rejet correspondant de l'altruisme éthique – dérive d'une conception de l'être humain comme étant motivé exclusivement par l'intérêt personnel, combinée à la thèse selon laquelle un système moral ne doit pas dépasser les capacités motivationnelles de l'être humain. Ainsi, selon C. D. Broad, si l'égoïsme psychologique était vrai, c'est-à-dire que les êtres humains *dans les faits* n'étaient motivés que par leur intérêt personnel, alors la seule doctrine morale qui pourrait être vraie serait celle de l'égoïsme éthique.

Si l'Hédonisme Psychologique, par exemple, était vrai, la seule position morale dont il vaudrait la peine de discuter serait une conception égoïste de l'Hédonisme Ethique. Car personne ne peut avoir l'obligation d'essayer de faire quelque chose qu'il est psychologiquement impossible de faire.¹⁷⁷

De la même manière, selon Jean-Claude Wolf, la doctrine normative de l'égoïsme éthique dérive de l'observation empirique de l'égoïsme humain.

L'égoïsme éthique est le produit de l'observation que les êtres humains se considèrent habituellement comme le centre du monde.¹⁷⁸

Nous retrouvons encore la même idée chez Richard Kraut ; selon lui, il serait possible de défendre l'égoïsme éthique à l'aide de la théorie psychologique de l'égoïsme, combinée à la « contrainte humaniste ». Si le philosophe considère toutefois que cet argument ne fonctionne finalement pas, cela n'est pas à cause de la *forme* de l'argument, structurée autour de la « contrainte humaniste », mais à cause du fait que l'égoïsme psychologique est *faux* en tant que théorie descriptive.

Une manière de défendre l'égoïsme éthique est d'affirmer l'égoïsme psychologique et ensuite de proposer que nos obligations ne peuvent pas excéder nos capacités ; si nous ne pouvons pas nous empêcher de maximiser notre propre bien-être, nous ne devons pas nous tenir à des standards moins égoïstes. Mais

¹⁷⁷ "If psychological Hedonism, e.g. had been true, the only ethical theory worth discussing would have been an egoistic form of Ethical Hedonism. For one cannot be under an obligation to attempt to do what is psychologically impossible" (BROAD, 1949).

¹⁷⁸ "Der ethische Egoismus ist das Produkt einer Beobachtung, nämlich dass sich die Menschen gewöhnlich für das Zentrum der Welt halten." (WOLF, 2004, p. 513)

cette défense est largement rejetée, parce que l'égoïsme psychologique paraît fournir une conception trop simple du comportement humain.¹⁷⁹

Il reste vrai néanmoins qu'un partisan de l'égoïsme psychologique pourrait soutenir que les doctrines morales altruistes sont fausses – et que la seule doctrine possiblement vraie est celle de l'égoïsme moral.

1.3. LE PROBLÈME DU SOPHISME NATURALISTE

Le reproche adressé à l'ensemble de ces doctrines morales par le biais de la « contrainte humaniste » consiste dans le fait que ces doctrines ne tiennent pas compte de certaines tendances ou défaillances motivationnelles humaines. Selon ce reproche, s'il n'est pas possible pour un être humain d'être motivé à faire une action particulière – s'il est contraire à la nature humaine d'être motivé à faire une action particulière – alors il n'est pas un devoir moral pour un être humain de faire cette action.

La première difficulté que soulève cette objection est la suivante : la conclusion normative – i.e. « les êtres humains n'ont pas le devoir moral de faire X » – est obtenue sur la base d'une prémisse *descriptive* – « un être humain *ne peut pas* désirer faire X ».¹⁸⁰ Or, un tel passage d'un énoncé descriptif à un énoncé normatif est traité, en métaéthique, comme étant un passage *fallacieux*. Ce passage fallacieux est qualifié de « sophisme naturaliste ». Il s'agit d'un *sophisme* du fait qu'une conclusion *normative*, au sujet de ce qui *doit être le cas*, est obtenue sur la base de prémisses exclusivement descriptives, au sujet de ce qui *est le cas*, alors que les règles de la logique n'autorisent pas une telle transition. Ce sophisme est dit « naturaliste » du fait que, la plupart du temps, les énoncés descriptifs dont il est question sont des énoncés *naturels*. Néanmoins, il conviendrait mieux d'adopter le qualificatif de « sophisme descriptiviste », car il est possible que l'énoncé descriptif porte sur une propriété non naturelle et non morale, par exemple « Dieu est tout-puissant », « les rousses ont des pouvoirs magiques » et « l'âme humaine est immortelle », etc. Dans tous les cas, ce sophisme consiste à traiter ce qui *est le cas* (descriptif) comme ayant une portée normative – comme déterminant ce qui *devrait être le cas*.

¹⁷⁹ "One way to defend ethical egoism is to affirm psychological egoism and then to propose that our obligations cannot outstrip our capacities; if we cannot help seeking to maximize our own wellbeing, we should not hold ourselves to less selfish standards. But this defense is widely rejected, because psychological egoism seems too simple a conception of human behavior." (KRAUT, 1998).

¹⁸⁰ Nous admettons comme *descriptifs* l'ensemble des énoncés modaux *aléthiques*.

Le sophisme naturaliste concerne la validité *logique* du raisonnement. Il ne s'agit toutefois pas de la seule variante d'erreur naturaliste que peut commettre une théorie morale. En effet, il est possible d'accuser une théorie morale de commettre une erreur naturaliste (ou descriptiviste) à trois niveaux : (A) au niveau logique (« La loi de Hume » ou « sophisme naturaliste »), (B) au niveau sémantique (« L'erreur naturaliste » au sens étroit) et (C) au niveau ontologique et/ou épistémologique (« le fossé faits-valeurs ») (FRITZ, 2009). Dans ce qui suit, nous nous intéresserons exclusivement aux erreurs *logiques* et *sémantiques*, puisque c'est à ces niveaux que se place notre argumentation.

1.3.1. LE SOPHISME NATURALISTE SUR LE PLAN LOGIQUE

L'impossibilité *logique* de dériver une conclusion normative sur la base d'une prémisse descriptive a été formulée déjà au 18^{ème} siècle par David Hume dans le passage suivant :

Dans tous les systèmes de morale que j'ai rencontrés jusqu'ici, j'ai toujours remarqué que l'auteur procède quelques temps selon la manière ordinaire de raisonner, qu'il établit l'existence de Dieu ou qu'il fait des remarques sur la condition humaine ; puis tout à coup j'ai la surprise de trouver qu'au lieu des copules *est* ou *n'est pas* habituelles dans les propositions, je ne rencontre que des propositions où la liaison est établie par *doit* ou *ne doit pas*. Ce changement est imperceptible ; mais il est pourtant de la plus haute importance. En effet, comme ce *doit* ou ce *ne doit pas* expriment une nouvelle relation et une nouvelle affirmation, il est nécessaire que celles-ci soient expliquées : et qu'en même temps on rende raison de ce qui paraît tout à fait inconcevable, comment cette nouvelle relation peut se déduire d'autres relations qui en sont entièrement différentes (HUME, 1973 (1740)).¹⁸¹

Dans ce passage, David Hume accuse ses opposants de commettre une erreur de raisonnement, un « paralogisme ». Cette erreur consiste à déduire, sur la base d'une ou plusieurs prémisses portant exclusivement sur la manière dont le monde *est*, une conclusion au sujet de la manière dont le monde *devrait-être*. Il s'agit d'une erreur logique, car le raisonnement introduit sans le justifier « une nouvelle relation », celle du devoir moral, qui n'était pas contenue dans les prémisses. On qualifie aujourd'hui de « loi de Hume » ou de

¹⁸¹ Livre III, chap. 1, section i.

« Guillotine de Hume » la règle selon laquelle tout passage entre une ou plusieurs prémisses purement descriptives et une conclusion normative est logiquement interdite.

**La « loi de Hume »
au niveau logique**

Il est logiquement interdit de dériver une conclusion normative (au sujet de ce qui devrait être le cas) à partir de prémisses exclusivement descriptives (au sujet de ce qui est le cas). Toute doctrine morale qui viole cette interdiction logique commet le *sophisme naturaliste* (descriptiviste): elle établit un passage de l'être au devoir-être qui est fallacieux (CLAVIEN, 2007, p. 223).

L'impossibilité logique de dériver une conclusion normative sur la base d'une ou plusieurs prémisses descriptives apparaît immédiatement lorsque nous formalisons le type d'argument dont il est question. Prenons par exemple l'argument fallacieux suivant :

(1) La guerre cause beaucoup de souffrance.

(2) $\therefore$ La guerre est immorale

Cet argument est visiblement invalide. En effet, il prend la forme suivante (logiquement invalide) :

(3) a est N

(4) $\therefore a$ est M

où le terme N désigne une propriété *naturelle* (descriptive) – i.e. causer beaucoup de souffrance, alors que le terme M désigne une propriété *morale* (normative) – i.e. être immorale.

Un argument qui obtient de cette manière une conclusion morale (normative) sur la base d'une prémisse descriptive est déficient sur le plan logique. Toutefois, cette déficience est générale ; elle ne provient pas du fait que le concept M désigne spécifiquement une propriété morale. En effet, du fait que la guerre cause de la souffrance, il ne s'ensuit pas *logiquement* que la guerre est immorale, pas plus que, du fait qu'un être humain a deux bras, il ne s'ensuit qu'il a deux jambes. Du point de vue de la logique (syllogistique), l'important est que la conclusion de l'argument fallacieux contient un terme qui n'apparaissait pas déjà dans les prémisses, ce que la logique interdit de manière générale.

Comme l'affirme William Frankena, le problème du « sophisme naturaliste » consiste donc à commettre une faute très générale de raisonnement, consistant à introduire dans la conclusion un terme qui n'était pas déjà présent dans les prémisses, mais ne concerne pas particulièrement le caractère *moral ou normatif* de la conclusion :

[L']argument, dans la forme stricte dans laquelle il est formulé, est fallacieux. Mais il n'est pas fallacieux parce qu'un terme éthique apparaît dans la conclusion alors qu'il n'apparaît pas dans la prémisse. Il est fallacieux parce que tout argument de la forme « A est B, alors A est C » est invalide, lorsqu'il est considéré sous cette forme stricte. Par exemple, il est invalide de soutenir que Crésus est riche parce que Crésus a de la fortune.¹⁸²

Les règles fondamentales de la logique interdisent donc que les énoncés *moraux* puissent être dérivés d'énoncés *descriptifs* – notamment des énoncés empiriques formulés par les disciplines empiriques telles que la physique, la biologie, la psychologie ou l'économie, mais aussi des énoncés descriptifs issus de la métaphysique – par ex. « Dieu a ordonné que... ». La philosophie morale apparaît ainsi comme une discipline *autonome* sur le plan logique.

1.3.2. LES THÉORIES AFFECTÉES PAR L'ACCUSATION DE « SOPHISME NATURALISTE »

Cette critique logique du « sophisme naturaliste », si elle est vérifiée, affecte un très grand nombre de théories morales ou de jugements moraux particuliers – largement répandus – dont voici quelques exemples :

¹⁸² "[The] argument, taken strictly as it stands, is fallacious. But it is not fallacious because an ethical term occurs in the conclusion which does not occur in the premise. It is fallacious because any argument of the form " A is B, therefore A is C " is invalid, if taken strictly as it stands. For example, it is invalid to argue that Croesus is rich because he is wealthy." (FRANKENA, 1939, p. 468)

(1) L'éthique évolutionniste du laissez-faire¹⁸³

- a. La nature favorise la survie du plus fort et la disparition du plus faible.
- b. ∴ L'Etat doit favoriser la survie du plus fort et la disparition du plus faible – notamment en promulguant, par voie législative, ou en protégeant de façon constitutionnelle, une économie politique de laissez-faire.¹⁸⁴

(2) L'utilitarisme économique

- a. L'augmentation de la richesse collective permet d'augmenter le bonheur général.
- b. ∴ Il faut augmenter la richesse collective.

(3) L'utilitarisme négatif

- a. La guerre cause beaucoup de souffrance.
- b. ∴ La guerre est immorale.

(4) L'éthique ascétique

- a. Le monde est fait de difficultés et de souffrances.
- b. ∴ il est bon de vivre avec difficultés et souffrances.

(5) Le rejet de l'homosexualité

- a. Les organes sexuels ont pour fonction de permettre la reproduction de l'espèce.
- b. ∴ l'homosexualité est une perversion, un vice. Un être humain ne doit pas d'entretenir de relations homosexuelles.¹⁸⁵

(6) L'anti-véganisme

- a. Il est naturel pour l'homme de manger de la viande.
- b. ∴ il n'est pas immoral pour l'homme de manger de la viande.

(7) Le positivisme légal

- a. Il est interdit par la loi de voler.
- b. Il est immoral de voler.

¹⁸³ Nous reprenons la formulation proposée par Scott Brewer (BREWER, 1993, p. 280).

¹⁸⁴ Ce fut la position défendue par Herbert Spencer (1879), comme le rapporte Michael Ruse (1993).

¹⁸⁵ Elizabeth Anscombe est connue pour s'être opposée à la fois à la contraception et à l'homosexualité en de termes naturalistes. Selon elle, l'acte sexuel est destiné à la reproduction. Empêcher la reproduction (par la contraception), ou avoir un rapport sexuel où la reproduction est impossible (homosexualité), c'est faire un *mauvais* usage de la sexualité, c'est *mal* agir, puisque c'est agir *contre* la nature. La contraception et l'homosexualité sont donc, selon elle, immorales (ANSCOMBE E. , 1976).

(8) Le favoritisme des proches

- a. Il appartient à la nature humaine de préférer ses proches.
- b. ∴ il est moralement approprié de préférer ses proches.

(9) Le sexisme

- a. Les mères ont un instinct parental plus fort que les pères.
- b. ∴ Il est moralement correct d'attendre que ce soient les femmes qui, au sein du foyer, s'occupent des enfants.

(10) Le racisme

- a. Les hommes blancs sont plus intelligents que les hommes noirs.
- b. ∴ il est légitime que les hommes blancs aient des privilèges, dirigent la société et exploitent les hommes noirs.

Bien entendu, l'accusation de « sophisme naturaliste » nous intéresse plus spécifiquement pour la difficulté qu'elle soulève contre la « contrainte humaniste ». En effet, la contrainte tombe aussi sous l'accusation de « sophisme naturaliste », car, si elle est admise, le fait que les hommes sont *incapables* de désirer faire certaines choses – énoncé descriptif – a pour implication que les hommes *n'ont pas le devoir moral* de faire ces choses – énoncé normatif. Or, une telle transition constituerait une instance du « sophisme naturaliste » que nous venons de présenter, comme le montre l'exemple suivant :

(11) La « contrainte humaniste »

- a. Les êtres humains sont incapables de désirer abandonner leurs propres enfants.
- b. ∴ il est faux que les êtres humains ont le devoir moral d'abandonner leurs propres enfants.

Un tel argument dérive de manière fallacieuse une conclusion morale (normative) sur la base d'une prémisse descriptive. Un tel argument viole en conséquence la loi de Hume.

1.3.3. LA CONTESTATION LOGIQUE DU SOPHISME NATURALISTE

Certains philosophes ont contesté la loi de Hume, selon laquelle il est *logiquement* impossible, à l'aide uniquement des règles logiques syntaxiques, de dériver un énoncé normatif sur la base de prémisses purement descriptives. Arthur Prior, par exemple, a soutenu que, si nous considérons la logique de manière plus large que la seule syllogistique,

alors il devient possible de dériver des conclusions normatives sur la base de prémisses purement descriptives (PRIOR, 1960). Pour le prouver, il présente notamment les arguments logiquement valides suivants :

- a. Il est commun de boire du thé en Angleterre p
- b. Soit il est commun de boire du thé en Angleterre, soit les Néo-Zélandais doivent être tués. $\therefore p \text{ ou } q$

Ou encore

- a. Il n'existe pas d'individu mesurant plus de six mètres $\neg A$
- $\text{ou } \neg \exists x(Fx)$
- b. $\therefore$ Il n'existe pas d'individu mesurant plus de six mètres qui a l'interdiction de s'asseoir sur une chaise ordinaire. $\therefore \neg (A \wedge B)$
- $\text{ou } \neg \exists x(Fx \wedge Gx)$

Ou encore

- a. Tous les croque-morts sont des représentants de l'église. $\forall x (Fx \rightarrow Gx)$
- b. $\therefore$ Si les représentants de l'église ont le devoir d'être respectueux, alors les croque-morts ont le devoir d'être respectueux. $\therefore \forall x (Gx \rightarrow Hx)$
- $\rightarrow \forall y (Fy \rightarrow Hy)$

Dans chacun des cas, la conclusion obtenue possède toutes les caractéristiques d'une proposition normative. Il n'y a donc pas de raison de considérer que cette conclusion n'est pas proprement normative. Or, dans chacun de ces exemples, la conclusion normative dérive de prémisses exclusivement descriptives. Ces exemples invalident donc, selon Prior, la loi de Hume.¹⁸⁶

Bien que l'argument d'Arthur Prior ouvre une discussion très intéressante, nous ne nous y intéresserons pas dans ce qui suit pour deux raisons : premièrement, nous nous intéresserons au passage entre l'être et le devoir-être au niveau *sémantique* plutôt qu'au

¹⁸⁶ D'autres auteurs ont proposé d'autres arguments invalidant la « loi de Hume ». Mark Nelson (2007), par exemple, propose le cas suivant :

- (1) Staline a autorisé le massacre de Katyrí
- (2) $\therefore$ Il n'est pas le cas à la fois que si Staline a autorisé le massacre de Katyrí, alors Staline est mauvais et qu'il n'est pas le cas que Staline est mauvais.

niveau seulement syntaxique, car c'est à ce niveau que peut être défendue la « contrainte humaniste ».

En second lieu, il nous semble que l'argument de Prior manque quelque peu sa cible : comme Prior l'admet lui-même, son argument ne permet pas de dériver des propositions morales à partir de propositions descriptives *dans des raisonnements syllogistiques*. Or, c'est précisément dans ce type de raisonnement que les doctrines morales qui commettent le « sophisme moraliste » sont critiquées ; elles prétendent déduire des conclusions normatives à partir de prémisses descriptives *dans le cadre d'un raisonnement syllogistique*.¹⁸⁷

1.3.4. LE RAISONNEMENT NATURALISTE AU NIVEAU SÉMANTIQUE

Sur le plan logique, il semble donc bel et bien *impossible* de dériver une conclusion morale à partir de prémisses uniquement descriptives. Pour reprendre l'exemple précédent, il est toujours invalide d'opérer la déduction suivante :

(1)	La guerre cause beaucoup de souffrances	a est N
(2)	$\therefore$ La guerre est immorale	$\therefore a$ est M

Il existe cependant une manière de rendre valide la présente déduction. Pour ce faire, il suffit d'ajouter une prémisses intermédiaire, permettant de mettre en relation la caractéristique naturelle N de la prémisses et la caractéristique morale M de la conclusion. Cette prémisses supplémentaire prendrait la forme suivante :

- (5) Tout ce qui est N est M.

Ce qui signifie, dans notre exemple :

- (6) Tout ce qui cause beaucoup de souffrance est immoral.

Cette prémisses intermédiaire (6) permet de rendre logiquement valide le passage de la première prémisses (1) à la conclusion (2), comme le montre explicitement le raisonnement suivant :

¹⁸⁷ Il est vrai que le troisième exemple développé par Prior possède la structure d'un raisonnement syllogistique. Toutefois, la conclusion normative qu'il obtient n'est que conditionnelle ; « si les officiers d'église ont le devoir moral de faire X, alors les croque-morts ont le devoir moral de faire X ». Cette conclusion n'affirme rien de proprement normatif ; elle n'attribue aucun devoir moral à quiconque. En effet, la conclusion ne dit pas que les officiers d'église *ont* le devoir moral de faire X, mais seulement que *s'ils avaient* un tel devoir, alors les croque-morts l'auraient aussi.

- (1) La guerre cause beaucoup de souffrance
- (6) Tout ce qui cause beaucoup de souffrance est immoral
- (2) ∴ La guerre est immorale

Un tel raisonnement est logiquement valide ; cependant, il ne contredit pas la loi de Hume, selon laquelle il est fallacieux de dériver une conclusion normative sur la base de prémisses exclusivement descriptives. Il ne permet pas d'invalider la loi, car la prémisse intermédiaire est elle-même *normative*. En effet, l'énoncé « Tout ce qui cause beaucoup de souffrance est immoral » apparaît clairement comme un énoncé normatif (moral) – i.e. il satisfait les conditions pour être considéré comme un énoncé moral.¹⁸⁸ La conclusion (2) ne dérive donc pas uniquement d'énoncés descriptifs. Or, selon la loi de Hume, il est seulement interdit de dériver une conclusion éthique normative sur la base de prémisses *exclusivement* descriptives. Mais il n'est pas fallacieux de dériver une conclusion normative (morale) sur la base d'une prémisse elle-même normative (morale).

1.3.5. DIFFÉRENTS « PRINCIPES SÉMANTIQUES DE TRANSITION » (PST)

Nous avons affirmé que le précédent raisonnement (valide) s'appuyait sur un énoncé intermédiaire *normatif*, i.e. « tout objet ayant la propriété naturelle N possède la propriété morale M ». Admettons maintenant que cet énoncé – qui met en lien une propriété naturelle et une propriété normative – ne soit pas un énoncé moral, mais un énoncé *analytiquement vrai*, c'est-à-dire un énoncé vrai en vertu seulement de la *signification* des termes qu'il contient.¹⁸⁹ Il deviendrait alors possible de déduire (conceptuellement) un énoncé normatif à partir de prémisses exclusivement descriptives.

¹⁸⁸ Voir la section 3.1. « Énoncés descriptifs et normatifs » de la première partie du présent travail.

¹⁸⁹ La définition des énoncés analytiques est la suivante : les énoncés analytiques sont des énoncés vrais par définition, en vertu de la seule signification des termes qu'ils contiennent. Ils sont vrais en vertu uniquement de la sémantique. Par exemple, l'énoncé « un homme célibataire est un homme non marié » est vrai en vertu de la signification du terme « célibataire », puisque le terme de « célibataire » se définit par « le fait d'être non marié ». Les énoncés analytiques s'opposent aux énoncés synthétiques. Les énoncés synthétiques, de leur côté, ne sont pas vrais en vertu de la seule signification des termes. La sémantique ne suffit pas pour déterminer leur vérité ou fausseté. Pour savoir si un énoncé synthétique est vrai, il est nécessaire de recourir à autre chose qu'à la signification des mots. Par exemple, l'énoncé « les hommes célibataires sont majoritairement des hommes sans emploi » n'est pas vrai en vertu de la signification de « célibataire ». Le mot « célibataire » n'implique pas « sans emploi ».

Cette distinction entre les énoncés analytiques et synthétiques s'accompagne d'une autre distinction importante, à savoir celle entre les énoncés *a priori* et les énoncés *a posteriori*. L'opposition entre le caractère *a priori* ou *a posteriori* d'un énoncé relève de l'épistémologie, de la manière dont nous avons accès à la vérité

Pour défendre la possibilité de dériver (sémantiquement) un énoncé normatif sur la base de prémisses exclusivement descriptives, il faudrait pouvoir montrer qu'il existe un lien *conceptuel* entre un concept descriptif (naturel) N et un concept normatif (moral) M. Plusieurs propositions ont été formulées dans ce sens.

L'une d'elles, la plus répandue, a été de considérer que les notions morales (normatives) étaient *synonymes* de notions naturelles (descriptives).¹⁹⁰ Admettons par exemple que le concept de « être immoral » signifie ou *soit synonyme* de celui de « causer beaucoup de souffrance ». Le raisonnement suivant deviendrait donc valide en vertu de la signification du terme « immoral » :

(1) La guerre cause beaucoup de souffrance

De la prémisses (1), il résulte de la définition du terme « immoral » que

(2) ∴ La guerre est immorale

Que ce soit par la synonymie des termes ou par d'autres moyens, les propositions pour établir une relation *conceptuelle* entre les notions normatives et descriptives ont pris la forme de « principes », qui affirment l'existence d'une relation conceptuelle entre un concept descriptif (naturel) et un concept normatif (moral). Ces principes sont appelés des *principes sémantiques de transition* (PST) (FRITZ, 2009, pp. 84-84).

Principe sémantique de transition	Principe d'après lequel il existe une relation analytique entre un concept descriptif (naturel) N et un concept normatif

ou fausseté de l'énoncé. Un énoncé est *a priori* s'il est possible de savoir que l'énoncé est vrai sans « regarder le monde », sans faire appel à l'expérience. Pour savoir si un énoncé *a priori* est vrai, il suffit de faire un bon usage de ses capacités intellectuelles, de ses capacités de raisonnement. Les énoncés mathématiques sont des exemples typiques d'énoncés connaissables *a priori* : pour savoir s'il est vrai que « 5 + 6 = 12 », il est inutile d'observer le monde. Il suffit de réfléchir adéquatement, de faire un bon usage de ses capacités intellectuelles. A l'inverse, un énoncé est *a posteriori* si, pour déterminer sa vérité ou fausseté, il est nécessaire de regarder le monde, de faire appel à l'expérience. Le seul usage de la pensée ne suffit pas. Par exemple, pour savoir si l'énoncé « le revenu moyen des travailleurs et travailleuses suisses s'élève à 5218 CHF » est vrai, il est clairement insuffisant de réfléchir à la question. Il faut observer le monde, faire des recherches empiriques à ce sujet (JAQUET & NAAR, 2019, pp. 111-112).

¹⁹⁰ G. E. Moore a proposé un argument fameux contre les doctrines qui traitent des notions normatives et des notions naturelles comme étant synonymes (MOORE G. E., 1993 (1903), p. 68). Il accuse ces doctrines de commettre un « sophisme naturaliste », mais le sophisme dont il est question est différent du sophisme logique associé à la « loi de Hume ». Selon Moore, les partisans de la synonymie des notions morales et naturelles commettent l'erreur de raisonnement suivante : de l'observation du fait que tout objet du monde qui a la propriété morale M – i.e. « être immoral » – a également la propriété naturelle N – i.e. « causer de la souffrance », les partisans de la synonymie concluent fallacieusement que « être immoral » signifie « causer de la souffrance ».

(moral) M.

Différents « principes sémantiques de transition » ont été proposés dans le but de permettre le passage d'un ou plusieurs énoncés exclusivement descriptifs à un énoncé normatif. Nous en discernons quatre principaux : (1) la *synonymie* des concepts moraux et naturels, (2) le principe « il faut tenir ses promesses », (3) l'interprétation attributive de la notion de « bien » et (4) le principe « devoir-implique-pouvoir ». Il existe ainsi quatre manières de défendre l'idée qu'il existe une relation conceptuelle entre N et M. Nous nous intéresserons uniquement au dernier principe (4), mais il est important de rappeler qu'il ne s'agit que d'un principe parmi d'autres permettant de franchir, par la sémantique, le fossé logique qui sépare les énoncés descriptifs des énoncés normatifs – et qui sépare donc les sciences empiriques et la philosophie morale.

Dans le tableau ci-dessous, la première colonne répertorie les différents « principes sémantiques de transition » proposés dans la littérature philosophique. La seconde colonne explicite le contenu de l'énoncé qui est censé être « analytiquement vrai » et, finalement, la dernière colonne fournit un exemple du type de raisonnement que chaque « principe sémantique (analytique) de transition » permet de construire. Dans chacun des cas, l'existence d'un lien conceptuel entre la notion descriptive (naturelle) et la notion normative (morale) permet de construire un raisonnement valide dont les prémisses sont uniquement descriptives et la conclusion est normative (morale).

Les principes sémantiques de transitions (PST)	Enoncé analytiquement vrai	Exemple de raisonnement valide
(1) La synonymie : M signifie N	« La propriété d'être M est la même propriété que celle d'être N. »	• La guerre cause beaucoup de souffrance • ∴ La guerre est immorale
(2) La promesse ¹⁹¹	« Il faut tenir ses promesses »	• Jean a promis de rendre visite à sa grand-mère • ∴ Jean a le devoir (moral) de rendre visite à sa grand-mère
(3) Le bien attributif ¹⁹²	« S est bon <i>en tant que</i> membre de l'espèce E si S possède les traits caractéristiques des E et S est <i>mauvais</i> en tant que membre de l'espèce E si S ne possède pas les traits caractéristiques de l'espèce E ».	• Tous les E (espèce) sont F. • L'individu S, appartenant à l'espèce E, n'est pas F. • ∴ S n'est pas un <i>bon</i> E. Il est un E <i>déficient, malade, faible</i>, il ne correspond pas à ce qu'il <i>devrait être</i> en tant que E, « il y a quelque chose de mauvais chez S » (FOOT, 2001, p. 30)
(4) Le principe « devoir-implique-pouvoir »	« Si S ne peut pas faire X, alors S n'a pas le devoir moral de faire X »	• Un être humain ne peut pas sauter plus haut que la Lune • ∴ Un être humain n'a pas le devoir moral de sauter plus haut que la Lune.

¹⁹¹ Jean-Louis Gardies et John Searle ont soutenu le caractère analytique de l'énoncé « il faut tenir ses promesses » (GARDIES, 1987; SEARLE, 1964).

¹⁹² L'interprétation « attributive » de la notion de « bien » a été proposée par les partisans de l'éthique de la vertu, notamment (GEACH, 1956; FOOT, 2001; THOMPSON, 1995; HURSTHOUSE, 1999; ANSCOMBE E. , 1958). Selon ces philosophes, « être bon » n'est pas un *prédicat*. « Etre bon » n'est pas un terme qui peut être défini *indépendamment* du nom auquel il est rattaché. Il s'agit d'une notion *attributive* plutôt que *prédicative* ; l'adjectif ne reçoit de signification que par le biais du nom auquel il est rattaché.

Pour saisir la différence entre les notions *prédicatives* et *attributives*, comparons des exemples de concepts prédicatifs et attributifs qui ne font pas partie du domaine normatif. L'adjectif « rouge » est un prédicat ; sa signification ne dépend pas du terme auquel il est rattaché. Lorsque nous disons « ce cahier est rouge » ou lorsque nous disons « cette voiture est rouge », la notion de « rouge » a la même signification dans les deux cas. En revanche, l'adjectif « grand » est un concept *attributif*. Lorsque nous disons « ceci est un grand rat » ou « ceci est un grand arbre », la signification du terme « grand » dépend du nom auquel il est rattaché. Dans le premier cas, il signifie « grand-en-tant-que-rat » et dans le second cas il signifie « grand-en-tant-qu'arbre » (WILLIAMS, 1972, pp. 38-48).

Selon les éthiciens de la vertu, la notion normative « être (moralelement) bon » a faussement été interprétée comme une notion *prédicative* (similaire à « rouge ») alors qu'il s'agit d'une notion *attributive* (similaire à « grand »). Ainsi, selon eux, lorsque nous affirmons que quelque chose est moralelement bon, nous ne pouvons pas comprendre la signification de « bon » sans connaître le nom auquel il est rattaché. Or, lorsque nous utilisons la notion de moralelement bon, nous voulons dire « bon-en-tant-qu'être-humain ». Et puisque « être bon en tant qu'être humain » signifie « posséder les traits caractéristiques de l'être humain », il suffit de déterminer ce qu'est un être humain (énoncé synthétique *a posteriori*) pour savoir ce que cela veut dire d'être bon, d'être vertueux ou d'être une bonne personne. Les énoncés moraux dérivent d'énoncés descriptifs au sujet de la nature humaine.

Ces quatre « principes sémantiques de transition » permettent, si leur caractère *analytique* est vérifié, de construire un raisonnement valide qui contient uniquement des prémisses descriptives (naturelles) et une conclusion normative (morale). En d'autres termes, pour chacun de ces principes, s'il est prouvé qu'il est vrai en vertu uniquement de la signification des termes qu'il contient, alors il devient possible de rejeter la loi de Hume. Même si celle-ci vaut toujours sur le plan syntaxique, elle n'est plus valable sur le plan sémantique.

L'interprétation *analytique* de chacun de ces principes a été contestée.¹⁹³ Il existe toutefois une façon très générale de caractériser l'objection à l'ensemble des interprétations *analytiques* des principes ci-dessus. Les philosophes qui considèrent qu'il n'existe *aucune* relation conceptuelle entre les notions morales et descriptives accusent leurs opposants de commettre une *erreur naturaliste* ; selon eux, leurs opposants souffrent de *confusion conceptuelle*, en croyant percevoir des relations conceptuelles là où aucune relation de ce type n'existe.

L'accusation d' « erreur naturaliste »

Il n'existe pas de relation *conceptuelle* entre des concepts moraux et des concepts descriptifs. Les « principes sémantiques de transition » sont donc tous

¹⁹³ Ainsi, aujourd'hui, presque plus personne ne soutient la thèse de la *synonymie* des notions normatives et naturelles. Désormais, l'identité des propriétés naturelles et morales est considérée comme étant une thèse *synthétique a posteriori*, exprimant néanmoins une vérité *nécessaire*. De la même manière que nous découvrons *par expérience* que l'eau est identique à H₂O, nous découvrons par expérience que la propriété d'être immoral est identique, par exemple, à la propriété de susciter de la souffrance (BRINK, 1989). Concernant la promesse, beaucoup de philosophes considèrent que le principe « il faut tenir ses promesses » exprime une vérité *synthétique* : pour Adolphe Reinach (2004 (1913)), par exemple, il s'agit d'une vérité *synthétique a priori*, saisie par l'intuition. Pour les empiristes, et notamment David Hume, il s'agit plutôt d'une vérité *synthétique a posteriori*. Selon Hume, les hommes apprennent, dans la vie en communauté, qu'il serait avantageux pour eux de pouvoir échanger des biens et services à des temps différents. Ils inventent alors la convention de la promesse : une personne qui fait une promesse accepte de se soumettre à des pénalités si elle ne tient pas ses engagements. Le motif de l'intérêt personnel – i.e. la crainte des sanctions – suffit pour motiver les individus à tenir leurs promesses dans les communautés de petite taille où l'institution de la promesse existe. Toutefois, ce motif est insuffisant pour les communautés de plus grande dimension. Il est nécessaire d'introduire un « nouveau motif » pour expliquer le respect des promesses au sein de société plus grande. C'est la « sympathie » qui explique l'apparition de ce nouveau motif. En effet, ce processus psychologique de la sympathie suscite un sentiment désintéressé d'approbation de l'institution de la promesse – puisque celle-ci est bénéfique à la société dans son ensemble (COHON, 2006). Ce sentiment soutenu et développé par l'éducation. Or, la notion de « ought » ou de « nécessité morale » fait référence, chez Hume, précisément à cette expérience d'approbation ou de désapprobation désintéressée (BECK, 1974). C'est donc en faisant l'expérience de ce sentiment de désapprobation que les hommes forment l'idée qu'ils « doivent tenir leur promesse ». Le principe « il faut tenir ses promesses » est donc un principe *synthétique a posteriori* chez Hume.

faux. Toute doctrine qui prétend qu'il existe une telle relation analytique est *fausse* et commet une « erreur naturaliste ».

Comme nous le voyons, le problème de « l'erreur naturaliste » ne concerne plus tellement, aujourd'hui, la question purement *logique ou syntaxique* de la transition entre les énoncés descriptifs (naturels) et les énoncés moraux (normatifs). Elle concerne plutôt l'existence d'une relation *conceptuelle* entre des notions descriptives et des notions morales. Le débat concerne le niveau *sémantique*. Dans la suite, nous étudierons spécifiquement l'analyse conceptuelle du principe « devoir-implique-pouvoir » (4) et n'étudierons pas les autres principes – bien que ceux-ci occupent une place importante dans la discussion métaéthique actuelle. Si nous avons mentionné l'existence des autres principes, c'est pour la raison suivante : elle montre que, même si le principe DIP devait ne pas être analytiquement vrai – et donc, ne pas permettre d'établir une transition sémantique entre des énoncés descriptifs et normatifs, cela ne suffirait pas pour affirmer qu'il n'existe aucune relation conceptuelle entre ces énoncés.

Notons encore que, pour le moment, nous n'avons pas pris position sur le caractère analytique ou non du principe DIP. La question reste pour l'instant ouverte. Nous soutiendrons toutefois, plus loin, que le principe exprime bien une vérité conceptuelle.

1.4. LA CONTRAINTE HUMANISTE COMME APPLICATION DU PRINCIPE « DEVOIR-IMPLIQUE-POUVOIR »

Nous avons vu que la « contrainte humaniste », selon laquelle il ne peut pas exister de devoirs moraux incompatibles avec les motivations humaines, ouvre la possibilité d'une transition entre, d'un côté, des énoncés descriptifs et, de l'autre côté, des énoncés moraux. Bien entendu, cette relation est purement *négative* : elle *exclut* certaines obligations imaginables, mais elle n'en pose aucune. En effet, si les êtres humains ne peuvent pas être motivés à faire ce qu'exige une doctrine morale, ou si la manière de vivre exigée par une doctrine morale est trop exigeante pour les êtres humains, alors il est faux que les êtres humains ont le devoir de faire ce que la doctrine morale spécifique exige. Ainsi, par exemple, s'il est impossible pour les êtres humains de renoncer à la satisfaction de leurs intérêts personnels, alors il est faux qu'ils ont le devoir moral de faire une action contraire à leurs intérêts.

Comment les partisans de cette « contrainte humaniste » peuvent-ils défendre leur position ? Comment parviennent-ils à légitimer la transition entre, d'un côté, des énoncés descriptifs au sujet de la manière dont les hommes sont motivés à agir, et, de l'autre côté, des énoncés normatifs au sujet de ce que les hommes ont le devoir moral d'accomplir ? Une manière de défendre leur position consiste à traiter la contrainte humaniste comme étant le résultat d'une extension du principe métaéthique général « devoir-implique-pouvoir » aux possibilités motivationnelles.

Selon le principe DIP, il ne peut pas être moralement obligatoire de faire une action qu'il est impossible d'accomplir et il ne peut pas être moralement interdit de faire quelque chose qu'il est impossible de ne pas accomplir. Par exemple, les êtres humains ne peuvent pas avoir l'obligation morale de sauter plus haut que la Lune ou d'utiliser la téléportation pour sauver une personne qui se noie à mille kilomètres de distance, parce que les êtres humains sont physiquement incapables d'accomplir de telles actions. De la même manière, il ne peut pas être interdit aux êtres humains de respirer ou de penser, puisqu'il ne s'agit pas d'actions que les êtres humains peuvent éviter de faire. Sur le plan psychologique, des parents ne peuvent pas avoir le devoir moral d'accorder aux enfants des autres personnes la même valeur qu'à leurs propres enfants, parce qu'il est psychologiquement impossible pour un parent d'accorder la même valeur aux enfants d'autres personnes qu'à leurs propres enfants. Ce que les personnes ne peuvent pas faire ou ce qu'elles ne peuvent pas éviter de faire impose des limites à ce qui peut être exigé de ces personnes d'un point de vue moral.¹⁹⁴

La thèse défendue par le recours au principe DIP est la suivante : il existe des limites à ce qui peut être exigé, au nom de la morale, des êtres humains. Des normes qui sont trop exigeantes, notamment parce qu'elles sont impossibles à satisfaire pour un être humain, sont invalides et les énoncés moraux déontiques portant sur des devoirs moraux impossibles

¹⁹⁴ Le principe DIP entretient un rapport étroit avec le fameux « principe des possibilités contraires (PPC) », contesté par Frankfurt (1969), selon lequel : S est responsable de faire X seulement si S aurait pu agir autrement. Toutefois, ce rapport n'est pas celui de l'implication directe ; PPC n'implique pas immédiatement DIP. En effet, PPC ne concerne pas le devoir, mais seulement la responsabilité. Pour que PPC implique DIP, il faudrait admettre que si une personne n'est pas responsable de ne pas faire une action X, alors elle n'a pas le devoir de faire l'action X. Plus précisément, il faudrait admettre la thèse suivante : si S ne fait pas X dans des circonstances où S n'est pas responsable de ne pas faire X, S n'a pas le devoir de faire X. Or, cette thèse est discutable. Nous remercions le professeur Glauser pour avoir souligné et clarifié le rapport entre les deux principes.

à accomplir sont des énoncés *faux*.¹⁹⁵ Le principe DIP ainsi employé, i.e. pour établir des *limites* aux exigences morales, est utilisé dans le cadre d'une approche *empiriste* ; les faits au sujet des limites physiques, psychologiques, économiques, biologiques des êtres humains, établissent des limites concernant le contenu des exigences morales. Le principe DIP permet donc de franchir le fossé logique entre les énoncés descriptifs – notamment les énoncés empiriques – et les énoncés normatifs.¹⁹⁶

Dans sa formulation générale, le principe DIP s'énonce de la manière suivante :

Principe « devoir-implique-pouvoir » (DIP)	Nécessairement, si S a le devoir moral de faire X, alors S peut faire X
La version contraposée de DIP	Nécessairement, si S ne peut pas faire X, alors il n'est pas le cas que S a le devoir moral de faire X.

Etant donné que nous nous intéresserons aux limites *motivationnelles* de l'être humain, nous nous intéresserons à l'extension du principe DIP à des considérations *motivationnelles*. Nous étudierons donc la manière dont DIP est utilisé pour imposer des limitations *motivationnelles* aux exigences morales :

DIP-motivation	Nécessairement, si S a le devoir moral de faire X, alors S peut être motivé à faire X
La version contraposée de DIP-motivation	Nécessairement, si S ne peut pas être motivé à faire X, alors il n'est pas le cas que S a le devoir moral de faire X.

A l'aide de DIP-motivation, interprété analytiquement, il devient possible d'établir une transition sémantiquement valide entre des énoncés descriptifs au sujet des motivations humaines et une conclusion normative au sujet des devoirs moraux. Cette transition prendrait, par exemple, la forme suivante :

¹⁹⁵ En effet, si DIP exprime une vérité conceptuelle, alors « S doit faire X » implique que S peut faire X. Or, si S ne peut pas faire X, cela signifie que l'énoncé déontique en question est *faux*, puisqu'il implique quelque chose de faux.

¹⁹⁶ Le lien entre les énoncés descriptifs et normatifs est ici négatif : DIP *exclut* certains devoirs imaginables, sans en poser.

- (1) Les êtres humains sont motivés à agir exclusivement par intérêt personnel. Prémisse descriptive
- (2) Aider un étranger dans le besoin ne permet pas à un être humain de satisfaire son intérêt personnel. Prémisse descriptive

Des prémisses (1) et (2), sur la base du principe DIP-motivation, il résulte que :

- (3) ∴ Il n'est pas le cas que les êtres humains ont le devoir moral d'aider un étranger dans le besoin. Conclusion normative¹⁹⁷

Dans la littérature métaéthique, il est largement admis qu'une condition nécessaire pour pouvoir être motivé à agir est d'avoir un désir ; de façon un peu technique, il est admis que « S peut être motivé à faire X » est vrai à condition seulement qu'il existe, dans l'ensemble motivationnel de S, un désir que l'action X permet de satisfaire. Les désirs contenus dans l'ensemble motivationnel d'un individu sont simplement *tous les états mentaux* de cet individu qui ont une direction d'ajustement monde-esprit, c'est-à-dire toutes les « pro-attitudes » des agents.¹⁹⁸ Dans ce qui suit, nous utiliserons donc la notion de « désir » dans ce sens très large, comme désignant n'importe quelle « pro-attitude ».

La spécification du désir comme condition nécessaire de la motivation a pour conséquence que la « contrainte humaniste » prend une forme assez spécifique dans le débat métaéthique actuel. Elle prend la forme d'une thèse portant le nom d' « internalisme moral ».

Internalisme moral¹⁹⁹

Nécessairement, si S a le devoir moral de faire X,

¹⁹⁷ Certains philosophes contestent que les négations d'énoncés normatifs soient elles-mêmes des énoncés normatifs. Ils contestent donc que le principe DIP permette de tirer une conclusion normative. Pour une réponse à cette objection, voir l'annexe 2.

¹⁹⁸ Cela inclut aussi, comme le souligne Williams, « des dispositions d'évaluation, les tendances à des réactions émotionnelles, des allégeances personnelles et différents projets [...] incarnant les engagements de la personne » [“S can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent.” (WILLIAMS, 1981, p. 105).]

¹⁹⁹ L' « internalisme moral » est distinct de l' « internalisme du jugement moral » que nous avons vu précédemment. En effet, l'internalisme moral porte sur les conditions de vérité des énoncés moraux, alors que l'internalisme du jugement moral portait sur les conditions de réussite des *jugements* moraux. Autrement dit, l'internalisme moral concerne la vérité des énoncés moraux plutôt que les conditions de réalisation d'un état mental, tel que le jugement moral. Plus formellement, l'internalisme moral soutient que l'énoncé « S doit faire

alors il existe dans l'ensemble motivationnel de S un désir que l'action X permet de satisfaire.²⁰⁰

La version contraposée de l'internalisme moral :

Nécessairement, s'il n'existe pas dans l'ensemble motivationnel de S un désir que l'action X permet de satisfaire, alors il n'est pas le cas que S a le devoir moral de faire X.

L'internalisme moral ne constitue rien de plus qu'une reformulation de la « contrainte humaniste » en termes de désirs, étant acceptée la thèse selon laquelle un désir est nécessaire à la motivation.

De prime abord, l'internalisme moral semble peu vraisemblable. En effet, si l'internalisme moral est vrai, cela signifie qu'il suffit qu'une personne n'ait aucun désir qui serait satisfait par le fait d'accomplir une action pour que cette personne n'ait pas le devoir moral d'accomplir cette action. Ainsi, si Fernand n'a aucun désir qui serait satisfait par le fait d'arrêter de battre sa femme, alors il n'a pas l'obligation morale d'arrêter de battre sa femme (WILLIAMS, 1995, p. 39). Cela semble intuitivement peu plausible ; d'habitude nous considérons au contraire que, lorsqu'il s'agit de devoirs moraux, leur validité ne dépend aucunement des désirs individuels ; il est mal que Fernand batte sa femme, et celui-ci a le devoir moral d'arrêter de la battre, peu importe quels sont ses désirs personnels.

Nous verrons cependant qu'il est possible de rendre l'*internalisme moral* plus séduisant, soit (a) en montrant qu'il dérive de prémisses difficiles à rejeter, soit (b) en considérant que ce ne sont pas les désirs *individuels*, mais les désirs *humains* en général qui déterminent ce qui peut ou ne peut pas être exigé d'un individu humain particulier, *en vertu du principe DIP-motivation*.

Dans la littérature métaéthique, l'internalisme moral est bien souvent défendu par la première méthode (a) ; bien qu'en soi contre-intuitive, la doctrine de l'internalisme moral doit être admise en vertu du fait qu'elle dérive de prémisses qu'il est (prétendument)

X » est vrai seulement si S est motivé à faire X, alors que l'internalisme du jugement moral affirme que l'énoncé « S juge qu'il est bon/juste/vertueux de faire X » est vrai seulement si S est motivé à faire X.

²⁰⁰ Stéphane Lemaire formule ce principe de manière explicite : « [I]l paraît absurde d'affirmer qu'un agent doit faire ϕ dans une situation C si cet agent est incapable d'avoir dans cette situation un ensemble de désirs qui pourrait le conduire à faire ϕ . » (LEMAIRE, 2008, p. 150).

difficile de rejeter. De notre côté, nous considérons que la seconde méthode (b) est plus appropriée ; elle permet mieux de défendre l’internalisme moral. C’est pourquoi nous commencerons par exposer la première méthode et justifierons notre choix de ne pas approfondir le débat la concernant. Puis nous discuterons plus longuement la seconde méthode qui, à notre avis, offre le meilleur argument en faveur de l’internalisme moral. Nous concluons cependant que même cette seconde méthode, plus solide, ne permet pas de défendre l’internalisme moral – et la « contrainte humaniste ».

1.5. LA DÉFENSE TRADITIONNELLE DE L’INTERNALISME MORAL

Nous pensons que le meilleur argument en faveur de l’internalisme moral – et de la « contrainte humaniste » – est basé sur le principe DIP-motivation. Toutefois, il existe un autre argument en faveur de l’internalisme moral, fondé sur l’internalisme des *raisons d’agir*. Il s’agit, semblerait-il, de l’argument métaéthique le plus discuté aujourd’hui (FINLAY & SCHROEDER, 2017). Malgré son rôle central dans le débat actuel, nous estimons que cet argument ne repose pas sur des bases solides. Après avoir justifié notre évaluation, nous procéderons donc à l’étude de l’argument basé sur le principe DIP-motivation.

1.5.1. L’INTERNALISME DES RAISONS D’AGIR

L’argument central de la métaéthique en faveur de l’internalisme moral est fondé sur deux prémisses principales, à savoir le rationalisme moral (P1) et l’internalisme des raisons d’agir (P2). L’argument se présente de la manière suivante :

(P1)	Nécessairement, si S a le devoir moral de faire X, alors il existe une raison normative pour S de faire X.	Rationalisme moral
(P2)	Nécessairement, s’il existe une raison normative pour S de faire X, alors il existe, dans l’ensemble motivationnel de S, un désir que l’action X permet de satisfaire.	Internalisme des raisons d’agir
(C1)	∴ Nécessairement, si S a le devoir moral de faire X, alors il existe, dans l’ensemble motivationnel de S, un désir que l’action X permet de satisfaire.	Internalisme moral

Ensemble, les prémisses (P1) et (P2) impliquent que les vérités morales elles-mêmes dépendent des désirs des personnes. Elles impliquent donc l'internalisme moral (C1), selon lequel l'existence d'un devoir moral pour S dépend des désirs de S.

Le rationalisme moral (P1) affirme simplement que la qualité morale d'une action (positive ou négative) fournit nécessairement une raison de faire ou de ne pas faire cette action.²⁰¹ Cette thèse – qui peut valoir pour d'autres concepts normatifs que celui de l'obligation morale – paraît tout à fait intuitive ; en effet, on considère généralement que, si une action possède une qualité morale positive (correcte, bonne, vertueuse), alors cela donne une raison de l'accomplir. À l'inverse, si une action possède une valeur morale négative (incorrecte, mauvaise, vicieuse), alors cela fournit une raison de ne pas l'accomplir. Par exemple, si l'action de faire un don à une œuvre de charité est moralement louable, alors cette qualité morale de l'action constitue une raison de faire un don à une œuvre de charité. Réciproquement, si la participation au commerce d'armement est moralement condamnable, alors cette qualité morale de l'action constitue une raison de ne pas participer au commerce d'armement.

En résumé, le rationalisme moral n'affirme rien de plus que la thèse suivante : si une action est bonne, juste, obligatoire, vertueuse, cela compte en faveur de l'action. Si une action est mauvaise, injuste, interdite, vicieuse, cela compte en défaveur de cette action. Cela semble difficile à contester.

La seconde prémisse (P2) prend le nom d'internalisme des *raisons d'agir*.²⁰² Selon cette thèse, une personne n'a de raison d'accomplir une action que si ladite action permet de satisfaire l'un de ses désirs. Autrement dit, si accomplir une action ne permet de satisfaire aucun désir d'un individu S, alors S n'a pas de raison de faire cette action. Cette prémisse

²⁰¹ Selon les termes de Michael Huemer : « Les faits moraux, s'ils existent, fournissent nécessairement des raisons normative d'agir. S'il est mal de voler [...] alors chacun a une raison normative pour ne pas voler ». [“Moral facts, if there are any, necessarily provide normative reasons for action. If stealing is wrong [...] then one has a normative reason to avoid stealing.” (HUEMER, 2005, p. 156).]

²⁰² Les multiples variantes d'*internalisme* peuvent être résumées par le tableau suivant :

	Enoncé	Condition de vérité
Internalisme du jugement moral	« S juge qu'il est moralement approprié de faire X »	S est motivé à faire X = Faire X satisfait un désir de l'ensemble motivationnel de S
Internalisme moral	« S a le devoir moral de faire X »	
Internalisme de raisons d'agir	« S a une raison (normative) de faire X »	

(P2) est en soi moins intuitive que la prémisse (P1), mais elle repose sur un argument solide, dont chaque prémisse est intuitivement plausible.

1.5.2. L'ARGUMENT DE LA MOTIVATION

Le principal argument en faveur de (P2) s'intitule « l'argument de la motivation » (MARKOVITS, 2014). Il se décline de la manière suivante. En premier lieu, s'il existe des raisons d'agir, c'est-à-dire des considérations qui *comptent en faveur de* certaines actions, alors ces raisons doivent pouvoir, au moins dans certaines circonstances, *expliquer* l'action ou l'attitude des personnes dont elles sont les raisons (WILLIAMS, 1981; MARKOVITS, 2014). Dans un langage plus technique, des raisons normatives – i.e. des considérations qui *rendent appropriées* une action du point de vue objectif – doivent pouvoir être des *raisons* motivantes – i.e. des considérations *en vue desquelles* un sujet agit, c'est-à-dire des considérations qui rendent appropriée l'action *du point de vue de l'agent* et *en vue* desquelles un agent agit.

Par exemple, si le fait que la guerre cause des souffrances est une raison de ne pas entretenir une situation de guerre, alors il est possible, pour un commerçant, d'être motivé à renoncer au commerce d'armement *en raison* de la souffrance que cause la guerre. Dans cet exemple, la souffrance causée par la guerre est à la fois ce qui *rend appropriée* et ce qui *explique* l'action du commerçant qui renonce à son activité. La souffrance est à la fois la raison *normative* et la raison *motivante* de l'action du commerçant.

L'argument continue ainsi : étant donné que les raisons motivantes *expliquent* les actions, elles semblent occuper le rôle de *causes* de ces actions. Or, ce sont des états mentaux qui causent les actions. Cela signifie que les raisons motivantes sont des états mentaux, capables de motiver une personne rationnelle à agir. Et, nous l'avons vu, seuls les désirs sont en général considérés comme des états mentaux ayant une force motivationnelle. De ce fait, il est admis par de nombreux philosophes que, pour qu'une personne rationnelle soit motivée à agir, elle doit avoir deux états mentaux distincts, à savoir un désir et une croyance (au sujet de la manière de satisfaire le désir). Conjointement, ce désir et cette croyance causent l'action de la personne rationnelle (DAVIDSON, 1993).

Reprenons l'exemple du commerçant d'armes : la raison pour laquelle il renonce à son activité pourrait être qu'il *croit* que la guerre cause beaucoup de souffrance et qu'il a le *désir*

de supprimer la souffrance dans le monde. Ce désir et cette croyance causent son action de mettre un terme à son activité économique. Or, étant donné que la cause de son action est aussi sa raison motivante, et que sa raison motivante est aussi une raison normative, alors la raison *normative* d'arrêter le commerce est elle-même composée du désir et de la croyance du commerçant.

Le cas du commerçant peut être étendu à toute raison normative : une raison normative est nécessairement constituée d'un désir et d'une croyance. De ce fait, il ne peut exister de raison normative pour un individu S de faire une action X, s'il n'existe pas dans l'ensemble motivationnel de S un désir que l'action X permet de satisfaire. L'argument mène ainsi directement à l'internalisme des raisons.

En résumé, le raisonnement en faveur de l'internalisme des raisons d'agir prend la forme suivante :

- (P3) Toute raison *normative* est une raison *motivante* potentielle.
- (P4) Les raisons motivantes sont les causes des actions rationnelles.
- (P5) Les causes des actions rationnelles sont composées de deux états mentaux distincts, i.e. un désir et une croyance.
- (C2) ∴ Toute raison normative est composée d'un désir et d'une croyance.

Internalisme des raisons	∴ S'il existe une raison normative pour S de faire X, alors il existe dans l'ensemble motivationnel de S un désir que l'action X permet de satisfaire.

1.5.3. CRITIQUES DE L'ARGUMENT DE LA MOTIVATION

L'« argument de la motivation » en faveur de l'internalisme des raisons que nous avons présenté dans la précédente sous-section, repose sur trois thèses d'identification :

- (A) L'identification des raisons normatives à des raisons motivantes potentielles – prémisse (P3)
- (B) L'identification des raisons motivantes aux causes des actions rationnelles – prémisse (P4)
- (C) L'identification des causes des actions rationnelles à des états mentaux, en l'occurrence à une paire composée d'un désir et d'une croyance (P5)

Or, il est possible de critiquer l'argument à chacun de ces niveaux d'identification. La première critique porte sur le niveau le plus élevé, à savoir l'identification des raisons normatives à des raisons motivantes – la prémisse (P3). Pour certain-e-s philosophes, notamment Suzanne Mantel (MANTEL, 2017), les raisons normatives ne peuvent pas être des raisons motivantes, si ces dernières sont des états mentaux, car les unes et les autres ont des conditions d'identification trop différentes ; les raisons normatives sont identifiées de manière « grossière » (*coarse grained*) alors que les raisons motivantes sont identifiées de manière « fine » (*fine grained*).

Par exemple, imaginons Nathalie qui se promène dans la rue et rencontre un jeune homme aux allures de banquier. Ce dernier semble chercher son chemin. Il s'avère que le jeune homme en question est Emmanuel Macron, président de la république française, mais Nathalie ne le sait pas. Dans cet exemple, le fait qu'un jeune banquier cherche son chemin constitue pour Nathalie une raison (normative) de s'arrêter pour l'aider. Cette raison (normative) est *la même raison* que le fait que *le président de France* cherche son chemin. Les raisons normatives sont donc identifiées de façon *extensionnelle*. Par contre, si Nathalie est motivé à s'arrêter pour aider un jeune banquier à la recherche de son chemin, cela ne signifie pas qu'elle soit motivé à s'arrêter pour aider *le président de la république française* à chercher son chemin. Au contraire, il se peut même que, pour elle, le fait que le jeune homme en question soit aussi le président de France soit une raison *de ne pas* lui venir en aide.²⁰³ Les raisons *motivantes* sont donc identifiées par leur contenu *intensionnel*. Cette

²⁰³ Je remercie Laurent Jaffro pour m'avoir suggéré cet exemple.

différence dans les critères d'identification implique que les deux types d'entités ne peuvent pas avoir le même statut ontologique ; les raisons normatives ne peuvent donc *jamais* être des raisons motivantes – si les raisons motivantes sont des états mentaux (MANTEL, 2017).

[N]ous distinguons entre certaines types d'entités [...] principalement par la manière d'individuer ces entités respectives. Appliquées aux raisons, cette idée est que si la considération motivante « qu'il y a de l'eau dans la bouteille » est distincte de la considération motivante « qu'il y a du H₂O dans la bouteille », alors que la raison normative « qu'il y a de l'eau dans la bouteille » s'avère être précisément la même raison normative « qu'il y a du H₂O dans la bouteille », alors la raison normative qu'il y a de l'eau dans la bouteille ne peut pas être identique à la considération motivante qu'il y a de l'eau dans la bouteille. Cela indique qu'elles appartiennent à des catégories ontologiques distinctes [...].²⁰⁴

Il n'est donc pas problématique que les raisons normatives ne puissent pas servir à expliquer l'action des personnes : ce n'est pas le genre de chose qui peut expliquer une action particulière – contrairement aux raisons motivantes.

La deuxième critique concerne l'identification des raisons motivantes à des *causes* de l'action – i.e. la prémisse (P4). Un argument avancé contre cette identification consiste, par exemple, à montrer qu'il existe entre une *raison motivante* et l'action qu'elle motive une relation conceptuelle – et donc nécessaire (BECK, 1966). Or, s'il existe une relation conceptuelle entre les raisons motivantes et les actions qu'elles motivent, alors les raisons motivantes ne sont pas des causes de l'action, car il n'existe *jamais* de relation conceptuelle entre une cause et son effet – selon l'objection.

En effet, les causes d'un phénomène sont *par nature* distinctes conceptuellement de leur effet : deux événements ne peuvent entretenir une relation causale que si les deux événements peuvent être identifiés indépendamment l'un de l'autre et que la relation qui les relie est contingente. Etant donné que, contrairement aux causes et leurs effets, les

²⁰⁴ “[W]e distinguish between certain kinds of entities [...] primarily in terms of the individuation of the respective entities. Applied to reasons, the idea is that if the motivating consideration ‘that there is water in the bottle’ is distinct from the motivating consideration ‘that there is H₂O in the bottle’, whereas the normative reason ‘that there is water in the bottle’ turns out to be just the same normative reason as ‘that there is H₂O in the bottle’, then the normative reason that there is water in the bottle cannot be identical with the motivating consideration that there is water in the bottle. This indicates that they belong to different ontological kinds [...]” (MANTEL, 2017 p. 5).

raisons motivantes et les actions rationnelles qu'elles motivent ne peuvent pas être identifiées indépendamment l'une de l'autre, cela a pour conséquence, selon la critique, que les raisons motivantes ne peuvent pas être des causes des actions rationnelles.

Si le « désir d'aller à la librairie » était causalement relié à l'action d' « aller à la librairie », alors nous devrions nécessairement être capables de définir et d'identifier le désir sans référence à l'action afin d'ensuite établir l'existence d'une relation contingente entre les deux. Mais la manière dont nous identifions un désir qui est relié à l'action d'aller à la librairie ne consiste pas à découvrir une quelconque cause mentale qui se produit systématiquement avant l'action d'aller à la librairie, pour ensuite nommer cette cause comme étant un « désir d'aller à la librairie ». Le fait est que nous ne pouvons pas identifier la « cause » en question si ce n'est en vertu du fait qu'il s'agit d'un désir d'aller à la librairie, et ainsi la situation décrite comme « désirer aller à la librairie » entretient une relation logique, et non une relation contingente, avec l'action d'aller à la librairie [...].²⁰⁵

Les raisons motivantes ne sont donc pas des causes des actions et ne sont donc pas des états mentaux.²⁰⁶

La critique peut également porter sur le niveau de l'identification des causes d'une action à des états mentaux – la prémisse (P5). L'objection consiste alors à nier que, pour être des causes, les raisons motivantes doivent être des états mentaux. Autrement dit, le fait

²⁰⁵ If "desire to go to the bookstore" were causally related to "going to the bookstore", then it would be necessary that we be able to define and identify the former without reference to the latter in order subsequently to establish a contingent relation between them. But the way we identify the desire that is related to going to the bookstore is not by finding some mental cause that always occurs prior to one's going to a bookstore so that we subsequently name this cause "desire to go to the bookstore". The fact is that we cannot identify the "cause" in question except by virtue of the fact that it is a desire to go to the bookstore, and thus the situation described as "desiring to go to the bookstore" stands in a logical and not a contingent relation to going to the bookstore [...]" (BECK, 1966 p. 163)

²⁰⁶ Donald Davidson avait déjà pris position contre ce type d'argument conceptuel dans son article fondateur "Actions, Reasons and Causes" (DAVIDSON, 1963). Depuis cet article, la conception causaliste de l'action, défendue par Davidson, a dominé les débats, contre les anti-causalistes héritiers de Wittgenstein et Anscombe. Toutefois, l'argument conceptuel contre le causalisme réapparaît dans les débats aujourd'hui, comme le montre ce passage d'un article de Maria Alvarez : « Les explications humiennes ne citent pas des états mentaux qui sont les antécédents causaux des actions ; les généralisations pertinentes reflètent plutôt des connexions *conceptuelles* qui, comme le défendaient certains des philosophes auxquels s'opposait Davidson, existent entre les attributions de croyances et désirs et les descriptions du comportement des personnes comme étant une action de tel ou tel type ». ["Humean explanations do not cite mental states that are causal antecedents of actions; rather the relevant generalizations reflect the *conceptual* connections that, as some of the philosophers that Davidson was arguing against had claimed, exist between ascriptions of beliefs and wants, and descriptions of people's behaviour as being this or that kind of action." (ALVAREZ, 2007, p. 117)].

qu'une raison motivante soit une cause n'implique pas qu'elle soit un état mental. Admettons à nouveau que ce qui motive notre commerçant d'armes à arrêter son activité économique soit le fait que la guerre cause beaucoup de souffrances. La raison de son action est un fait dans le monde : la souffrance causée par la guerre. C'est la considération de *ce fait*, et non pas la considération de sa *croyance* que la guerre cause de la souffrance, qui est la raison qui motive le commerçant à agir. L'élément explicatif de l'action – et donc la *raison motivante* de l'action – est *le fait dans le monde* et non *l'état mental* au sujet de ce fait (DANCY, 2000; ALVAREZ, 2007).²⁰⁷

Nous avons vu trois objections contre l'« argument de la motivation » en faveur de l'internalisme des raisons d'agir, c'est-à-dire contre l'identification des raisons normatives à des désirs (et des croyances). La première objection refuse l'identification des raisons motivantes aux raisons normatives ; la seconde, refuse l'identification des raisons motivantes aux causes des actions rationnelles ; et la troisième, refuse l'identification des causes à des états mentaux. Si l'une de ces trois objections est concluante, alors cela suffit pour refuser l'argument de la motivation en faveur de l'internalisme des raisons d'agir et, indirectement, en faveur de l'internalisme moral.

En conclusion, nous avons vu que l'internalisme moral dépend de la vérité de cinq prémisses (P1)-(P5) qui chacune peut être critiquée et il suffit que l'une d'entre elle soit fausse pour que l'argument complet échoue. En outre, nous avons implicitement accepté la thèse motivationnelle humienne selon laquelle toute motivation requiert la présence d'un désir. Or, il est possible également de critiquer cette thèse et son rejet impliquerait également l'échec de l'argument en faveur de l'internalisme moral. L'internalisme moral ainsi défendu dépend au total de *six prémisses*, chacune individuellement discutable. La thèse semble donc bien fragile. En outre, pour la défendre, il faudrait argumenter en faveur de chacune de ces prémisses, ce qui nous ferait sortir du cadre restreint de la présente recherche.

²⁰⁷ Evidemment, cette réponse pose des difficultés pour rendre compte des cas où les croyances au sujet des raisons d'agir sont fausses : dans ce cas, qu'est-ce qui joue le rôle motivationnel, puisque les faits du monde sensés occuper le rôle de raisons d'agir n'existent pas ?

1.5.4. « DEVOIR-IMPLIQUE-POUVOIR » EN FAVEUR DE L'INTERNALISME DES RAISONS D'AGIR

Certains pourront contester que nous ayons présenté l'ensemble des arguments existants en faveur de l'internalisme des raisons d'agir (P2).²⁰⁸ Il existe en effet un autre argument que l'« argument de la motivation » en faveur de (P2). Cet autre argument a été avancé par Michael Huemer (2005, pp. 158-159).²⁰⁹ Nous ne trouvons pas que cet argument est convaincant, pour les raisons que nous verrons. En outre, nous ne considérons pas que cet argument soit très différent de l'argument basé sur DIP-motivation que nous étudierons par la suite. Nous ne le trouvons donc pas particulièrement intéressant.

Commençons toutefois par exposer l'argument dont il question²¹⁰ :

(P6) Si S a une raison normative de faire X, alors en l'absence d'une raison normative contraire, S doit faire X.

(P7) Si S doit faire X, alors S peut librement faire X.

(P8) Si S peut librement faire X, alors S peut être motivé à faire X

(C3) ∴ Si S a une raison normative de faire X, alors S peut être motivé à faire X.

Internalisme ∴ Si S a une raison normative de faire X, alors il existe dans l'ensemble des raisons motivationnel de S un désir que l'action X permet de satisfaire.

Cet argument ne nous semble pas convaincant ; en effet, comme le fait remarquer Shafer-Landau (2005, pp. 181-183), l'argument dépend d'une équivoque au sujet de la modalité de « pouvoir » utilisée dans (P7) et (P8). En effet, dans (P7) la notion de possibilité utilisée est clairement celle de la possibilité *physique*. Mais si tel est le cas, alors (P8) devient fausse : il est faux que nous pouvons être motivé à faire tout ce qui nous est physiquement accessible. Par exemple, il est possible qu'un individu soit physiquement capable de tuer ses propres enfants, mais qu'il soit incapable d'être motivé à le faire. Pour prendre des exemples

²⁰⁸ L'internalisme des raisons d'agir est, pour rappel, la thèse suivante : si S a une raison de faire X, alors il existe, dans l'ensemble motivationnel de S, un désir que l'action X permet de satisfaire.

²⁰⁹ Huemer, partisan d'un réalisme moral intuitionniste, fait un usage « moraliste » de DIP. En effet, il déduit une conclusion *descriptive* « S peut être motivé à faire X » sur la base d'une prémisse normative « S a une raison de faire X ». Nous reprendrons la forme générale de son argument pour défendre la thèse de l'existence d'une motivation proprement morale chez les êtres humains, mais sans faire de détour par l'internalisme des *raisons d'agir*.

²¹⁰ Nous reprenons ici la formulation raccourcie proposée par François Jaquet et Hichem Naar (2019, p. 171)

moins tragiques, il est possible que je n'arrive pas à me motiver à faire mon ménage, alors que je dois faire mon ménage et que je suis physiquement capable de faire mon ménage. Il est donc faux que la possibilité physique implique la possibilité motivationnelle.²¹¹

Si nous interprétons alors la notion modale de possibilité comme faisant directement référence à la possibilité motivationnelle, c'est la prémisse (P7) qui devient problématique. En effet, elle signifierait alors que, si S doit faire une action, alors S peut être motivé à faire cette action. Or, cela revient à affirmer l'internalisme moral (ou, dans ce cas, l'internalisme normatif) que nous cherchons justement à prouver. Utiliser cet argument en faveur de l'internalisme moral revient donc à commettre une pétition de principe.

Nous ne considérons donc pas que l'argument présenté ci-dessus en faveur de l'internalisme des *raisons d'agir* – et, indirectement, en faveur de l'internalisme moral – soit concluant. En outre, l'argument repose sur le principe « devoir-implique-pouvoir » (P7), que nous nous apprêtons à étudier dans le cadre d'un argument *direct* pour l'internalisme moral. Il nous semble qu'il est plus approprié de nous intéresser à l'usage de ce principe pour soutenir *directement* la thèse de l'internalisme moral plutôt qu'à son usage dans un argument permettant de soutenir seulement *indirectement* l'internalisme moral, c'est-à-dire en défendant d'abord l'internalisme des raisons d'agir.

En conclusion, nous ne nous intéresserons plus, dans la suite, à l'argument en faveur de l'internalisme moral basé sur les prémisses (P1)²¹² et (P2)²¹³ pour plusieurs raisons : en premier lieu, couvrir l'intégralité des arguments en faveur ou en défaveur de ces thèses dépasserait les limites de notre propos. En second lieu, le bref aperçu de la discussion sur ce sujet montre qu'il existe de nombreux arguments susceptibles de mettre à mal la prémisse

²¹¹ Cela n'exclut pas qu'il existe des situations où notre incapacité de *faire* quelque chose soit expliquée précisément par une incapacité *d'être motivé* à faire l'action en question. C'est même cette relation qui, selon nous, justifie l'extension de DIP à des considérations motivationnelles. Toutefois, une telle situation n'implique pas que toute possibilité physique implique la possibilité d'être motivé. Par exemple, il se peut que Nathalie soit incapable de se motiver à serrer la main d'Emmanuel Macron – tout en sachant qu'il s'agit d'Emmanuel Macron. Cela peut la rendre incapable de serrer la main du jeune banquier qui se tient devant elle. Dans ce cas, si DIP-motivation est accepté, elle n'a pas le devoir moral de *serrer la main du jeune banquier*. Il reste cependant vrai qu'elle est physiquement capable de serrer la main du jeune banquier. Si elle ne savait pas que le jeune banquier est en fait Emmanuel Macron, elle lui serrerait effectivement la main. L'événement « Nathalie sert la main d'Emmanuel Macron » n'est donc pas une impossibilité physique, bien que des limitations *motivationnelles* puissent empêcher sa réalisation.

²¹² Le rationalisme moral

²¹³ L'internalisme des raisons d'agir

(P2), notamment en attaquant (P3)²¹⁴, (P4)²¹⁵ ou (P5)²¹⁶. La thèse (P2) paraît donc assez fragile. Finalement, le principal argument retenu actuellement dans la discussion en faveur de (P2) est un argument qui fait appel à la notion de « devoir » et qui s'appuie sur le principe « devoir-implique-pouvoir ». Or, l'argument le plus fort – à notre avis – en faveur directement de l'internalisme moral (C1) est un argument qui fait un usage plus direct et mieux adapté du principe « devoir-implique-pouvoir ». Il n'est donc pas besoin de défendre l'internalisme moral en défendant d'abord l'internalisme des raisons d'agir (P2) par le biais de DIP. Le principe DIP peut être utilisé directement pour défendre l'internalisme moral (C1).

1.6. LE PRINCIPE DIP-MOTIVATION

Nous avons affirmé que le meilleur argument en faveur de l'internalisme moral consistait à étendre le principe DIP à la motivation humaine. Nous avons appelé DIP-motivation le principe résultant de l'extension du principe général DIP au cas des motivations humaines.

Les raisonnements obtenus en faisant usage de DIP-motivation prendraient, nous l'avons vu, la forme suivante :

(P1)	Les êtres humains sont motivés à agir exclusivement par intérêt personnel.	Prémisse descriptive (positive)
(P2)	Aider un étranger dans le besoin ne permet pas à un être humain de satisfaire son intérêt personnel	Prémisse descriptive (négative)
(P3)	Les êtres humains n'ont le devoir moral de faire une action X que s'ils peuvent être motivés à faire X	DIP-motivation
(C1)	∴ Il n'est pas le cas que les êtres humains ont le devoir moral d'aider un étranger dans le besoin	Conclusion normative

L'internalisme moral ne découle de rien de plus que du principe DIP-motivation et de la théorie humienne de la motivation, selon laquelle toute motivation requiert nécessairement la présence d'un désir.

²¹⁴ L'identification des raisons normatives à des raisons motivantes potentielles

²¹⁵ L'identification des raisons motivantes aux causes des actions rationnelles

²¹⁶ L'identification des causes des actions rationnelles à des paires de désirs et de croyances

(P3) Les êtres humains n'ont le devoir moral de faire une action X que s'ils peuvent être motivés à faire X.	DIP-motivation
(P4) Toute motivation dépend d'un désir ²¹⁷	Théorie humienne de la motivation
(C2) ∴ Les êtres humains n'ont le devoir moral de faire une action X que s'il existe dans l'ensemble motivationnel des êtres humains un désir que X permet de satisfaire.	Internalisme moral

Si le principe DIP-motivation est vrai et que la théorie humienne de la motivation est vraie, alors l'internalisme moral s'ensuit logiquement.²¹⁸

Étant donné que DIP-motivation est une extension du principe général DIP, nous devons commencer par défendre la version générale de DIP, avant de l'étendre au cas de la motivation. C'est pourquoi nous commencerons par étudier les arguments en faveur de DIP dans sa formulation générale, avant de nous intéresser à DIP-motivation.

²¹⁷ Plus précisément, la théorie humienne affirme la thèse suivante : S ne peut être motivé à faire X que s'il existe dans l'ensemble motivationnel de S un désir que l'action X permet de satisfaire.

²¹⁸ Nous admettons pour le moment que la théorie humienne de la motivation est vraie, bien que, selon nous, elle ne soit vraie que de manière triviale. Nous reviendrons sur ce point.

LE PRINCIPE « DEVOIR-IMPLIQUE-POUVOIR »

Nous avons vu que l'argument qui, selon nous, permet le mieux de défendre l'internalisme moral, est un argument fondé sur le principe « devoir-implique-pouvoir », lorsque celui-ci porte sur les possibilités motivationnelles. C'est pourquoi nous allons étudier ce principe. Nous allons procéder de la manière suivante : nous commencerons par présenter et clarifier la signification du principe. Nous discuterons ensuite du principal argument en faveur de l'interprétation de DIP comme exprimant une *vérité conceptuelle*. Nous verrons alors plusieurs objections qui ont été formulées à l'encontre de cette interprétation conceptuelle de DIP. Puis nous passerons en revue les interprétations alternatives du principe DIP, mais nous considérerons que ces propositions alternatives souffrent de difficultés rédhibitoires. Nous terminerons le chapitre par l'étude d'une objection plus spécifique contre l'usage de DIP-motivation en faveur de la « contrainte humaniste ». Selon cette objection, même si nous acceptons que DIP-motivation exprime une vérité conceptuelle, il reste que ce principe ne peut servir à rejeter des théories morales, *car il n'existe rien que l'être humain soit incapable de désirer ou de valoriser*. Autrement dit, selon cette dernière objection, le principe DIP-motivation est tout simplement vide. Le chapitre nous mènera alors à la conclusion que, bien que DIP-motivation exprime, selon nous, une vérité conceptuelle, il ne peut pas servir à rejeter, sur la base de considérations motivationnelles, certaines théories normatives.

Nous répondrons seulement dans le chapitre suivant, i.e. dans le chapitre intitulé « une défense du moralisme », aux objections soulevées contre l'interprétation conceptuelle de DIP. Nous y verrons aussi que, si DIP-motivation ne permet pas d'obtenir des conclusions normatives sur la base de prémisses descriptives (au sujet des motivations humaines), il permet, à l'inverse, de tirer des conclusions descriptives (au sujet des motivations humaines) sur la base de prémisses normatives. Avant de nous intéresser à cela, commençons néanmoins par présenter ce fameux principe DIP.

1. LES VARIÉTÉS DE POSSIBILITÉS

Le principe DIP soutient que, s'il n'est pas possible pour un individu d'accomplir une action, alors cet individu n'a pas le devoir moral de l'accomplir. Bien entendu, l'interprétation du principe DIP dépend entièrement de la signification des notions de « devoir » et de « pouvoir ».

En ce qui concerne la signification de la notion de « devoir », nous nous contenterons de la notion morale, puisqu'il s'agit de celle qui nous intéresse. Certes, le principe DIP peut être appliqué à des notions normatives non morales, mais nous ne nous en préoccupons pas dans ce travail. Nous nous contenterons d'étudier les notions normatives spécifiquement morales, c'est-à-dire les notions normatives relatives à des règles portant sur la manière dont les êtres humains se traitent les uns les autres, éventuellement sur la manière dont les êtres humains se traitent eux-mêmes.²¹⁹

En ce qui concerne la notion modale de « possibilité », elle peut également recevoir de nombreuses interprétations ; métaphysique, physique, psychologique, motivationnelle, ou encore, éventuellement, biologique ou économique. Pour chacune de ces interprétations, le principe DIP implique que, si une action est *impossible* à accomplir, alors cette action ne peut pas être l'objet d'un devoir moral.

Le tableau suivant présente quelques exemples de devoirs moraux qui, si nous acceptons DIP, seraient invalidés par le fait que l'action est *impossible* à réaliser – que ce soit sur le plan logique, métaphysique, physique, psychologique, biologique, motivationnel, etc. La première colonne énumère les différents types pertinents d'impossibilité et la seconde colonne donne un exemple de devoir moral qui se trouverait invalidé par le fait que l'action est impossible à accomplir.²²⁰

²¹⁹ Bien entendu, il est possible de vouloir intégrer dans le domaine moral le traitement que les êtres humains font subir à d'autres espèces. Nous n'y faisons pas référence ici, car il s'agit d'un objet sujet à controverse. A l'inverse, tout le monde s'accorde sur le fait qu'une composante centrale de la morale consiste à définir la manière dont il est approprié, pour les êtres humains, de se traiter les uns les autres.

²²⁰ Il est important de préciser l'élément suivant : tout le monde admet, dans les discussions sur DIP, que ce principe ne concerne pas les cas où l'impossibilité d'accomplir son devoir est le résultat d'une *faute antérieure* de la part de l'agent (BÉLANGER, 2011, pp. 113-115; SINNOTT-ARMSTRONG, 1984, p. 260). Si l'agent est fautif de l'impossibilité d'agir dans laquelle il se trouve, alors il n'est pas le cas que l'agent ne peut pas accomplir son devoir. Il aurait pu accomplir son devoir, mais il s'est mis par lui-même, de manière fautive, dans l'incapacité de l'accomplir. Dans ce cas, on parle, à la suite de Michael Stocker, d'incapacité coupable (STOCKER, 1971, p. 114).

Types d'impossibilité	Devoirs moraux invalidés par DIP
Logique	« Tu dois travailler et ne pas travailler ! » ²²¹
Métaphysique	« Tu dois changer le passé! »
Physique	« Tu dois sauter plus haut que la Lune ! » « Tu dois travailler 23 heures par jour, tous les jours de l'année ! » « Il est interdit de causer des effets externes négatifs. » ²²² « Tu dois éteindre le soleil ! » (GARDIES, 1972, p. 77) « Tu dois aider les générations passées ! »
Biologique	« Il faut empêcher son corps de vieillir ! » « Il est interdit de satisfaire tout désir sexuel ! » « Vis en dehors de toute communauté ! » [Pour un homme adulte] : « Deviens un chimpanzé, une femme, un enfant ! » ²²³
Economique	« Il faut baisser le prix des loyers sur le marché et diminuer le nombre de logement disponibles ! » « Il faut supprimer complètement le chômage ! » « Il faut vendre la totalité du maïs mondial à un prix acceptable ! » (O'CONNOR & THOMPSON, 2009, p. 70) « Il faut rendre la nourriture rare et très demandée accessible à tous les portemonnaie ! » « Le gouvernement doit (A) établir un taux fixe d'échange avec l'étranger, (B) supprimer les barrières au flux de capitaux et (C) avoir une politique monétaire indépendante. »
Politique	« Il faut imposer un service militaire de longue durée à quelques individus – plutôt qu'un service militaire de courte durée à tous les individus » ²²⁴

²²¹ Il nous semble que nous pourrions ajouter des impossibilités « sémantiques » du type : « Tu dois chanter sans émettre de sons », « tu dois réfléchir sans penser », « tu dois observer en fermant les yeux » ou « sauter sans décoller les pieds du sol », « tu dois laver ce qui n'est pas sale », « tu dois tuer un mort », etc.

²²² Stéphane Chauvier soutient que, puisqu'il est impossible « de réaliser un monde dans lequel toutes les actions humaines seraient immunisées contre la production d'effets externes négatifs » (CHAUVIER, 2013, p. 97), alors il est faux qu'il est obligatoire de ne jamais produire de tels effets externes. Cela implique, selon l'auteur, d'octroyer aux individus une *permission* de produire des effets externes négatifs, dans des circonstances spécifiques (CHAUVIER, 2013, p. 98).

²²³ « Oui, l'homme constitue d'une certaine manière son propre projet et il se fait lui-même. Mais il a des limites à ce qu'il peut être. Le "projet" est biologiquement prédéterminé pour tous les hommes ; il est celui de devenir un homme. Il ne peut adopter pour lui-même le projet de devenir un chimpanzé. Ou même une femme. Ou un bébé. » (MASLOW, 2004, p. 68).

²²⁴ Selon Jon Elster, même si le service militaire de longue durée, imposé uniquement à quelques individus, est la solution la meilleure techniquement, elle n'est pourtant pas « politiquement faisable » : « La solution technique optimale, telle que d'imposer à quelques individus un long service militaire, plutôt que d'imposer à tous les individus un service militaire de courte durée, peut être politiquement impossible à réaliser à cause des tendances universalistes inhérentes à la société moderne. » [“The technically optimal solution, such as imposing a long military service on some individuals rather than a short service on everyone,

Psychologique	«Tu dois calculer en 10 secondes de tête 234 x 456!» « Tu dois ne jamais éprouver de la jalousie ! »
Motivationnelle	«Tu dois tuer tes propres enfants!» «Tu dois renoncer au bonheur !» « Tu dois renoncer à la satisfaction de tes propres intérêts ! »

De toutes ces possibilités, c'est bien entendu la dernière qui nous intéressera dans ce travail. Il s'agit des limites aux exigences morales qui proviennent de limites à ce que les êtres humains peuvent être motivés de faire.

Avant de procéder, notons que nous indiquons des exemples d'impossibilités motivationnelles dans notre tableau uniquement à titre d'exemples de *ce à quoi ressembleraient* des impossibilités motivationnelles *si elles existaient*, mais nous soutiendrons par la suite que de telles impossibilités n'existent pas en fait. En effet, notre propre position sera la suivante: nous soutenons que DIP permet d'exclure les normes qu'il est logiquement, biologiquement, physiquement, psychologiquement, économiquement impossible de respecter. Nous soutenons également que *s'il existait des incapacités motivationnelles*, alors ces incapacités supprimeraient également les devoirs moraux. Nous arguerons cependant qu'il n'existe pas, selon nous, de telles incapacités motivationnelles.

Selon la version *motivationnelle* de DIP, ce que les sujets moraux peuvent ou ne peuvent pas être motivés à accomplir restreint le contenu des devoirs et interdictions morales auxquelles ces sujets sont soumis. Certains théoriciens considèrent que les restrictions proviennent de possibilités et d'impossibilités *universelles*, alors que d'autres considèrent que les restrictions proviennent de possibilités et d'impossibilités *individuelles*. Pour les premiers, ce que *les êtres humains* ne peuvent pas désirer, parce que cela est contraire à la nature ou essence de l'être humain, ne peut pas constituer le contenu d'un devoir moral. Par exemple, il ne peut pas être un devoir moral pour l'être humain de renoncer à toute source de bonheur, puisqu'il est dans la nature des êtres humains de désirer être heureux.

Pour les seconds théoriciens, les obligations morales sont limitées par les incapacités motivationnelles *individuelles*. Ainsi, une personne ne peut pas avoir une obligation de faire

may be politically impossible because of universalistic tendencies inherent in modern society." (ELSTER, 1978, p. 53)].

une action, si cette action ne permet pas de satisfaire au moins quelque désir de cette personne. Prenons Marc, par exemple, qui n'a aucun désir qui pourrait être satisfait pas le fait de venir en aide aux pauvres ; dans ce cas, il n'a pas le devoir moral de venir en aide aux pauvres. Comme nous l'avons vu, cette interprétation *individuelle* de DIP-motivation a pour conséquence indésirable que les devoirs moraux varient d'une personne à l'autre en fonction des désirs individuels. C'est pourquoi nous favoriserons l'interprétation *générale* ou *universelle* des notions de *possibilités*. C'est ce qui est possible ou impossible au niveau de l'espèce humaine qui définit les limites aux interdictions et devoirs moraux. La plupart des théoriciens qui abordent cette question des limitations motivationnelles aux exigences morales font d'ailleurs usage de l'interprétation universelle de la notion de possibilité (NUROCK, 2011; ESTLUND, 2011; SCHEFFLER, 1992).

Au sujet des limites provenant de la motivation, précisons l'élément suivant : l'incapacité motivationnelle est un genre d'incapacité dans lequel les individus sont incapables de faire ce qu'ils font *tout en sachant ce qu'ils font*. Il s'agit donc d'une incapacité qui est définie *intensionnellement*. Il s'agit

d'une incapacité à faire quelque chose consciemment. Avec les incapacités simplement physiques, par contraste, si je ne peux pas ϕ , alors il n'existera jamais un événement dans le monde qui sera un exemple de moi faisant ϕ . Si Rambo ne peut pas soulever un poids de 200 kg, alors il ne soulèvera pas un tel poids, voilà tout. [...] Ce caractère absolu de l'incapacité s'applique également à de nombreuses incapacités psychologiques. Si je ne peux pas, dans ma tête, calculer en 10 secondes le produit de deux nombres, qui chacun contient cinq chiffres, alors il s'agit de quelque chose que je ne ferai jamais, peu importe la manière dont la tâche m'est présentée et peu importe les conditions dans lesquelles je suis invité à le faire.

D'autres incapacités consistent en l'incapacité de faire ϕ tout en sachant que l'on est en train de faire ϕ . Cela peut inclure des choses telles que marcher sur une planche étroite au-dessus des gorges de l'Avon ; avoir une conversation détendue et amicale avec l'amant de sa femme ; et manger du rat rôti. ²²⁵

²²⁵ "[Such incapacity is an] incapacity to do a certain thing knowingly. With a straightforward physical incapacity, by contrast, if I cannot ϕ , then there is no event that the world will contain that will be a ϕ -ing by

Dans l'ensemble de ces situations, il n'est pas impossible que l'action soit *produite* : mais il est impossible qu'elle soit produite *et* que la personne soit consciente de ce qu'elle est en train de faire.

Par exemple, il est possible qu'un tueur à gages puisse être motivé à tuer le maire de la ville de Blackburn²²⁶, mais il n'aurait pas pu se motiver à le faire s'il avait su que le maire était son propre père. En effet, (nous supposons) il est impossible pour un être humain de vouloir tuer son propre père. Ainsi, s'il est vrai que les êtres humains ne peuvent pas avoir le devoir moral de tuer leur propre père, cela ne provient pas du fait qu'il est physiquement impossible pour un être humain de tuer son propre père, mais parce qu'il est impossible pour un être humain de (être motivé à) *tuer son propre père* tout en sachant qu'il s'agit de son propre père.

2. L'ARGUMENT SÉMANTIQUE EN FAVEUR DU PRINCIPE DIP

Nous avons spécifié, dans les précédentes sections, la signification du principe DIP. Ce principe doit servir, nous l'avons vu, à défendre la « contrainte humaniste », selon laquelle une théorie morale trop exigeante pour l'être humain est déficiente. Nous allons maintenant étudier différents arguments en faveur de DIP. Nous commencerons par exposer un argument sémantique en faveur de DIP. Selon cet argument, le principe DIP exprime une vérité *conceptuelle* au sujet de la notion d'obligation. Il s'agit de l'argument le plus fort en faveur du principe DIP. En effet, si DIP exprime une vérité *conceptuelle*, cela implique que toute théorie morale trop exigeante est *fausse* ; une telle théorie attribue aux êtres humains des capacités qu'ils n'ont pas en fait. Les autres arguments en faveur de DIP – i.e. les arguments par *présupposition*, *pragmatique* ou *moral* – sont moins forts : selon ces

me. If Rambo cannot lift a 500lb weight, then he will not lift it, period. [...] This absoluteness applies to many psychological incapacities as well. If I cannot, in my head, work out in ten seconds the product of two numbers, each of which consists of five different digits, then that is something I shall not be doing, however the task is represented to me and whatever the conditions under which I am invited to do it. [...] Other incapacities take the form of one's being unable to ϕ if one knows that one is ϕ -ing. These might include such things as walking a narrow plank over the Avon gorge; having relaxed and friendly conversation with the lover of one's spouse; and eating roast rat." (WILLIAMS, 1995, pp. 48-49)

²²⁶ Blackburn est une ville quasi-réelle d'Angleterre, dans laquelle on peut avoir le beurre et l'argent du beurre. ["Blackburn is a quasi-real town in England, in which you can have your cake and eat it too" (DENNETT & STEGLICH-PETERSEN, 2008)].

arguments, si une théorie morale ne respecte pas DIP, elle n'est pas fausse, mais *déficiente* sous un autre aspect – par exemple, elle est inefficace ou immorale.

Etant donné que l'argument sémantique est le plus fort en faveur de DIP, nous commencerons par le présenter. Nous exposerons ensuite quatre difficultés auxquelles fait face cet argument sémantique. Nous proposerons alors trois conceptions alternatives du principe : une conception comme « présupposition », une conception pragmatique et une conception morale. Nous verrons toutefois que chacune de ces conceptions de DIP rencontre ses propres difficultés.

2.1. LA SÉMANTIQUE MINIMALE DE LA NOTION DE « DEVOIR »

Un des principaux arguments en faveur de DIP est un argument *sémantique* : la notion de « devoir » ou « d'obligation »²²⁷ implique sémantiquement la notion de « pouvoir ». Quiconque comprend correctement la signification du terme « devoir » reconnaît que « devoir » implique « pouvoir ». Ainsi, si cette conception de DIP est correcte, il ne ferait pas sens, il serait contradictoire, de dire « John ne peut pas payer 5000 CHF d'impôts chaque mois, mais néanmoins John *doit* ou *a l'obligation morale* de payer 5000 CHF d'impôts chaque mois ». Une telle phrase contiendrait une contradiction, car, selon l'argument sémantique, la définition de la notion de « devoir » implique que le devoir *peut* être accompli.

Nous pouvons comparer cette analyse sémantique du terme « devoir » avec l'analyse sémantique de la notion de « célibataire » ou d'« être humain » ; il serait absurde, parce que contradictoire, de dire « John est célibataire, mais néanmoins John est marié », parce que la *définition* du terme « célibataire » implique celle de « non marié ». De la même manière pour la définition de « humain ». Si nous admettons que la définition d'un « être humain » est celle d'un « animal rationnel », alors il serait contradictoire de dire que Jack est un être humain, mais que néanmoins Jack n'est pas un animal (ou qu'il n'est pas rationnel). Il en va de même pour la notion de « devoir » ; si l'analyse sémantique est vérifiée, il est contradictoire de dire que John a le devoir de faire X bien que John soit incapable de faire X, simplement parce que la *définition* du terme « devoir » implique la notion de pouvoir.

²²⁷ Dans ce qui suit, nous utiliserons les notions de « devoir » et d'« obligation » de manière interchangeable.

Il faut maintenant prouver que l'analyse sémantique de la notion de « devoir » implique bien celle de « pouvoir ». Pour le montrer, nous allons suivre le raisonnement de Gardies dans son ouvrage *L'erreur de Hume* (1987).

L'analyse de la notion du devoir ou de l'obligation morale que nous allons proposer repose sur la sémantique des mondes possibles. Le principe de l'analyse sémantique du concept de « devoir », par le biais de la sémantique des mondes possibles, est le suivant : il existe un contenu sémantique « minimal » des notions de devoir et d'obligation, telles que nous les utilisons dans le langage ordinaire. Or, ce contenu sémantique minimal suffit pour valider DIP sur le plan sémantique. De plus, les sémantiques minimales proposées permettent de rendre compte de manière appropriée de certains usages que nous faisons des notions normatives dans le langage ordinaire. Ainsi, en plus d'être minimales, les sémantiques proposées *éclaircent* la manière dont nous utilisons les termes déontiques.

Cette signification minimale est la suivante : peu importe ce qui fait qu'un acte est bon ou mauvais, juste ou injuste, peu importe, en d'autres termes, la nature du bien et du mal, les notions morales nous permettent de partager les mondes possibles entre, d'un côté, les mondes moralement permissibles/positifs et, de l'autre côté, les mondes moralement impermissibles/négatifs. Peu importe la nature de la distinction entre le bien et le mal, le juste et l'injuste, le moralement obligatoire et l'interdit, nous savons qu'un monde dans lequel le mal est réalisé est un monde moralement négatif ou inadmissible et qu'un monde dans lequel le bien est réalisé est un monde moralement positif ou admissible.²²⁸ Dès lors que nous admettons cette analyse minimale des notions morales, en termes de mondes possibles *admissibles et inadmissibles*, il est possible d'attribuer des conditions de vérité minimales aux énoncés moraux, sans se prononcer sur la nature ontologique du bien et du mal.

De la même manière que la sémantique des mondes possibles a permis d'attribuer des conditions de vérité minimales aux énoncés modaux à l'aide de la notion d'*accessibilité* au monde actuel, sans connaître la nature ontologique de cette relation d'accessibilité, il est possible d'attribuer des conditions de vérité minimales aux énoncés moraux à l'aide de la

²²⁸ Il peut aussi exister des mondes *moralement neutres*, qui sont positifs sous certains aspects, et négatifs sous d'autres aspects, et dont les aspects positifs et négatifs s'équilibrent. Pour notre propos, nous nous contentons de séparer les mondes positifs et négatifs *selon que* le sujet accomplit ou non un devoir moral particulier.

relation d'*admissibilité* et de *permissibilité* par rapport au monde actuel, sans attendre d'avoir déterminé la nature métaphysique exacte de cette relation (GARDIES, 1987, pp. 45-46). En effet, il a été possible d'attribuer des conditions de vérité minimales à l'énoncé « il est nécessaire que p » dans m_1 en attribuant à cet énoncé les conditions de vérité suivantes : dans tous les mondes *accessibles* à m_1 , p est vraie. Dans cette sémantique, la nature ontologique de la relation d'*accessibilité* reste indéterminée. De la même manière, nous pouvons attribuer des conditions de vérité à l'énoncé « il est moralement obligatoire que S fasse X » à l'aide de la relation de « permissibilité » ; « S a le devoir de faire X » est vraie en m_1 , si dans tous les mondes possibles *permissibles*, S fait X .²²⁹

Une fois acceptée l'idée d'une sémantique *minimale* des notions morales, il faut encore montrer que cette sémantique implique bien le principe DIP, et qu'elle correspond à un usage ordinaire des concepts moraux. C'est ce que nous allons montrer dans la prochaine sous-section.

2.2. LES DEUX SÉMANTIQUES DU DEVOIR MORAL

Selon Gardies, il existe deux versions de la sémantique minimale du « devoir », basée sur la sémantique des mondes possibles, brièvement exposée dans la section précédente (GARDIES, 1987, pp. 38-43). Selon lui, chacune de ces sémantiques correspond à un usage différent – et concurrent – que nous faisons dans le langage ordinaire de la notion de « devoir » ou « d'obligation ». Or, étant donné, comme nous le verrons, que ces sémantiques impliquent toutes les deux que le principe DIP est vrai, nous pouvons en conclure que, dans toutes les sémantiques disponibles, le principe DIP est vrai. Nous considérons donc que DIP est un principe vrai en vertu de la seule sémantique.

Le raisonnement général prend donc la forme disjonctive suivante : il existe deux sémantiques possibles de la notion de « devoir ». Chacune correspond à un usage particulier de la notion de « devoir » dans le langage ordinaire. Ces deux sémantiques ont toutes les deux DIP parmi leurs théorèmes. Si nous admettons qu'il n'existe pas d'autres sémantiques

²²⁹ Nous reprenons ainsi l'argument de J. L. Gardies : « Déjà la considération des simples modalités ontiques obligeait à introduire la distinction du vrai et du faux dans d'autres mondes encore que celui de notre expérience immédiate, le nécessaire et l'impossible devenant par exemple le respectivement vrai ou faux dans tous les mondes métalinguistiquement possibles ; il suffisait de considérer des mondes métalinguistiquement non seulement possibles, mais reconnus comme moralement ou juridiquement positifs (ou dans d'autres cas négatifs), pour pouvoir donner aux modalités déontiques, obligatoire, interdit, permis et facultatif, un fondement rigoureusement vérifonctionnel. » (GARDIES, 2006, p. 16).

de la notion de « devoir », alors nous pouvons conclure que DIP est vrai dans toutes les sémantiques disponibles capables de rendre compte du discours normatif commun. Donc, DIP est analytiquement vrai.

Il nous faut maintenant exposer ces deux sémantiques de la notion de « devoir » et montrer que chacune valide effectivement le principe DIP.

2.2.1. LA PREMIÈRE SÉMANTIQUE DE L'OBLIGATION MORALE

La première sémantique est très minimaliste. Selon cette sémantique, l'énoncé moral « S a le devoir moral de faire X » est vrai dans le monde actuel si et seulement si :

- (1) Dans tous les mondes possibles²³⁰ moralement *admissibles*, S fait X.
- (2) Il existe un monde possible moralement admissible.

Cette sémantique implique-t-elle le principe « devoir-implique-pouvoir » ? Elle le fait. Pour le vérifier, en suivant Gardies (1987, pp. 39-40), il suffit de faire l'hypothèse que le principe DIP est faux et de montrer que cette hypothèse mène à une contradiction. Par un tel raisonnement par l'absurde, il est ainsi possible de prouver que la sémantique proposée a bien DIP parmi ses théorèmes.

Supposons donc que DIP soit faux. Supposons donc que « devoir » n'implique pas « pouvoir ». Cela signifierait que l'énoncé « S a le devoir moral de faire X et S ne peut pas faire X » soit vrai. Nous pouvons alors établir le raisonnement par l'absurde suivant :

- « Il est obligatoire que S fasse X » :

Pour tout *m admissible*

VRAI	FAUX
S fait X	

- « Il est impossible que S fasse X » :

Pour tout *m*

VRAI	FAUX
	S fait X

²³⁰ Nous l'avons vu, la nature de la possibilité peut varier selon que nous considérons la possibilité métaphysique, physique, psychologique ou motivationnelle.

- Etant donné qu'il existe au moins un monde possible admissible, alors :

Il existe un m où

VRAI	FAUX
S fait X	S fait X

Ce raisonnement montre donc que la sémantique minimaliste de la notion de « devoir » proposée ci-dessus implique le principe DIP, car la négation de ce principe forcerait à admettre un monde possible dans lequel la proposition « S fait X » est vraie et la proposition « il n'est pas le cas que S fait X » est vraie également, c'est-à-dire dans lequel les propositions p et $\neg p$ sont vraies toutes les deux. Pour éviter d'admettre des mondes contenant de telles contradictions, il faut donc rejeter l'hypothèse et admettre le principe DIP.

Comme le fait remarquer Gardies, le principe DIP n'est pas le seul principe validé dans la sémantique proposée. La sémantique valide en effet un autre principe, à savoir le « principe de nécessité », selon lequel le nécessaire implique l'obligatoire :

Le principe de nécessité²³¹ S'il est nécessaire que S fasse X, alors il est obligatoire que S fasse X

Le fait que la sémantique proposée valide également ce principe de nécessité est important, car il s'agit d'un principe qui, selon Gardies, est admis dans le langage normatif ordinaire. De ce fait, nous obtenons confirmation que la sémantique des mondes possibles correspond bien à la manière dont nous utilisons le langage normatif dans le discours ordinaire.

A nouveau, la preuve de la validité du « principe de nécessité », dans la sémantique proposée, s'établit par un raisonnement par l'absurde (GARDIES, 1987, p. 41) : nous supposons, à l'intérieur de la sémantique proposée, que le principe est faux, et montrons que nous obtenons une contradiction. Nous supposons ainsi qu'il est vrai que « il est nécessaire que S fasse X et il n'est pas obligatoire que S fasse X ».

²³¹ Nous reviendrons plus loin sur le caractère contre-intuitif de ce principe.

- « Il n'est *pas obligatoire* que S fasse X » :

Il existe au moins un *m admissible* où

VRAI	FAUX
	S fait X

- « Il est nécessaire que S fasse X » :

Dans tous les *m*

VRAI	FAUX
S fait X	

- DONC :

Il existe un *m* où

VRAI	FAUX
S fait X	S fait X

Etant donné que la négation du « principe de nécessité » mène à une contradiction dans la sémantique proposée, cela signifie que la sémantique valide ce principe de nécessité. Autrement dit, si nous acceptons la sémantique, alors nous acceptons que la nécessité de faire une action implique l'obligation de la faire.

Selon nous, contrairement à ce que soutient Gardies, le « principe de nécessité » ne semble pas, de prime abord, très intuitif. Il fait notamment face à de nombreux contre-exemples ; le fait que nécessairement les êtres humains ne courent pas aussi vite que la vitesse de la lumière ne semble pas impliquer que les êtres humains ont l'obligation morale de ne pas courir à la vitesse de la lumière. Du fait que les êtres humains ne sont nécessairement pas des chimpanzés, il ne semble pas s'ensuivre qu'ils *ont le devoir ou l'obligation* ne pas être des chimpanzés.

Il apparaît néanmoins que nous faisons parfois usage d'une sémantique conforme à ce principe. C'est le cas, par exemple, lorsque nous identifions une *obligation* à l'absence de possibilité d'agir autrement. Lorsqu'un magistrat dit « nous sommes obligés de couper dans les dépenses », il veut dire que « nous n'avons pas le choix », « il est nécessaire de le faire ». L'obligation provient donc de l'impossibilité de faire autrement. L'idée selon laquelle « chacun doit/a le devoir de travailler » est souvent défendue par l'argument selon lequel il est (physiquement ou biologiquement) nécessaire que les êtres humains travaillent. De la même manière, les êtres humains *doivent* se nourrir, vivre en société, former des couples,

dans un sens qui implique qu'ils *n'ont pas le choix* : un être humain est ainsi fait qu'il ne peut pas vivre s'il n'agit pas ainsi. Nous retrouvons une sémantique de ce type par exemple dans les écrits de Philippa Foot :²³²

La manière dont un individu doit être est déterminée par ce qui est nécessaire pour son développement, sa survie et sa reproduction : dans la plupart des espèces, cela inclut la défense et, dans certaines espèces, cela inclut l'éducation des enfants (FOOT, p. 33.).²³³

Si le lecteur continue de douter – comme nous le faisons nous-mêmes – que le langage ordinaire admet une sémantique aussi contre-intuitive que celle du « principe de nécessité », force est d'admettre que, sur le plan philosophique, plusieurs théories ont, à leur manière, adopté le « principe de nécessité » dans leur sémantique de l'obligation morale. Le cas le plus connu, auquel Gardies fait lui-même référence, est celui de Leibniz, selon lequel « nécessité fait loi ».

Le philosophe allemand parvient, selon Gardies, à cette conception de la relation entre la nécessité et l'obligation morale de la manière suivante :

- (P1) Ce qui est nécessaire est nécessaire pour chaque homme.
- (P2) Ce qui est nécessaire pour chaque homme est nécessaire pour l'homme vertueux.
- (P3) (Par définition) Ce qui est nécessaire pour l'homme vertueux est obligatoire.
- (C1) ∴ Ce qui est nécessaire est obligatoire.²³⁴

Nous trouvons aussi, selon Wilson (1982), une sémantique validant le « principe de nécessité » chez John Stuart Mill. Dans sa théorie morale, un être humain doit (*must*) faire X, s'il est vrai universellement et en vertu des lois de la nature que tous les êtres humains font

²³² L'usage de l'expression : « Tu ne peux pas faire cela ! », par exemple, « tu ne peux pas tuer ton concurrent pour gagner la course ! » offre l'indice d'un usage sémantique qui identifie impossibilité et devoir négatif, et donc nécessité et devoir.

²³³ "The way an individual should be is determined by what is needed for development, self-maintenance, and reproduction: in most species involving defense, and in some the rearing of the young." (FOOT, 2001, p. 33)

²³⁴ « [L]e necessarium pris simpliciter, c'est ce qui est nécessaire en tout cas et pour tous; il est donc aussi nécessaire à l'homme bon; et ce qui est nécessaire à l'homme bon est par définition le debitum » (GARDIES & KALINOWSKI, 1974, p. 101).

X » (WILSON, 1982, p. 68). Ainsi, dans sa doctrine, chaque être humain a le devoir moral de poursuivre son propre bonheur, étant donné que, en vertu des lois de la nature, tous les êtres humains poursuivent le bonheur. L'obligation morale dérive donc bien de la nécessité (nomologique).²³⁵

Certains philosophes ont ainsi fait usage de la notion de « devoir » conformément au « principe de nécessité ». La première sémantique minimaliste proposée offre en ce sens l'avantage de correspondre à un usage existant de la notion de « devoir ». Comme nous allons le voir, il est possible de développer une seconde sémantique minimaliste qui, elle aussi, correspond à un certain usage ordinaire de la notion de « devoir » – un usage, à notre avis, beaucoup plus intuitif. Cette seconde sémantique valide également le principe DIP.

2.2.2. LA SECONDE SÉMANTIQUE DE L'OBLIGATION MORALE

Nous avons vu dans la section précédente qu'une sémantique basée sur des conditions de vérité minimalistes et se rapportant à un usage ordinaire de la notion d'obligation permet de dériver le principe DIP. Nous avons vu aussi que le fait que plusieurs philosophes adoptent une sémantique dans laquelle la nécessité implique l'obligation nous autorise à prendre au sérieux cette première sémantique, validant aussi bien le « principe de nécessité » que DIP.

La seconde sémantique est à peine plus complexe que la première. En effet, elle reprend les deux mêmes conditions que la première sémantique, auxquelles elle ajoute une condition supplémentaire. Les conditions de vérité des énoncés d'obligation sont, dans cette seconde sémantique, les suivantes :

« S a le devoir moral de faire X » est vrai dans le monde actuel SSI :

- (1) Dans tous les mondes possibles moralement admissibles, S fait X.
- (2) Il existe au moins un monde possible admissible
- (3) Dans au moins un monde possible (inadmissible), S ne fait pas X

²³⁵ Diderot également accepte cette doctrine selon laquelle les êtres humains ont l'obligation morale de poursuivre leur propre bonheur, en vertu du fait qu'ils le désirent nécessairement : « Il n'y a qu'un devoir, c'est d'être heureux. Puisque ma pente invincible, inaliénable est d'être heureux, c'est la source unique de mes vrais devoirs, et la seule base de toute bonne législation. » (DIDEROT, 1963).

Cette seconde sémantique valide également le principe DIP. La preuve est exactement la même que la preuve que nous en avons donnée dans la sous-section précédente, et nous ne la renouvellerons donc pas ici.

La différence entre cette nouvelle sémantique et la précédente provient du fait qu'elle valide un autre principe que le « principe de nécessité ». Ce nouveau principe exprime la caractéristique des devoirs et obligations d'être « précaires », c'est-à-dire de toujours pouvoir être violés. Ce principe de la « précarité du devoir » rend compte du fait qu'un devoir authentique est tel qu'il est possible d'y désobéir.

Le « principe de précarité » S'il est nécessaire que S fasse X, alors il n'est pas le cas que S a le devoir moral de faire X.

Ce principe est impliqué par la troisième condition (3). Celle-ci a en effet pour conséquence que tout sujet moral a la possibilité de ne pas accomplir son devoir moral, de ne pas respecter son obligation morale. Cela implique que s'il n'est pas possible pour S d'échouer à accomplir son devoir, alors S n'a tout simplement pas de (véritable) devoir. Dans cette nouvelle sémantique – contrairement à la première – la nécessité (modalité aléthique) exclut le devoir moral (modalité déontique).

Cette seconde sémantique rend compte du fait qu'il appartient à la nature des obligations de pouvoir être violées. Une obligation qui ne peut pas être violée ne constitue pas une *obligation* véritable. Par exemple, étant donné qu'il est impossible pour un individu de ne pas respirer, il est absurde de dire qu'il a *le devoir moral* de respirer ; il ne peut de toute façon pas désobéir, puisqu'il ne peut pas décider d'arrêter de respirer – du moins pendant un temps long. De la même manière, s'il tout être humain nécessairement poursuit déjà son propre bonheur, l'idée d'un *devoir moral de poursuivre le bonheur* serait absurde.²³⁶

Selon Gardies, et nous le suivons sur cette voie, le « principe de précarité » correspond à la manière dont nous employons les concepts déontiques dans le langage ordinaire. Nous considérons, en général, que si un devoir ne peut pas être violé, alors il ne s'agit pas d'un véritable devoir. Cette seconde sémantique des mondes possibles rend ainsi compte,

²³⁶ C'est même précisément ce qui sépare les lois de la nature et les lois morales ou les normes en général : on ne peut pas désobéir aux premières alors qu'on peut désobéir aux secondes (von WRIGHT, 1963, chap. 1., sect. 6.).

comme la première, d'un certain usage ordinaire du langage normatif. La sémantique proposée semble donc tout à fait correspondre à la signification ordinaire de la notion d'obligation. Et puisque, elle aussi, valide le principe DIP, nous avons à nouveau la preuve que DIP est admis dans le discours normatif ordinaire.

Pour prouver que cette seconde sémantique – à peine plus étoffée que la première – valide le « principe de précarité », il est possible de reprendre le schéma des preuves précédentes : il s'agit de faire l'hypothèse que le principe est faux, pour ensuite montrer que, dans la sémantique proposée, cette hypothèse mène à une contradiction.

Supposons donc la négation du « principe de précarité ». Sa négation prendrait la forme suivante : « Il est nécessaire que S fasse X et S a le devoir moral de faire X ». Cette proposition donnerait lieu aux tableaux de vérité suivants :

- « Il est nécessaire que S fasse X » :

Dans tous les m

VRAI	FAUX
S fait X	

- « S a le devoir moral de faire X » :

Dans tous les m admissibles

VRAI	FAUX
S fait X	

et

Il existe un monde m_1 où

VRAI	FAUX
	S fait X

- DONC :

Dans le monde m_1

VRAI	FAUX
S fait X	S fait X

A nouveau, la présence de la contradiction doit nous amener à rejeter l'hypothèse qui y mène, à savoir la négation du « principe de précarité ». Ainsi, la seconde sémantique de la notion de « devoir », beaucoup plus proche – à notre avis – de l'usage ordinaire que la

première, valide à la fois le principe DIP et un autre principe, à savoir le « principe de précarité ».

En conclusion, il existe deux sémantiques minimales, qui chacune rend compte d'un usage que nous faisons, dans le langage ordinaire, de la notion de « devoir ». Ces deux sémantiques, bien que concurrentes, valident toutes les deux le principe DIP.

Il nous semble par ailleurs très difficile de contester le contenu minimal de ces deux sémantiques ; en effet, il semble incontestable qu'une des fonctions du langage moral – et de la pensée morale – consiste, au minimum, à départager entre les mondes possibles moralement admissibles et ceux qui ne le sont pas. Or, les deux sémantiques proposées rendent simplement compte de cette fonction ; en soi, elles n'impliquent rien sur le plan normatif ; elles sont neutres concernant la nature ontologique du bien et du mal, de la justice et de l'injustice, de la vertu et du vice.²³⁷

Malgré son caractère, de prime abord, difficilement contestable, l'analyse sémantique de DIP a rencontré plusieurs objections. Dans la prochaine section, nous allons les présenter. Nous verrons ensuite d'autres manières dont les philosophes ont tenté de rendre compte du fait que DIP possède une forte valeur intuitive. Nous soutiendrons, cependant, que ces autres manières de défendre DIP rencontrent leurs propres difficultés et ne peuvent donc pas, selon nous, remplacer l'argument conceptuel en faveur du principe.

3. LES OBJECTIONS CONTRE L'ANALYSE CONCEPTUELLE DE DIP

Différentes objections ont été avancées contre l'analyse *conceptuelle* de DIP, et en conséquence contre la logique déontique dans son ensemble, puisque celle-ci admet, dans sa forme classique, le principe DIP. Nous pouvons distinguer quatre *types* d'objections : le premier type (3.1.) consiste à fournir des *contre-exemples* au principe DIP. Le second type de critique (3.2.) fait remarquer que l'analyse conceptuelle de DIP implique que certaines situations ordinaires, dont nous faisons l'expérience dans le monde actuel, devraient être

²³⁷ Les logiques déontiques actuelles, plus récentes que celle de Gardies, admettent d'ailleurs des sémantiques similaires ; ces sémantiques ne se prononcent pas sur la nature du bien et du mal, mais se contentent d'admettre que, parmi tous les mondes possibles, certains sont moralement meilleurs que d'autres et il est possible de déterminer quel est le meilleur des mondes possibles – qui n'est pas nécessairement le monde actuel (WEDGWOOD, 2006).

considérées comme logiquement *impossibles*. Or, comme ces situations existent, elles ne sont pas impossibles. Si DIP implique que ces situations actuelles sont impossibles, alors DIP ne peut pas être une vérité conceptuelle. La critique suivante (3.3.) soutient qu'il faut rejeter l'analyse conceptuelle de DIP car celle-ci contrevient à la loi de Hume, selon laquelle il n'est pas possible de dériver un énoncé normatif à partir d'un énoncé exclusivement descriptif. Pour finir, la dernière objection (3.4.) contre l'analyse conceptuelle soutient que celle-ci, contrairement à ce qu'elle devrait faire, ne respecte pas un principe de « neutralité déontique ».

3.1. LES CONTRE-EXEMPLES

Une première méthode pour critiquer l'analyse conceptuelle de DIP consiste à avancer des contre-exemples. Ces contre-exemples décrivent des situations dans lesquelles il n'est pas possible pour une personne de faire ou ne pas faire une action, et pourtant nous considérons que la personne *a le devoir moral* de faire ou de ne pas faire ladite action. Sinnott-Armstrong propose l'exemple suivant (1984, p. 251):

La coupe de cheveux du directeur Il reste vrai que j'ai le devoir moral de ne pas rire de la nouvelle coupe de cheveux de mon directeur, même si je ne peux pas m'empêcher de rire.

Un autre contre-exemple au principe DIP est celui du banquier raciste : un banquier a été éduqué de manière à détester les personnes d'Asie de l'Est et a continué à cultiver son racisme par la lecture de tracts et de livres confirmant son point de vue et en participant à des groupes racistes. Par son éducation et ses actions, le banquier est devenu incapable de prendre des décisions impartiales à l'égard de la race. Il n'est pas possible de remédier à son racisme au moment où le banquier doit décider d'accorder le crédit. Le banquier refuse le prêt à un client asiatique, alors qu'il a accordé des prêts semblables à des caucasiens ayant un profil similaire. Bien que le banquier soit incapable de choisir autrement, nous considérons qu'il a le devoir moral de ne pas discriminer le client sur la base de son appartenance raciale (MONGE, STRUDLER, & KIM, 2015).

Le banquier raciste Il reste vrai que le banquier raciste a le devoir moral de traiter ses clients asiatiques de manière équitable, même s'il est incapable de les traiter équitablement.

De ce type de contre-exemples, un philosophe comme Sinnott-Armstrong tire la conclusion que le principe DIP n'est certainement pas vrai conceptuellement. « Devoir » n'implique pas analytiquement « pouvoir ». L'énoncé « S doit faire X » n'est pas rendu *faux* par le fait que S ne peut pas faire X.

Un partisan de la « contrainte humaniste », fondée sur le principe DIP, peut répondre de la manière suivante : si la contrainte est dite *humaniste*, c'est précisément parce qu'elle concerne les possibilités ou impossibilités d'action déterminées par la nature humaine, et non pas les possibilités ou impossibilités d'action relative à un individu spécifique – ou même à un groupe particulier d'individus (NUROCK, 2011; ESTLUND, 2011; SCHEFFLER, 1992). Ce serait d'ailleurs en ce sens qu'il faudrait comprendre l'analyse conceptuelle de DIP : dans le langage ordinaire, lorsque nous utilisons la notion de « devoir », la notion de « pouvoir » qui en dérive est traitée comme relative à l'être humain et non pas comme relative à un individu ou à un groupe particulier. Ainsi, lorsque nous interprétons le principe DIP comme analytiquement vrai, la notion de « pouvoir » qui est impliquée concerne les capacités ou incapacités humaines universelles, et non les capacités et incapacités d'un agent moral spécifique.

Ainsi, Sinnott-Armstrong a raison lorsqu'il affirme que, si un individu particulier – ou même un groupe d'individus – est incapable de faire quelque chose, cela ne fait pas disparaître le devoir. Dans de tels cas, il n'est pas *contradictoire* de dire que l'individu S a le devoir moral de faire X, bien que S soit incapable de faire X. Etant donné que l'incapacité n'est pas universellement partagée, elle ne fait pas disparaître le devoir.

Pour reprendre l'exemple précédent, je conserve le devoir moral de ne pas rire au sujet de la nouvelle coupe de cheveux de mon directeur, même si je suis incapable de me retenir, car les êtres humains ne sont pas universellement incapables de se retenir de rire. De la même manière, le banquier raciste conserve le devoir moral de traiter équitablement ses clients, même s'il est lui-même *incapable* de les traiter équitablement. Etant donné que son incapacité n'est pas une incapacité caractéristique de l'être humain, elle ne permet pas de supprimer le devoir moral.

Cependant, il peut sembler, à première vue, que cette réponse des partisans de la « contrainte humaniste » fragilise l'interprétation *conceptuelle* du principe DIP. En effet, s'il peut sembler crédible que nous interprétions « devoir » comme impliquant « pouvoir », il ne semble pas très crédible que nous interprétions « devoir » comme impliquant « pouvoir-relatif-à-l'être-humain ». Si l'avantage de la preuve de l'analyse conceptuelle de DIP consistait dans le fait qu'elle reposait sur un contenu sémantique *mince* (*thin*), cet avantage est désormais perdu, puisque « possible » signifie « possible-pour-l'être-humain ». Le contenu sémantique s'est épaissi et rend le principe DIP intuitivement moins convaincant.

Toutefois, la réponse des partisans de la « contrainte humaniste » ne nous semble pas illégitime. En effet, une théorie morale a pour but de déterminer le contenu des devoirs moraux *des êtres humains* en général, et non pas de certains êtres humains particuliers. Ainsi, lorsque nous utilisons la notion de « devoir », cette notion est elle-même interprétée comme relative à une catégorie d'agents – ici les êtres humains. Elle n'est pas utilisée comme portant sur des individus particuliers. Si nous acceptons que la notion de « devoir » est elle-même relative à un type d'individus, et non pas relative à un individu particulier, alors il n'est plus étonnant que la notion de « devoir moral » que nous utilisons, i.e. « devoir-moral-des-êtres-humains », implique conceptuellement la notion de « pouvoir-relatif-aux-êtres-humains ».

Le principe DIP devrait donc être reformulé de façon plus générale, pour tenir compte du fait qu'il s'applique aux êtres humains en général – ou à toute autre catégorie d'êtres – et non pas à des êtres humains particuliers – ou à un membre particulier d'une catégorie d'êtres.

DIP-Général Si les êtres du type T ont le devoir moral de faire X, alors il est possible pour les êtres du type T de faire X.

Imaginons qu'il existe un devoir moral pour les êtres humains de ne pas exprimer des pensées négatives susceptibles de blesser autrui. Etant donné qu'il n'existe pas d'incapacité humaine générale de garder ses pensées pour soi, le devoir est valide et l'existence d'un individu particulier qui ne serait pas capable de garder ses pensées pour lui n'invalide pas le principe moral général.

Imaginons maintenant qu'il existe des êtres raisonnables non humains, appelons-les « Wittgensteiniens », qui ont pour particularité que toutes leurs pensées sont automatiquement exprimées publiquement ; il leur est *impossible* de ne pas exprimer leurs pensées. Cette incapacité est inhérente à leur nature. Dans ce cas, il est *faux*, parce que contradictoire, selon l'analyse conceptuelle, de dire que les Wittgensteiniens ont le devoir moral de ne pas exprimer les pensées susceptibles de blesser autrui, quand bien même il est impossible pour les Wittgensteiniens de ne pas exprimer publiquement leurs pensées – même celles qui sont blessantes. Etant donné que « devoir-pour-les-Wittgensteiniens » implique « pouvoir-pour-les-Wittgensteiniens », le fait que les Wittgensteiniens sont incapables de ne pas exprimer leurs pensées a pour conséquences que les Wittgensteiniens n'ont pas le devoir moral de ne pas exprimer leurs pensées blessantes.

Nous pouvons appliquer le même raisonnement au sujet des limitations motivationnelles. En effet, DIP-motivation concerne les possibilités ou impossibilités motivationnelles *de l'être humain* en général et non les possibilités ou impossibilités d'un individu particulier.

DIP-motivation-général	Si les êtres du type T ont le devoir moral de faire X, alors il est possible pour les êtres du type T d'être motivés à faire X.

Nous admettrons pour le moment, donc, qu'il existe une manière crédible de défendre l'interprétation conceptuelle du principe DIP-Général. Cette interprétation permet, à l'aide de la sémantique du concept de « devoir », de justifier la « contrainte humaniste » en vertu de laquelle une théorie morale qui contredirait DIP-Général est une théorie morale invalide.

3.2. L'IMPOSSIBILITÉ DES DILEMMES MORAUX ET DE LA FAIBLESSE DE LA VOLONTÉ

La seconde objection contre l'interprétation conceptuelle de DIP-général est la suivante ; il existe des situations tout à fait ordinaires qui, si le principe DIP exprime bien une vérité conceptuelle, devraient être considérées comme contradictoires, et, donc, *impossibles*. Mais elles ne peuvent pas être contradictoires, puisque nous en faisons l'expérience dans notre quotidien. Dès lors, au lieu de les considérer comme contradictoires – et donc impossibles – il faut rejeter l'interprétation conceptuelle de DIP

Il existe deux types de situations qui devraient être considérés comme *impossibles* car *contradictoires* si le principe DIP était une vérité conceptuelle. Le premier type de situation est celui des dilemmes moraux (3.2.1.) et le second type est celui de la *faiblesse de la volonté* (3.2.2.).

3.2.1. LES DILEMMES MORAUX

Les dilemmes moraux se produisent soit lorsqu'une personne a le devoir moral à la fois de faire l'action A et le devoir moral de ne pas faire l'action A ou lorsqu'une personne a le devoir moral de faire l'action B et elle a aussi le devoir moral de faire une autre C, mais il ne lui est pas possible de faire à la fois B et C.

Concernant le premier cas de dilemmes moraux, imaginons un mari dont l'épouse souffre d'une maladie mortelle pour laquelle il n'existe qu'un seul médicament et nous savons que ce médicament fonctionne très bien. Cependant, le mari ne dispose pas de moyens financiers suffisants pour acheter le traitement de son épouse. En revanche, étant infirmier, il peut accéder à la réserve de médicaments de l'hôpital. Nous acceptons, pour l'argumentation, que le mari a le devoir moral de faire tout ce qui est possible pour sauver sa femme – encore jeune – de la mort. Il a donc le devoir moral de se procurer le médicament. Le seul moyen d'y parvenir étant de le voler, le mari a le devoir moral de voler le médicament. Cependant, étant donné qu'il est immoral de voler, notre époux a aussi le devoir moral de *ne pas* voler le médicament en question. Il a donc à la fois le devoir moral de voler et celui de ne pas voler le médicament.²³⁸

Le second type de dilemme moral est souvent illustré par l'exemple suivant (OGIEN, 2001, p. 17): deux enfants jumeaux sont en train de se noyer dans le lac, à équidistance du

²³⁸ Si nous cherchons des exemples plus ordinaires, nous pouvons penser au devoir moral de travailler – et de se déplacer au travail – et au devoir moral de ne pas polluer. Il est possible qu'une personne ait le devoir moral de conduire sa voiture pour satisfaire son obligation morale de travailler, et qu'elle ait simultanément le devoir moral de ne pas conduire sa voiture pour satisfaire son obligation morale de ne pas polluer. Elle a le devoir à la fois de conduire et de ne pas conduire son véhicule. Nous pourrions aussi reprendre le cas proposé par Lemmon et repris par Ogien : « L'exemple de Lemmon, que j'ai adapté, lui-même adapté de Platon, est approximativement le suivant. Un ami vous laisse un fusil de chasse en dépôt. Vous promettez de le lui rendre dès qu'il repassera le chercher. Mais lorsque cet ami revient pour récupérer son arme, il est dans un état de totale dépression et vous craignez qu'il s'en serve pour autre chose que du gibier (sa femme, ses enfants, lui-même, etc.). Dans un cas de ce genre, dit Lemmon, on doit et on ne doit pas restituer le fusil. (LEMMON, 1962, p. 150) » (OGIEN, 2001, p. 18). Pour finir, de tels conflits d'obligation apparaissent, selon nous, de manière assez flagrante dans l'éthique environnementale ; il est possible en effet que le devoir moral de protéger l'environnement soit incompatible avec le devoir moral d'aider les populations pauvres à sortir de la pauvreté, que la préservation de l'environnement soit incompatible avec la prospérité économique, etc.

bord. La seule personne présente sur la plage est un touriste allemand. Ce touriste est capable de nager et ne prend pas de risques en essayant de venir en aide aux enfants, mais il lui est impossible de sauver les deux enfants. Le problème est alors le suivant : le touriste a le devoir moral de venir en aide au premier enfant. En outre, étant donné que la situation du second enfant est exactement la même que celle du premier, le touriste a également le devoir moral de sauver la vie du second enfant. Néanmoins, le touriste ne peut pas les sauver tous les deux. En conséquence, peu importe ce que choisit de faire notre touriste allemand, il échouera à satisfaire une de ses obligations morales ; s'il sauve le premier enfant, il échoue à sauver le second. S'il sauve le second, il échoue à sauver le premier.

Si nous acceptons le principe DIP, de telles situations de dilemme moral sont *conceptuellement impossibles*. Il est *conceptuellement impossible* que le mari ait le devoir moral de voler et de ne pas voler le médicament et il est conceptuellement impossible que le touriste ait le devoir moral de sauver les deux jumeaux.²³⁹ Toutefois, comme le montrent ces exemples, les dilemmes moraux sont possibles, puisqu'il existe, dans les faits, de véritables situations de dilemmes moraux. Il existe donc des situations dans lesquelles nous avons le devoir de faire des actions qu'il nous est impossible de faire. Cela implique en conséquence que l'interprétation conceptuelle de DIP doit être rejetée.

Le raisonnement à l'encontre de l'interprétation conceptuelle de DIP se présente donc de la manière suivante (INNALA, 2000, p. 403):

1. Il existe des situations de véritables dilemmes moraux.²⁴⁰
2. S'il existe des situations de véritables dilemmes moraux, alors la logique déontique doit pouvoir les représenter.

²³⁹ La contradiction logique ne provient pas du fait qu'il serait contradictoire d'avoir le devoir de faire quelque chose et son contraire. Les énoncés déontiques $O(A)$ et $O(\neg A)$ ne sont pas en soi contradictoires (contrairement aux énoncés $O(A)$ et $\neg O(A)$).

Même à l'aide du principe DIP, ils ne deviennent pas contradictoires, lorsque traités séparément ; $O(A)$ implique qu'il est possible de faire A et $O(\neg A)$ implique qu'il est possible de faire non- A , ce qui n'est pas contradictoire. C'est l'obligation portant sur la conjonction des actions qui implique contradiction : $O(A \wedge \neg A)$ implique une contradiction parce qu'elle implique qu'il est possible de faire $A \wedge \neg A$, donc qu'il existe un monde où A et $\neg A$ sont réalisées, ce qui est contradictoire (SAYRE-MCCORD, 1986, p. 181). Nous reviendrons sur ces considérations dans le dernier chapitre du travail.

²⁴⁰ Les dilemmes moraux véritables sont ceux qui n'ont pas seulement l'apparence d'être insolubles, mais ceux qui le sont vraiment. Nous utilisons donc les expressions de « dilemmes moraux véritables » et « dilemmes moraux insolubles » comme étant équivalentes. Un dilemme moral insoluble se décrit de la manière suivante : « [D]ans le cas des dilemmes moraux [insolubles], l'agent ne peut éviter de manquer à l'un de ses devoirs et il n'est rien qu'il aurait pu et qu'il aurait dû faire auparavant pour éviter de se trouver dans cette situation » (BÉLANGER, 2011, p. 16).

3. La logique déontique ne parvient pas à représenter les situations de véritables dilemmes moraux, étant donné que la logique déontique accepte DIP parmi ses axiomes.
4. ∴ La logique déontique n'est pas appropriée.

3.2.2. LA FAIBLESSE DE LA VOLONTÉ

Si le principe DIP exprimait une vérité conceptuelle, alors il faudrait considérer comme conceptuellement impossible un deuxième type de situation dont nous faisons pourtant l'expérience dans le monde actuel. Il s'agit de situations de *faiblesse de la volonté*. Les cas de faiblesse de la volonté sont très répandus – peut-être plus encore que les cas de dilemmes moraux ; il s'agit de situations dans lesquelles nous savons ce que nous devons faire, mais nous ne parvenons cependant pas à nous motiver à le faire. Nous savons que nous devons consommer et voyager moins, que nous devons recycler, que nous devons faire des actes de charité, que nous devons être sincères en tout temps avec nos amis, mais nous ne pouvons pas nous motiver à agir en conformité avec notre jugement. Le problème est alors le suivant ; si DIP est une vérité conceptuelle, alors il implique que dans les situations de faiblesse de la volonté, nos devoirs moraux *s'évanouissent*. Puisque nous ne pouvons pas à nous motiver à recycler, à faire un don, à diminuer nos voyages de loisir, alors nous n'avons pas le devoir moral de recycler, de faire un don ou de diminuer nos voyages de loisir. Or, il paraît absurde que, du seul fait que nous ne puissions pas à nous motiver à faire quelque chose, alors nous n'avons pas le devoir moral de la faire.

Pour pouvoir maintenir l'interprétation conceptuelle de DIP, il faudrait alors conclure que les situations de faiblesse de la volonté sont en fait *impossibles*²⁴¹ ; si nous avons le devoir de donner à la charité, de recycler et de diminuer nos voyages de loisir, alors nous sommes capables de nous motiver à donner à la charité, à recycler et à diminuer nos voyages de loisir. Malgré les expériences de situations qui *ont l'apparence* d'un cas de faiblesse de la volonté, celle-ci est en fait conceptuellement impossible ; nous n'avons donc jamais d'expérience *véridique* d'un cas de faiblesse de volonté dans l'accomplissement de nos devoirs. Autrement dit, si DIP est une vérité conceptuelle, soit les devoirs moraux

²⁴¹ Une autre manière serait d'adopter une version *conditionnelle* de DIP: il faudrait admettre que « devoir-implique-pouvoir », *sauf* en cas de faiblesse de la volonté. En cas de faiblesse de la volonté, le devoir serait maintenu, malgré l'incapacité de se motiver à agir. Toutefois, ce procédé nous semble *ad hoc*.

dépendent de la motivation, soit les situations de faiblesse de la volonté sont tout simplement impossibles – malgré les apparences. Etant donné que ces deux conséquences sont inacceptables, il vaut mieux, selon les critiques, renoncer à l'interprétation conceptuelle de DIP (OGIEN, 2001, p. 20).

En résumé, si le principe DIP exprime une vérité conceptuelle, alors les dilemmes moraux et la faiblesse de la volonté sont (logiquement) impossibles. Or, ils sont possibles – i.e. nous en faisons l'expérience. Donc DIP n'est pas une vérité conceptuelle. Comme le dit Ruwen Ogien :

[...] il n'y a pas de lien interne ou conceptuel entre « devoir » et « pouvoir ». On peut très bien concevoir des devoirs auxquels il est impossible de se conformer pour des raisons psychologiques ou parce qu'ils sont en conflit avec d'autres devoirs, mais qui restent, néanmoins, des devoirs authentiques (OGIEN, 2001, p. 20).

3.3. LA VIOLATION DE LA LOI DE HUME

La dernière critique formulée contre DIP est que, si le principe exprime une vérité conceptuelle au sujet du devoir, cela implique qu'un énoncé normatif peut être déduit à partir de prémisses exclusivement descriptives. Ainsi, l'argument suivant deviendrait (sémantiquement) valide :

(P1) S ne peut pas faire X.	Prémisse descriptive
(C1) ∴ S n'a pas le devoir moral de faire X	Conclusion normative

Etant donné que le principe DIP exprime une vérité conceptuelle, analytiquement vraie, la signification des termes employés dans (P1) permet de déduire directement (C1). La sémantique de la notion du « devoir » autorise donc une transition directe d'une prémisse exclusivement descriptive à une conclusion normative, sans devoir introduire une prémisse normative intermédiaire.

Cette transition directe du descriptif au normatif pose des difficultés car elle transgresse la fameuse loi de Hume, selon laquelle il est logiquement interdit de dériver une conclusion normative sur la base de prémisses exclusivement descriptives. Ainsi, selon

notamment Ruwen Ogien, étant donné que DIP viole la loi de Hume, largement acceptée, il vaut mieux renoncer à DIP plutôt que renoncer à la loi de Hume (OGIEN, 2001).

Cette critique nous paraît néanmoins inappropriée, étant donné que l'impossibilité *logique* d'établir un argument valide partant de prémisses exclusivement descriptives pour obtenir une conclusion normative n'exclut aucunement qu'une transition *sémantique* soit possible. Si nous soutenons que le principe DIP est analytiquement vrai, le fait qu'il n'exprime pas une vérité purement logique n'a pas d'importance.

Le principe DIP n'est d'ailleurs pas le seul « principe sémantique de transition » (PST) qui a été proposé.²⁴² Nous l'avons vu, d'autres principes sémantiques permettent d'établir une transition conceptuelle entre les domaines descriptif et normatif ; (i) l'identité conceptuelle (« M signifie N »), (ii) la promesse (« il faut tenir ses promesses »), et (iii) le bien attributif (« être vertueux » se définit par « avoir les caractéristiques propres à l'être humain »). Dans chacun de ces trois cas, ainsi que dans le cas de DIP, les philosophes²⁴³ qui ont développé une interprétation analytique de ces principes avaient pour but précisément de montrer que : *bien que la transition reste invalide sur le plan syntaxique, et qu'elle viole donc la loi de Hume, elle est néanmoins légitime sur le plan sémantique*. La *signification* des termes autorise une transition directe du descriptif au normatif. Face à une telle thèse, il est parfaitement inapproprié (et inoffensif) de rétorquer que la validité sémantique de DIP viole, sur le plan logique, la loi de Hume.

3.4. L'ABSENCE DE NEUTRALITÉ NORMATIVE DE LA LOGIQUE DÉONTIQUE

L'argument qui, selon nous, est le plus décisif contre l'interprétation conceptuelle du principe DIP est le suivant : la logique déontique a pour fonction d'identifier le contenu conceptuel minimal des notions normatives, c'est-à-dire les invariables conceptuels, tout en restant neutre à l'égard de la nature du bien et du mal, c'est-à-dire à l'égard des critères substantiels permettant de séparer les mondes possibles moralement positifs et les mondes possibles moralement négatifs. Comme le souligne Jean-Louis Gardies :

²⁴² Pour rappel, un PST est un principe qui affirme l'existence d'un lien conceptuel entre une notion descriptive et une notion normative. Ces principes étant analytiquement vrais, ils rendent possible une transition sémantique entre des énoncés exclusivement descriptifs et une conclusion normative.

²⁴³ Pour la référence de ces auteurs, nous renvoyons le lecteur au point (1.3.5.) du chapitre de « L'erreur moraliste » du présent travail.

Ce que permettent les sémantiques des mondes possibles est en effet quelque chose de si simple qu'il risque même de laisser à l'observateur l'impression d'un tour de passe-passe philosophique : en assimilant l'obligatoire à ce qui est vrai dans tous les mondes possibles reconnus comme positifs, il renvoie en quelque sorte l'élément proprement normatif à un autre niveau et permet ainsi de raisonner en termes strictement extensionnels. [...] [II] se dégage ainsi la possibilité d'une procédure extensionnelle, grâce à cette sorte de rejet du problème philosophique à un autre étage. La question de savoir ce qui est bien ou mal reste entière ; mais si l'on admet que l'attitude morale consiste à exercer dans l'ensemble des mondes possibles une partition entre ces deux sous-ensembles, quelles que puissent être les frontières qui président à cette partition (car c'est là que réside le relativisme), on se rend compte que la rationalité de nos raisonnements déontiques ne dépend pas de la situation de ces frontières du bien et du mal. La preuve en est que nous pouvons en donner ces schémas extensionnels (GARDIES, 1987, pp. 44-45).

Autrement dit, la logique déontique a pour fonction de *décrire* le langage moral, et sa structure *a priori*, sans présupposer d'engagements normatifs particuliers. Un système de logique déontique qui *favoriserait* une position morale plutôt qu'une autre ne pourrait pas prétendre avoir identifié la structure logique commune à tout système moral. Il ne constituerait donc pas un système logique adéquat.²⁴⁴

²⁴⁴ Les systèmes logiques de la logique déontiques sont aussi parfois accusés d'être incompatibles avec certaines positions métaéthiques ; étant donné notamment qu'ils attribuent des conditions de vérité aux énoncés moraux, les systèmes de logique déontique paraissent exclure le non-cognitivisme. Toutefois, des non-cognitivistes aussi important que R. M. Hare et von Wright ont soutenu la possibilité de faire une logique déontique dans le cadre d'une métaéthique non cognitiviste. Pour cela, il suffit de considérer l'énoncé « S doit faire X » comme décrivant l'existence d'une norme (d'un ordre) ordonnant à S de faire X. Si la valeur de vérité est attribuée à l'énoncé plutôt qu'à la norme elle-même, l'énoncé étant vrai si la norme existe ou si l'ordre a été donné, alors il devient tout à fait possible de construire une logique déontique pour une doctrine métaéthique non cognitiviste. Cependant, dans ce cas, il devient difficile d'introduire des foncteurs proprement déontiques : étant des descriptions de l'existence d'une norme, de tels énoncés admettent uniquement les règles de la logique propositionnelle. Autrement dit, il n'y a plus de règles logiques spécifiques au domaine normatif. Comme l'affirme Innala : « Considérons l'énoncé déontique typique de la forme « il doit être le cas que... ». Selon la première interprétation prescriptive, cet énoncé est considéré comme une norme ou une obligation de se comporter d'une certaine manière. Selon la seconde lecture, descriptive, l'énoncé est interprété comme l'affirmation qu'il existe une norme ou une obligation d'agir d'une certaine manière. Les deux lectures sont considérées comme radicalement différentes parce que l'existence de normes et d'obligation est un fait social observable alors que les normes ou obligations elles-mêmes sont considérées comme étant incapables de recevoir des valeurs de vérité. A fortiori, la logique déontique est aujourd'hui interprétée non plus comme la logique des normes ou obligations elles-mêmes, mais comme la logique des propositions-normatives, i.e. des phrases contenant le concept de « devoir » interprétée de façon descriptive.

Le développement de la logique déontique reflète un désir de saisir le fondement commun, la structure sous-jacente, de notre discours moral. Il a été motivé par un « principe de neutralité déontique » qui soutient que la logique déontique doit être neutre entre différentes théories morales concurrentes, et qui promet d’offrir un terrain de combat neutre pour le débat moral.²⁴⁵

Le problème provient du fait que les énoncés de la logique déontique *ont* des conséquences normatives ; le principe DIP, notamment, a pour conséquence d’exclure un certain nombre de positions normatives pourtant considérées comme étant des positions morales cohérentes. Comme le fait remarquer Charles Pigden :

Une logique déontique doit fournir un cadre à l’intérieur duquel n’importe quelle morale cohérente doit pouvoir être exprimée, même si elle est repoussante. [L’]incapacité [de DIP] à s’appliquer à toutes les théories démontre que Devoir-Implique-Pouvoir n’est pas une vérité logique. Parce que toutes les vérités de la logique déontique sont *tenues de* s’appliquer à toutes les théories normatives.²⁴⁶

Par exemple, DIP exclut les doctrines morales qui admettent la possibilité de dilemmes moraux. Le principe exclut aussi, de manière plus générale, toutes les théories normatives qui contiennent des exigences morales qu’il est impossible d’accomplir – qu’elles apparaissent sous forme de dilemme moral ou autrement.

[...] Cependant, toutes les versions de cette vue largement admise ont le même défaut fondamental, à savoir, que l’interprétation descriptive ne génère pas de relation logique (par exemple des théorèmes déontiques) qui sont spécifiques aux énoncés déontiques. » [“Consider a typical deontic sentence of the form ‘It ought to be the case that...’. On the first prescriptive reading it is interpreted as a norm or an obligation to behave in a certain way. On the second descriptive reading it is interpreted as a statement that there exists a norm or an obligation to behave in that way. These two readings are said to be radically different because the existence of norms or obligations is an uncontroversial social fact whereas norms or obligations themselves are thought to be incapable of assuming truth-values. A fortiori, deontic logic is now interpreted, not as a logic of norms or obligations themselves, but as a logic of norm-propositions, i.e. descriptively interpreted ought-sentences. However, all versions of this widely held view seem to have the same fundamental flaw, namely, that the descriptive interpretation does not generate logical relations (e.g. deontic theorems) that are in any sense peculiar to deontic sentences.” (INNALA, 2000, p. 404)]

²⁴⁵ “The development of deontic logic reflects a desire to capture the common ground, the underlying structure, of our moral discourse. It has been driven by a “principle of deontic neutrality” which holds that deontic logic ought to be neutral between competing moral theories, and which promises a neutral battleground for moral debate.” (SAYRE-MCCORD, 1986, p. 179)

²⁴⁶ “[A] deontic logic should provide a framework within which any coherent morality can be expressed, no matter how repellent. [The] failure to cut across all theories demonstrates that Ought-Implies-Can is not a logical truth. For the truths of deontic logic are required to cut across all moral theories” (PIGDEN, 1990, p. 15).

Nous avons vu qu'il existe de nombreuses doctrines morales qui peuvent être accusées d'exiger l'impossible sur le plan motivationnel. En outre, il existe même des doctrines morales qui contiennent des normes qu'il est physiquement, ou même métaphysiquement, impossible à respecter. C'est le cas, notamment, des doctrines morales de Luther²⁴⁷ et Calvin, qui admettent la prédestination divine (QUILLIET, 2007):

Le principe que « devoir » implique « pouvoir » fait face aux doctrines morales des religions qui, comme le Calvinisme, admettent la prédestination. Puisque ces doctrines sont certainement aussi soutenables que les doctrines du péché originel²⁴⁸, le principe de neutralité déontique exige d'abandonner la supposition que ce qui est obligatoire est possible.²⁴⁹

Il est soutenu par Schotch et Jennings (1981) qu'« il est admis parmi les philosophes moraux » que « devoir » doit être interprété de telle manière à valider Devoir-implique-(logiquement)-pouvoir (p. 149) » – certainement parce que « ce principe s'applique à toutes les théories morales (p. 156). Mais l'existence du système moral de Luther démontre que Devoir-Implique-Pouvoir n'est *pas* une vérité logique.²⁵⁰

Ainsi, il est possible de refuser à DIP le statut de vérité conceptuelle par le biais du raisonnement suivant :

1. Les principes logiques admis en logique déontique ne doivent pas exclure certaines doctrines morales.

²⁴⁷ « Il est montré à l'homme, par de telles paroles, ce qu'il doit, et non ce qu'il peut. Il est dit à Caïn qu'il doit dominer son péché et tenir son désir au-dessous de lui. Mais cela, il ne l'a pas fait, et il ne l'a pas pu, en tant qu'il était déjà maîtrisé par un pouvoir souverain qui n'était pas le sien. » Cette citation de Luther provient du livre de Bernard Quilliet (2007, p. 265). Elle confirme le refus, explicite, par Luther, du principe DIP.

²⁴⁸ La doctrine catholique du péché originel reste compatible avec le fait que l'homme *peut* faire le bien, s'il le décide. En effet, bien que le péché originel ait pour conséquence que l'homme est par nature mauvais, Dieu lui accorde gratuitement la *grâce*, c'est-à-dire la capacité de faire le bien. L'homme peut ensuite accepter ou refuser cette aide divine (QUILLIET, 2007, p. 265). Avec la prédestination telle qu'elle est acceptée par Luther et Calvin, il ne reste plus de place pour la participation de l'individu. Dieu a décidé et déterminé à l'avance quels sont les hommes qui feront le bien et ceux qui ne le feront pas.

²⁴⁹ "The principle that « ought » implies « can » runs head-on into the ethical doctrines of those religions, such as Calvinism, which are committed to predestination. Since these doctrines are surely as tenable as the doctrines of the original sin, the principle of deontic neutrality demands abandoning the assumption that what is obligatory is possible." (SAYRE-MCCORD, 1986, p. 188).

²⁵⁰ "[I]t is urged by Schotch and Jennings (1981) that 'it is accepted among moral philosophers' that 'Ought' should be so interpreted as to validate Ought-Implies-(logically)-Can (p. 149) – presumably because 'that principle cuts across all moral theories' (p. 156). But the very existence of Luther's moral system demonstrates that this is false." (PIGDEN, 1990, p. 15).

2. Le principe DIP exclut certaines doctrines morales – notamment celles qui admettent l’existence de véritables dilemmes moraux ainsi que les doctrines théologiques de la prédestination.
3. ∴ Le principe DIP ne doit pas être considéré comme un principe de logique déontique.

Le fait que le principe DIP n’est pas normativement neutre, puisqu’il exclut des théories morales pourtant cohérentes, montre bien pourquoi, selon l’objection, ce principe ne peut pas être traité comme une vérité de *logique déontique*.²⁵¹

Selon les critiques de la logique déontique, DIP n’est pas le seul axiome qui viole le principe de neutralité axiologique. Tous les axiomes de la logique déontique auraient, selon eux, des conséquences normatives. De ce fait, selon les critiques, il faut conclure que la logique déontique ne peut pas être une discipline indépendante de la philosophie morale normative.

Cela ne signifie toutefois pas que la logique déontique doive être abandonnée ; si la logique déontique devait être abandonnée sous prétexte qu’elle n’est pas normativement neutre, alors il faudrait également abandonner la logique aléthique et la logique temporelle, puisque les axiomes de l’une et l’autre ne sont pas neutres concernant la *métaphysique* des mondes possibles ou la métaphysique du temps (INNALA, 2000, pp. 397-98). La reconnaissance du caractère non neutre des axiomes de la logique déontique doit plutôt nous inciter à accepter la conséquence suivante : il est nécessaire de fournir des arguments

²⁵¹ La critique ne concerne d’ailleurs pas uniquement les systèmes déontiques qui admettent le principe DIP. Comme le fait remarquer Sayre-McCord, n’importe quelle sémantique des concepts déontiques – ou, plus généralement, des concepts normatifs, a des conséquences normatives, aussi minimalistes soient-elles. De ce fait, lorsque nous exigeons la neutralité axiologique de la logique déontique, nous excluons en fait toute logique déontique. « Parce que nos concepts déontiques sont sujets à controverse de bout en bout, presque rien ne survit [à l’exigence de neutralité] – bien que communément utilisé pour des attaques chirurgicales, l’exigence de neutralité axiologique est plutôt appropriée pour un bombardement dévastateur » [“Because our deontic concepts are controversial through and through, almost nothing survives-though commonly used for surgical strikes, the principle of deontic neutrality is better suited for devastation bombing.” (SAYRE-MCCORD, 1986, p. 180)]. Le fait qu’il existe des logiques déontiques capables d’être neutres concernant la possibilité des dilemmes moraux ne signifie donc pas que ces doctrines ont atteint la neutralité normative. « L’application du principe de neutralité déontique nous laisse donc sans logique déontique [...] Tout ce qui reste de la logique déontique, une fois épurée pour ne pas contredire le principe de neutralité déontique, est la logique propositionnelle, le modus ponens, et l’affirmation que tout ce qui est obligatoire sous certaines circonstances reste obligatoire dans n’importe quelle circonstances logiquement équivalentes. » [“Consistent enforcement of the principle of deontic neutrality effectively leaves us without a deontic logic. [...] All that remains of deontic logic, once it is trimmed so as not to offend the principle of deontic neutrality, is propositional logic, modus ponens, and the claim that anything which is obligatory under certain circumstances is obligatory under all logically equivalent circumstances.” (SAYRE-MCCORD, 1986, p. 193)].

indépendants de la logique déontique en faveur ou en défaveur de l'acceptation de chaque principe déontique.

Nous avons vu quatre objections contre l'interprétation conceptuelle de DIP : (1) A l'objection des contre-exemples, nous avons dit que les partisans de la contrainte humaniste pouvaient répondre par le fait que DIP ne concerne que les capacités et incapacités relatives à la nature humaine et ne concerne pas les capacités ou incapacités individuelles. Nous avons ensuite présenté deux autres objections : (2) l'impossibilité des dilemmes moraux et (3) l'impossibilité de la faiblesse de la volonté. Nous réservons les réponses à ces critiques au chapitre suivant. Pour finir (4), nous avons vu que DIP pose problème car il ne respecte pas l'idéal de neutralité déontique. Nous avons vu, toutefois, que seul un abandon complet de la logique déontique pourrait permettre d'éviter *absolument* toute conséquence normative. Or, cela semble un prix difficile à payer.

Pour le moment, nous allons supposer que l'interprétation conceptuelle de DIP est problématique, et nous allons donc étudier les solutions alternatives. Or, nous verrons que les autres interprétations posent aussi des difficultés.

4. QUELLE INTERPRÉTATION DE DIP ?

Dans la section précédente, nous avons présenté quelques raisons de douter de l'interprétation de « devoir-implique-pouvoir » comme étant un principe analytique ou sémantique. Le principe DIP imprègne néanmoins profondément le discours normatif ordinaire. Il semble que nous l'acceptons implicitement dans notre manière ordinaire de parler. Il ne suffit donc pas de rejeter le principe sur le plan conceptuel. Il faut pouvoir expliquer pourquoi, bien que le principe ne soit pas vrai en vertu de la seule analyse des concepts, il reste très largement admis dans le discours normatif. Plusieurs solutions peuvent être envisagées pour rendre compte de l'acceptation implicite du principe DIP dans le langage normatif. Chacune de ces solutions offre l'avantage, non négligeable, d'être compatible avec l'existence de situations de dilemmes moraux et de faiblesse de la volonté.

4.1. LA PRÉSUPPOSITION

Une première manière d'interpréter la relation entre les notions de « devoir » et de « pouvoir » consiste à la considérer comme une relation de *présupposition* : ainsi, lorsque

nous faisons usage du langage normatif, nous *présupposons* que l'action sur laquelle porte le langage normatif peut être réalisée. Dans le cas de la relation de présupposition – contrairement au cas de la relation *d'implication conceptuelle* – le fait qu'une action est impossible à accomplir n'a pas pour conséquence de rendre *faux* l'énoncé normatif qui affirme que l'action en question est obligatoire. Le fait qu'une action est impossible à accomplir a plutôt pour conséquence que la question « l'action est-elle obligatoire ? » ne se pose même pas.²⁵² L'énoncé normatif affirmant qu'une action impossible est obligatoire n'est donc ni vrai ni faux, puisque la question de sa vérité ou fausseté n'apparaît même pas.

Interprétation de DIP	Valeur de vérité de l'énoncé « S a le devoir moral de faire X » lorsqu'il est impossible pour S de faire X
Conceptuelle	FAUX
Présupposition	NI VRAI, NI FAUX

Cette interprétation de la relation entre « devoir » et « pouvoir » comme étant celle de la présupposition pose toutefois une difficulté, soulevée par Sinnott-Armstrong. En effet, il peut exister différentes raisons pour lesquelles un énoncé normatif est *faux*. Par exemple, l'énoncé « S doit sauter plus haut que la Lune » peut être faux parce qu'il n'y a pas de raisons de sauter plus haut que la Lune. Certes, il est *en outre* impossible pour S de sauter plus haut que la Lune, mais cette impossibilité n'est pas la raison pour laquelle l'énoncé normatif « S doit sauter plus haut que la Lune » est faux. La raison pour laquelle l'énoncé normatif est faux est qu'il n'existe pas de raisons d'obliger quelqu'un à sauter plus haut que la Lune. L'énoncé normatif « S a le devoir de sauter plus haut que la Lune » resterait donc *faux*, même s'il était possible pour S de sauter plus haut que la Lune.

La possibilité qu'un énoncé normatif soit *faux* indépendamment de la possibilité ou de l'impossibilité de faire l'action pose une difficulté pour l'interprétation de DIP en termes de *présupposition*. En effet, selon cette interprétation, si une action est impossible à accomplir, alors l'énoncé normatif selon lequel cette action est obligatoire n'est *ni vrai ni faux*. Mais

²⁵² Pigden propose de comparer ce cas avec celui de l'énoncé « Le roi de France est sage » : « Si quelqu'un affirme que « Le roi de France est sage », il présuppose que le roi de France existe. En revanche, si nous admettons que le roi de France n'existe pas, alors il n'est pas faux qu'il est sage. La question « Le roi de France est-il sage ? » ne se pose même pas. ["If it is true that the King of France is wise, he exists. But if it is false that he exists, it is not false that he is wise. Rather, the supposition lacks truth value: 'the question does not arise'." (PIGDEN, 1990, p. 3)].

nous avons vu que, dans certains cas, il est *faux* que S a le devoir moral de sauter plus haut que la Lune, bien qu'il soit *impossible* pour S de faire son devoir ; dans de tels cas, la fausseté de l'énoncé s'explique par autre chose que par l'impossibilité de faire l'action et l'énoncé *resterait faux* même s'il était possible de faire l'action.

Si « devoir » présupposait « pouvoir », alors il ne pourrait jamais être *faux* qu'un agent doit faire ce qu'il est impossible de faire. Mais il est parfois faux [qu'un agent doit faire ce qu'il est impossible de faire]. S'il n'existe pas de raison pour Smith de sauter plus haut que la Lune, alors il est faux que Smith a le devoir de le faire, puisque cela resterait *faux* même si Smith pouvait sauter aussi haut que cela.²⁵³

Autrement dit, il est possible qu'un énoncé normatif portant sur une action impossible à accomplir soit *faux*, ce que la présente interprétation de DIP n'autorise pas – puisque, selon elle, l'énoncé devrait toujours être ni vrai ni faux. L'interprétation de DIP en termes de présupposition est donc erronée.

4.2. L'INEFFICACITÉ DU DISCOURS MORAL

Les deux prochaines interprétations de DIP reposent sur une distinction entre, d'un côté, la validité d'un devoir moral et, de l'autre côté, l'attribution de la *responsabilité*. Selon ces deux interprétations, l'incapacité de faire une action ne supprime pas un devoir moral. En revanche, une incapacité supprime la *responsabilité* de la faute morale. Puisque l'individu est *incapable* de faire X, alors il n'est pas *responsable* de ne pas avoir fait X. De ce fait, bien que la personne ait toujours le devoir moral de faire X, il est inapproprié de la *blâmer* pour avoir échoué à faire son devoir. Selon la première interprétation, le blâme a pour fonction de modifier le comportement. Le blâme est donc *inutile* et, de ce fait, inapproprié, lorsqu'il n'est pas possible de modifier le comportement de la personne. Selon la seconde interprétation, le blâme est inapproprié en cas d'incapacité, car il est injuste de blâmer une personne et *immoral* de la sanctionner pour une action qu'elle ne pouvait pas accomplir.

Pour Sinnott-Armstrong, la relation appropriée entre « devoir » et « pouvoir » est celle de l'implication *conversationnelle*. Selon cette dernière interprétation de DIP, le principe est

²⁵³“If 'ought' presupposed 'can', then it could never be false that an agent ought to do what cannot be done. But it can be false. If there is no reason for Smith to jump over the moon, then it is false that Smith ought to do so, since it would not be true even if Smith could jump that high.” (SINNOTT-ARMSTRONG, 1984, p. 254)

vrai en vertu de l'*usage* ou des *effets* que nous cherchons à obtenir à l'aide du langage normatif.

Afin de tester si « devoir » implique conversationnellement « pouvoir », nous devons déterminer le but des énoncés qui utilisent « devoir ». Ce test est compliqué, puisque « devoir » peut être utilisé pour plusieurs fonctions. « Devoir » peut être utilisé pour conseiller des actions actuelles ou pour blâmer un agent pour des actions passées ou pour délibérer au sujet du futur.²⁵⁴

Si les notions normatives sont utilisées pour exprimer le conseil ou le blâme, par exemple, alors elles ne peuvent pas remplir leur fonction si l'agent conseillé ou blâmé ne peut de toute façon pas accomplir l'action. Dans un tel cas, l'agent a toujours le devoir prudentiel ou moral de faire l'action, mais il est inutile de lui dire qu'il *devrait* le faire (conseil) ou qu'il *doit* le faire (ordre), s'il est impossible pour l'agent d'adapter son comportement au conseil ou à l'ordre. Par exemple, si S a le devoir de faire deux actions incompatibles, et qu'il ne lui est donc pas possible de satisfaire à toutes ses obligations, cela ne signifie pas qu'il n'a pas le devoir de faire ces deux actions. Cela signifie seulement qu'il est inutile de faire usage du langage normatif dans ce contexte, que ce soit dans le but de conseiller ou dans le but de blâmer la personne.

Puisque devoir implique au maximum conversationnellement pouvoir, il peut être vrai qu'un agent ait le devoir de faire les deux actions, bien qu'il soit inutile de le lui dire dans le but de le conseiller, et bien que les agents ne soient pas blâmables s'ils ne se sont pas mis eux-mêmes en situation de dilemme.²⁵⁵

Sinnott-Armstrong utilise cet argument, par exemple, pour répondre au relativisme métaéthique. En effet, selon le relativisme métaéthique, certaines communautés *ne peuvent pas* vouloir accomplir certaines actions, et cela a pour conséquence qu'elles n'ont pas le devoir de les accomplir. Le principe DIP, selon les relativistes culturels, servirait de preuve

²⁵⁴ "In order to test whether 'ought' conversationally implies 'can', we must determine the purposes of utterances using 'ought'. This task is complicated, since 'ought' can be used for many purposes. 'Ought' can be used to advise present acts or to blame an agent for past acts or to deliberate about the future." (SINNOTT-ARMSTRONG, 1984, p. 257)

²⁵⁵ "Since 'ought' at most conversationally implies 'can', it can be true that an agent ought to do both acts, even though saying so cannot serve purposes such as advising, and the agents are not blameworthy unless they got themselves into the dilemma." (SINNOTT-ARMSTRONG, 1984, p. 260)

pour montrer que les vérités morales varient d'une communauté à l'autre et que le relativisme moral est donc vrai.

Toutefois, selon Sinnott-Armstrong, le fait que certaines cultures ne peuvent pas envisager d'accomplir certaines actions – i.e. qu'il leur est *impossible* de vouloir accomplir certaines actions – n'implique pas qu'elles n'ont pas le *devoir* moral de faire ces actions. Leur incapacité rend seulement inadapté l'usage du langage moral dans le but de les conseiller ou de les blâmer, mais elle ne rend pas *invalide* le devoir moral.

L'affirmation que « devoir » implique « pouvoir » constitue une prémisse de l'argument suivant en faveur du relativisme culturel : si un agent est endoctriné dans la culture des Iks, alors l'agent ne peut pas concevoir le fait de s'occuper de ses enfants au-delà de l'âge de trois ans, et il est de ce fait incapable de s'en occuper. S'il ne peut pas s'en occuper, alors il n'est pas vrai qu'il a le devoir de s'en occuper. Mais les parents dans notre culture ont le devoir de s'occuper de leurs enfants au-delà de l'âge de trois ans. En conséquence, le fait que les parents ont le devoir de s'occuper de leurs enfants est relatif à la culture. [...] Cependant, de nombreux anthropologistes ont le sentiment, avec raison, que cette réponse est sévère et qu'elle sous-estime la force de l'endoctrinement culturel. Une meilleure réponse consisterait à dire que, étant donné que « devoir » au maximum implique conversationnellement « pouvoir », même les parents Iks ont le devoir de s'occuper de leurs enfants, mais il est inutile de le leur dire, et ils ne sont pas blâmables de ne pas le faire. Ainsi, nous pouvons reconnaître que la culture joue un rôle pertinent sans néanmoins admettre le relativisme culturel.²⁵⁶

Si nous acceptons cette conclusion, alors il est faux qu'un individu cesse d'être soumis à un devoir moral du seul fait qu'il ne peut pas l'accomplir – que cette incapacité soit métaphysique, physique, psychologique ou motivationnelle. L'unique conséquence de

²⁵⁶ “[T]he claim that 'ought' implies 'can' is a premise in the following argument for moral relativism: if an agent is indoctrinated into Ikien culture, then the agent cannot conceive of caring for her children after the age of three, so she cannot care for them. If she cannot care for them, then it is not true that she ought to care for them. But parents in our culture ought to care for their children well beyond the age of three. Therefore, whether or not parents ought to care for their children is relative to the culture. [...] However, many anthropologists feel, with some reason, that this response is harsh and underestimates the force of cultural indoctrination. A better response is that, since 'ought' at most conversationally implies 'can', even Ikien parents ought to care for their children, but there is no point in saying so to their faces, and they are not blameworthy for not doing so. Thus, one can recognize the relevance of culture without admitting moral relativism.” (SINNOTT-ARMSTRONG, 1984, p. 261).

l'incapacité à agir ou à être motivé à agir consiste dans le fait qu'il est inutile de faire usage du langage moral pour modifier ou influencer le comportement et que les personnes cessent d'être *blâmables* pour n'avoir pas respecté leur devoir. Comme le montre le tableau ci-dessous, l'incapacité d'un individu n'a aucune influence sur la valeur de vérité d'un énoncé déontique. L'unique conséquence d'une incapacité est que le discours moral devient *inefficace* – puisqu'il n'est pas possible d'influencer le comportement des personnes visées par le discours moral – et que le *blâme* devient inapproprié.

Interprétations de DIP	Valeur de vérité de l'énoncé « S a le devoir moral de faire X » lorsque S ne peut pas faire X
Conceptuelle	FAUX
Présupposition	NI VRAI, NI FAUX
Pragmatique	VRAI L'usage du langage moral, notamment pour blâmer, est inapproprié, car inefficace.

Le problème avec l'interprétation pragmatique est le suivant : il est vrai qu'il est inutile de conseiller ou d'ordonner à quelqu'un de faire une action qu'il est incapable d'accomplir. En effet, notre discours risque bien de rester sans effet sur le comportement de la personne. Toutefois, cette interprétation est insuffisante, car elle ne rend pas compte de la manière dont nous évaluons certaines situations d'impossibilité.

Prenons le cas du propriétaire terrien qui, un matin, attache, à l'aide d'une solide corde, son esclave à un poteau au milieu d'un champ de blé. Il ordonne ensuite à l'esclave de labourer le champ pour le soir, puis l'abandonne à sa solitude. Revenant le soir et voyant que l'esclave n'a pas labouré le champ – il ne le pouvait pas, puisqu'il était attaché au poteau – le propriétaire terrien décide de punir l'esclave pour sa désobéissance en le fouettant.

Dans cet exemple, nous considérons que le blâme et la punition sont inappropriés. Toutefois, il ne semble pas que leur caractère inapproprié provienne du fait que le blâme et la punition sont inutiles, i.e. qu'ils ne permettront pas de changer le comportement de l'esclave. Étant donné que l'esclave *ne peut pas* labourer le champ, il n'est pas seulement *inutile* de le blâmer et de le punir pour son échec, mais il est *cruel* et *immoral* de le faire. Si le

propriétaire a un comportement inapproprié, ce n'est pas parce que ses actions sont *inutiles* pour faire changer le comportement de son esclave, mais c'est parce que ses actions à l'égard de l'esclave sont cruelles, perverses et inhumaines.

Faire usage du blâme et de sanctions, en cas de violation d'une norme morale impossible à satisfaire, cela n'est pas seulement inutile, c'est en plus immoral. C'est pourquoi de nombreux auteurs ont favorisé l'interprétation *morale* du principe DIP ; selon eux, ce principe exprime une *vérité morale* (BÉLANGER, 2011; OGIEIN, 2001; HINTIKKA, 1969).

4.3. L'IMMORALITÉ DES PRINCIPES MORAUX QU'IL EST IMPOSSIBLE DE RESPECTER

La dernière interprétation du principe DIP consiste à le considérer comme un principe *moral* ; il est *immoral* d'exiger d'une personne une action qu'il lui est impossible d'accomplir.²⁵⁷ Les systèmes moraux qui contiennent des exigences qu'il est impossible pour un être humain de satisfaire sont des systèmes moraux *cruels* et *inhumains*.

Selon cette approche, le principe DIP dérive ainsi d'une *méta-norme*, i.e. une norme de second ordre, qui énonce un critère pour sélectionner des normes de premier ordre. Cette méta-norme stipule qu'un système normatif ne doit pas contenir de normes cruelles et inhumaines, notamment des normes que des êtres humains ne sont pas capables d'accomplir. La méta-norme dont dérive DIP exprimerait un « principe d'humanité » selon lequel les normes utilisées pour guider le comportement des êtres humains doivent être des normes « adaptées » à l'être humain, sous peine d'être cruelles et inhumaines.²⁵⁸

À supposer qu'il serait [...] logiquement possible d'exiger l'impossible, il n'en résulterait pas que nous devrions le faire. À quoi pourrait bien ressembler un monde où l'impossible serait exigé ? À un monde bizarre (pour dire le moins), cruel et inhumain (pour dire ce qu'il serait assez exactement, puisque ces normes seraient, en principe, assorties de sanctions négatives et donc de souffrances injustifiées).

²⁵⁷ Sur le plan formel, la différence entre l'interprétation conceptuelle et morale s'exprime de la façon suivante : la première prend la forme $O(A) \rightarrow \Diamond(A)$, alors que la seconde prend la forme $O(O(A) \rightarrow \Diamond(A))$.

²⁵⁸ On le voit clairement ici : nous retrouvons une formulation de la « contrainte humaniste ». Cependant, au lieu que cette contrainte humaniste *dérive* de DIP (ce qui est le cas si DIP exprime une vérité conceptuelle), c'est ici la relation contraire qui s'applique ; DIP dérive d'un principe général d'humanité.

À supposer, donc, que la méta-norme « devoir implique pouvoir » n'ait pas de soutien logique, elle resterait justifiée, néanmoins, par ce que j'appelle un principe d'humanité. « Devoir implique pouvoir » est une méta-norme qui nous enjoint d'exclure les candidats à être des normes, qui, s'ils étaient retenus, rendraient notre monde cruel ou inhumain. Si nous jugeons que « devoir implique pouvoir » est inattaquable, c'est, à mon avis, pour cette raison (OGIEN, 2001, pp. 20-21).

Si, dans le langage ordinaire, nous respectons le principe DIP, cela n'est pas, selon Ogien, en vertu d'une relation *conceptuelle* entre « devoir » et « pouvoir », mais en vertu du fait que nous admettons un principe *moral* fondamental, à savoir le « principe d'humanité », selon lequel il est cruel d'exiger de quelqu'un de faire quelque chose d'impossible – et de le sanctionner en cas d'échec.

Si nous adoptons cet argument *moral* en faveur de DIP, alors la possibilité ou l'impossibilité de faire l'action n'affecte pas la *vérité* des énoncés normatifs ; il est ainsi possible que l'énoncé normatif « S a le devoir de faire X » soit *vrai* quand bien même il est impossible pour S de faire X. L'impossibilité pour S de faire X n'a pour conséquence ni de rendre l'énoncé « S doit faire X » faux, ni de le rendre *ni vrai ni faux*. La seule conséquence de l'impossibilité de satisfaire le devoir consiste dans le fait que l'usage du langage moral, le blâme et les sanctions deviennent *immoraux*.

Comme le montre le tableau ci-dessous, le fait qu'une personne ne peut pas faire une action a la même incidence sur la valeur de vérité d'un énoncé déontique dans l'interprétation pragmatique et dans l'interprétation morale de DIP ; dans les deux cas, l'incapacité de l'agent n'a *pas* d'incidence sur la valeur de vérité de l'énoncé déontique. Par contre, dans un cas – i.e. dans celui de l'interprétation pragmatique – l'usage du langage normatif devient inapproprié car il est *inefficace*. Dans le second cas – i.e. celui de l'interprétation *morale* – l'usage du langage normatif devient inapproprié car il est *immoral* de blâmer une personne pour avoir échoué à faire une action qu'elle ne pouvait de toute façon pas accomplir.

Le tableau ci-dessous résume les implications de chacune des interprétations proposées de DIP sur la valeur de vérité des énoncés déontiques :

Interprétations de DIP	Valeur de vérité de l'énoncé « S a le devoir moral de faire X » lorsque S ne peut pas faire X
Conceptuelle	FAUX
Présupposition	NI VRAI, NI FAUX
Pragmatique	VRAI L'usage du langage moral, notamment pour blâmer, est inapproprié, car inefficace.
Morale	VRAI L'usage du langage moral, notamment pour blâmer, est inapproprié, car immoral.

Il existe, selon nous, deux avantages majeurs qui peuvent être avancés en faveur de l'interprétation morale de DIP – par rapport aux autres interprétations : (1) contrairement à l'interprétation conceptuelle, l'interprétation morale n'implique pas de violer la loi de Hume. Contre les interprétations conceptuelle et conversationnelle (2), le fait que, dans les premières discussions (théologiques) du principe, les arguments utilisés en sa faveur étaient proprement moraux, permet de soutenir l'idée que l'acceptation ou le refus de DIP dépend de l'adoption d'une certaine position morale. En ce sens, l'interprétation morale de DIP correspond à la manière dont, historiquement, le problème de l'adéquation des exigences morales aux capacités humaines a été formulé.²⁵⁹

4.3.1. LE RESPECT DE LA LOI DE HUME

Une manière de légitimer l'interprétation *morale* du principe DIP consiste, d'abord, dans le fait que si DIP est un principe *normatif*, alors le problème de la violation de la loi de Hume ne se pose plus ; il n'est plus question d'un passage logiquement invalide entre des énoncés exclusivement descriptifs et une conclusion normative. En effet, si DIP est un principe normatif, alors le raisonnement obtient une conclusion normative à partir d'une prémisses elle-même normative, ce qui n'est pas problématique du point de vue logique :

- (P3) Les êtres humains ne peuvent pas être motivés à faire X Prémisses descriptives
- (P4) Si les êtres humains ne peuvent pas être motivés à faire X Prémisses normatives

²⁵⁹ En outre, il est évident que l'interprétation morale de DIP offre l'avantage d'expliquer le caractère *non moralement neutre* de tout système de logique déontique.

à faire X, alors ils n'ont pas le devoir moral de faire X.

(C3) ∴ les êtres humains n'ont pas le devoir moral de faire X. Conclusion normative

(P4) étant une prémisse normative, la conclusion (C3) n'est pas déduite de prémisses exclusivement descriptives. La loi de Hume est donc respectée.

Nous avons toutefois déjà fait remarquer que le recours à la loi de Hume pour rejeter DIP ne nous paraît pas adapté (section 3.3.). En effet, rejeter DIP sous prétexte qu'il ne respecte pas la loi de Hume, alors que la loi de Hume ne concerne précisément pas la sémantique du discours normatif, revient à commettre une pétition de principe. De manière réciproque, donc, le fait que l'interprétation morale de DIP permet de satisfaire la loi de Hume ne nous semble pas constituer un véritable avantage par rapport à l'interprétation conceptuelle.

4.3.2. LE DÉBAT INITIAL : PÉCHÉ ORIGINEL, PRÉDESTINATION ET LA CRUAUTÉ DE DIEU

L'interprétation *morale* de DIP offre également l'avantage de correspondre à la manière dont le principe fut utilisé initialement dans un débat théologique au sujet du rôle joué par la Grâce divine pour permettre à l'être humain d'accomplir son devoir.²⁶⁰ Dans ce débat – que nous ne pouvons que décrire de manière très générale et imprécise dans ces lignes – les uns soutenaient que les êtres humains étaient soumis à tous les commandements divins, même à ceux qu'ils ne pouvaient pas respecter. Ce fut l'avis par exemple de Martin Luther²⁶¹, comme le montre Pigden :

²⁶⁰ Dans les grandes lignes, il est possible de distinguer trois positions : les partisans de la prédestination – dont fait partie Luther – considèrent que l'homme n'est pas maître de son destin, et que, s'il fait le bien, ce n'est que parce que Dieu en a décidé ainsi. L'homme ne choisit, en ce sens, ni de faire le bien, ni de faire le mal. S'il fait le mal, c'est parce qu'il est le fils du péché originel et qu'il ne dispose pas des ressources en lui-même pour sortir du péché. S'il fait le bien, c'est parce que Dieu l'a prédestiné ainsi. A l'autre extrême, les pélagiens attribuent à l'homme un libre arbitre total : si l'homme fait le bien, c'est parce qu'il l'a choisi, et s'il fait le mal, c'est aussi parce qu'il l'a choisi. Pour finir, entre les deux, les Catholiques – dont fait partie Erasme – considèrent aussi, avec les luthériens, que l'homme ne peut pas faire le bien par lui-même ; fils du péché originel d'Adam, il a besoin de l'aide divine, de la grâce divine, pour faire le bien. Toutefois, il est libre de recevoir cette grâce ou de la refuser. Il porte donc, dans une certaine mesure, la responsabilité de son destin (QUILLIET, 2007).

²⁶¹ Le débat nous semble encore plus subtil qu'il n'y paraît : en effet, Dieu n'aurait pas créé l'homme en le rendant incapable de satisfaire son devoir moral. Dieu a donné la capacité à l'homme de faire le bien, mais l'homme a perdu lui-même cette capacité par le péché originel. En ce sens, l'homme est responsable de son

Que dire au sujet de la loi et des commandements [divins] ? Ils ne peuvent être respectés sans aide par les hommes. En effet, l'obéissance complète est impossible même pour ceux qui sont assistés par la Grâce. Il n'existe pas de rédemption par les œuvres. Tu ne peux pas, par exemple, « aimer Dieu de tout ton cœur et de toute ta force et de toutes des possibilités ». Ce n'est simplement pas possible (p. 164-165). Mais bien que tu ne puisses pas obéir, il est certain que tu le *dois*. C'est pourquoi tu es légitimement damné si tu ne fais pas partie des élus, et pourquoi l'élection est un acte de miséricorde plutôt qu'un acte de justice. En conséquence, Devoir n'implique pas Pouvoir.²⁶²

De leur côté, les opposants de Luther, notamment Erasme, furent un usage proprement *moral* de DIP : selon eux, si Dieu formule des commandements impossibles à satisfaire, et punit et récompense les êtres humains en raison de la non-conformité de leurs actions, alors il faudrait admettre non pas que Dieu commet une faute de logique, mais que Dieu est un être cruel et inhumain.

Les opposants de Luther, aussi bien les auteurs scholastiques tels que Gabriel Biel, contre lesquels Luther réagissait, et les écrivains tels qu'Erasme, qui réagissait contre Luther, étaient préoccupés par l'image de Dieu qui semblait émerger s'il était nié que « *Devoir* » implique « *Pouvoir* ». C'est dans leurs écrits que le principe Kantien a bien pu émerger explicitement pour la première fois. Le problème était simple : si on nie que « *Devoir* » implique « *Pouvoir* », est-ce que Dieu peut être juste ? Un être qui édicte des lois qu'il est impossible pour ses

incapacité à faire le bien, et Dieu n'est pas cruel dans le choix de ses commandements. Par la Grâce, Dieu accorde son pardon à l'homme pécheur – il fait un acte de miséricorde plutôt qu'un acte de justice. Il faut toutefois noter qu'une telle spécification de la morale luthérienne la rend tout à fait compatible avec le principe DIP, ce qui annule l'argument de Pigden (1990). En effet, DIP affirme qu'un individu ne peut pas être soumis à une obligation qu'il lui est impossible d'accomplir, sauf s'il s'est lui-même rendu incapable de faire son devoir – ce que Michael Stocker nomme « l'incapacité coupable » (STOCKER, 1971, p. 114). Or, cela implique que la théorie luthérienne, basée elle aussi sur le péché originel, est compatible avec la thèse selon laquelle un individu ne peut pas être soumis à un principe moral qu'il ne peut pas accomplir, sauf s'il est lui-même responsable de son incapacité à accomplir son devoir. Etant donné que l'être humain est responsable de son incapacité à respecter son devoir, à cause du péché originel, son devoir persiste même s'il est désormais incapable de l'accomplir, et Dieu ne commet pas un acte de cruauté, car il ne commande rien qui ne soit inaccessible à l'homme préalablement au péché originel.

²⁶² "What of the law and the commandments? These are impossible to unaided men. Indeed, complete obedience is impossible even to those assisted by Grace. There is no salvation by works. You cannot, for instance, 'love the Lord thy God with all thy heart and all thy strength and all thy might.' It simply isn't on (p. 164-165). But though you cannot obey, you certainly ought to do so. Which is why you are righteously damned if you are not among the elect, and why election is an act of mercy rather than strict justice. Hence Ought does not imply Can." (PIGDEN, 1990, p. 7).

sujets de respecter, et qui les punit pour leur désobéissance, serait un monstre. Le monde ressemblerait à une dictature dans les mains d'un tyran capricieux qui délibérément formule les lois de la terre de sorte que chacun est coupable de violer la loi, de sorte qu'il serait impossible d'être un citoyen obéissant. Il pourrait ainsi punir ceux qu'il choisit, et récompenser ceux qu'il choisit, en fonction seulement de sa volonté.²⁶³

Cet argument se retrouve dans les passages suivants d'un texte d'Erasme, cité par White :

A nouveau, en ce qui concerne les préceptes, si un Seigneur ordonnait constamment à un esclave attaché par les pieds à une roue de carrier (*treadmill*), « va là, fais cela, cours, reviens », tout en formulant des menaces effroyables s'il désobéit et sans le libérer, et s'il est même prêt à le fouetter s'il désobéissait, est-ce que l'esclave ne pourrait pas légitimement qualifier son maître de fou et de cruel de battre un homme à mort pour ne pas avoir fait quelque chose qu'il était incapable de faire ?

Plus encore, lorsque Dieu charge l'homme avec tant de commandements qui ne servent pas d'autres fonctions que de lui faire haïr Dieu encore plus, et d'être de ce fait encore plus terriblement damné, est-ce que cela ne rend pas Dieu encore plus mauvais que le tyran Dionysos de Sicile, qui délibérément a édicté des lois qu'il soupçonnait que la majorité ne respecterait pas [...] ?²⁶⁴

²⁶³ "Luther's opponents, both the scholastic authors like Gabriel Biel, whom he was reacting against, and writers like Erasmus, who were reacting against him, were concerned with the picture of God that seemed to emerge if you denied that 'Ought' implies 'Can'. It is in their writings that the Kantian principle may well emerge explicitly for the first time. The issue here is simple: if you deny 'Ought' implies 'Can', can God be just? A Being who enacted laws for his subjects which it was impossible for them to obey, and then punished them for disobedience, would be a monster. The world would then be like a dictatorship in the hands of a capricious tyrant who deliberately frames the laws of the land in such a way that everyone is guilty of breaking the law, so that it is impossible to be a law-abiding citizen. He could then punish whom he chose, and reward whom he chose, purely according to his whim. [...]" (WHITE, 2005, pp. 16-17)

²⁶⁴ "Again, as concerns the precepts, if a lord were constantly to order a slave who was bound by the feet in a treadmill, 'Go there, do that, run, come back', with frightful threats if he disobeyed and did not meanwhile release him, and even made ready the lash if he disobeyed, would not the slave rightly call the master either mad or cruel who beat a man to death for not doing what he was unable to do?

Furthermore, when God burdens man with so many commandments that serve for no other purpose than to make him hate God more and be more terribly damned, do not they make him worse than the tyrant Dionysius of Sicily, who deliberately made many laws that he suspected the majority would not keep in the absence of restraint, and at first took no notice, and then when he saw that everybody was breaking them, he

Selon Erasme, si Dieu punit et récompense pour des actes qui ne sont pas à la portée de l'homme, alors Dieu est un monstre cruel et inhumain, il est une brute, un tyran capricieux et sadique. Or, cette image de Dieu est incompatible avec la définition de Dieu comme un être bon et bienveillant. De ce fait, il faut admettre que Dieu respecte le « principe d'humanité » et n'exige de ses sujets que des actes que ceux-ci sont capables d'accomplir.²⁶⁵

Comme les passages précités le montrent, l'usage de DIP dans le débat théologique entre Luther et Erasme fut donc bien un usage proprement *moral*.

Il n'y a pas d'absurdité conceptuelle à être soumis à des lois que nous ne pouvons pas tenir. Mais c'est moralement monstrueux, le plus malsain des rêves des hommes malades. Dieu est dépeint comme une brute et un tyran qui prend plaisir à la misère de ses créatures et qui sauve seulement celles qu'il a prédestinées à se prosterner.²⁶⁶

Pigden a donc raison de considérer que l'historique de l'usage du principe DIP dans les débats philosophiques confirme qu'il faut favoriser l'interprétation *morale* du principe.

4.4. CRITIQUE DE L'INTERPRÉTATION MORALE

Nous avons soutenu que, parmi les interprétations alternatives à l'interprétation conceptuelle de DIP, l'interprétation *morale* du principe est la plus convaincante. Nous allons cependant voir que cette interprétation rencontre également une importante difficulté. En effet, dans cette interprétation, DIP exprime la croyance *morale* selon laquelle il est immoral d'exiger d'une personne quelque chose qu'elle est incapable d'accomplir. Le problème est que, selon l'argument du « dilemme Darwinien » développé par Sharon Street, cette croyance morale n'est pas justifiée.

began to summons them to punishment, in this way bringing everyone into his power (Erasmus, *Diatribes*, Epilogue)." (WHITE, 2005, pp. 17-18).

²⁶⁵ Bien entendu, admettre que Dieu lui-même est soumis au « principe d'humanité » soulève des problèmes concernant sa toute puissance. Nous ne rentrerons toutefois pas dans ces considérations.

²⁶⁶ "There is no conceptual absurdity in the idea that we are bound by a law we cannot keep. But it is morally monstrous, the sickest of sick men's dreams. God is depicted as a bully and a tyrant who delights in the misery of his creatures and only saves those he has predetermined to grovel." (PIGDEN, 1990, p. 22).

4.4.1. LE DILEMME MORAL DU RÉALISME MÉTAÉTHIQUE

L'interprétation proprement morale de DIP présuppose qu'il existe des vérités morales (très générales). Selon cette approche, l'énoncé « il est immoral d'exiger des êtres humains le respect de normes impossibles (ou très difficiles) à respecter » est *vrai*. Et cet énoncé est considéré comme vrai indépendamment de ce qu'en pensent les agents moraux. L'interprétation morale de DIP suppose donc qu'il existe des vérités morales objectives – i.e. elle suppose le réalisme moral.²⁶⁷

Du fait qu'elle présuppose le réalisme moral²⁶⁸, nous pouvons formuler une objection à l'encontre de l'interprétation *morale* du principe DIP – c'est-à-dire à l'encontre de l'interprétation selon laquelle DIP dérive d'un principe moral « d'humanité », suivant lequel il ne faut pas adopter des normes cruelles ou inhumaines. Cette objection se base sur une critique bien connue, formulée par Sharon Street, à l'encontre des positions métaéthiques réalistes, notamment à l'encontre du réalisme des valeurs. Selon cette critique, étant donné l'effet des croyances morales sur les chances de survie des individus, il faut admettre que l'évolution a eu une influence sur le contenu de nos croyances morales.²⁶⁹ La sélection

²⁶⁷ Ce réalisme moral peut être un réalisme des valeurs. Le principe DIP incarne ainsi la valeur d'humanité.

²⁶⁸ Bien entendu, il est possible de rendre compte de l'énoncé « il est immoral d'exiger d'un individu d'importants sacrifices » ou de l'énoncé « il est immoral d'exiger d'un individu des actions impossibles à accomplir » dans le cadre d'une théorie métaéthique non réaliste. Cependant, dans ce cas, nous rencontrons plusieurs difficultés. Par exemple, si nous adoptons une interprétation expressiviste du principe, alors il n'est plus possible de construire un raisonnement valide à l'aide de ce principe. En effet, dans l'expressivisme, la signification de l'énoncé « il est immoral d'exiger des actions impossibles à accomplir » est simplement « Bouh ! (exiger des actions impossibles à accomplir) ». Or, comme on le sait, il est difficile de construire un véritable raisonnement sur la base d'une exclamation dépourvue de valeur de vérité. En conséquence, si l'interprétation expressiviste est adoptée, il n'est plus possible d'exclure certaines normes morales sur la base d'un argument fondé sur DIP-motivation. Si nous adoptons une doctrine non objectiviste (subjectiviste, relativiste) selon laquelle « il est immoral d'exiger d'un individu trop de sacrifices » est un énoncé vrai selon l'attitude actuelle de certains individus à l'égard des sacrifices, alors l'énoncé « il est immoral d'exiger d'importants sacrifices » signifie simplement « les êtres humains désapprouvent les normes qui requièrent des sacrifices trop importants ». Dans ce cas, le raisonnement que l'on peut construire à l'aide du principe serait, par exemple, le suivant : (1) les êtres humains désapprouvent les normes qui exigent d'importants sacrifices. (2) La norme utilitariste exige d'importants sacrifices. (3) ∴ Les êtres humains désapprouvent la norme utilitariste. Aucun passage du descriptif au normatif n'a été opéré. Pour finir, plus la description des individus en question est idéale ou hypothétique, plus la doctrine se rapproche en fait d'une forme de réalisme moral ou normatif, dans quel cas notre argument s'applique à nouveau. En outre, comme le montre notamment l'article de Selim Berker (2014), une théorie constructiviste des énoncés moraux risque bien de rencontrer les mêmes difficultés qu'une théorie réaliste face aux explications évolutionnistes.

²⁶⁹ Comme le fait remarquer Sharon Street, admettre que l'évolution a exercé un rôle déterminant dans la formation des jugements moraux n'implique pas de nier que d'autres forces ont exercé une influence, notamment l'éducation, la société ou l'histoire (STREET, 2006, p. 14). Il est cependant probable, selon elle, que l'influence exercée par la société ou l'éducation explique *les variations* interculturelles de jugements moraux

naturelle a joué un rôle déterminant dans la formation de nos croyances morales.²⁷⁰ Par exemple :

Personne ne croit que les parents doivent exterminer leurs descendants [...] parce que cette croyance ferait (aurait fait) disparaître les porteurs de ce gène. Au contraire, nous croyons que les parents ont le devoir de prendre soin de leur progéniture.²⁷¹

D'autres jugements moraux, comme par exemple l'exigence de chasteté et de la fidélité, qui concerne principalement les femmes, s'explique également par des pressions évolutives ; les hommes ont tendance à prendre soin de leur progéniture uniquement s'ils sont assurés qu'il s'agit bel et bien de leur progéniture (WILLIAMS, 1995, p. 107).²⁷² Un autre exemple paradigmatique est celui de l'inceste²⁷³ ; le fait que nous jugeons l'inceste comme immoral s'explique par le fait qu'il est évolutivement avantageux d'éprouver du dégoût à l'égard de l'inceste.²⁷⁴

plutôt que les constantes. Les constantes, de leur côté, semblent pouvoir être expliquées par les forces de l'évolution (STREET, 2006, p. 119).

²⁷⁰ L'influence sur les *jugements* moraux est indirecte. Les pressions de l'évolution exercent une influence directe sur nos *tendances évaluatives fondamentales*, c'est-à-dire des expériences fondamentales d'attraction ou d'aversion. Ces tendances ont, à leur tour, exercé une influence sur la formation des jugements moraux articulés. Admettre une influence de l'évolution dans la formation des jugements moraux implique uniquement d'admettre que, si les tendances évaluatives fondamentales de l'être humain avaient été différentes, alors les jugements moraux des êtres humains auraient été différents (STREET, 2006, p. 120).

²⁷¹ "No one believes that parents should exterminate their offspring, for instance, because this belief (would have) led its bearers' genes to disappear. Instead, we believe that parents should take care of their progeny." (JAQUET, 2018, p. 1152).

²⁷² Nous trouvons dans des textes religieux l'idée que l'infidélité de la femme est plus grave que l'infidélité de l'homme. De ce fait, seul l'homme est en droit de rompre les fiançailles : « Si chacun des fiancés fornique de son côté après les fiançailles, il appartient à l'homme seul de résilier les fiançailles : car les délits sont très inégaux et un préjudice beaucoup plus grave menace l'homme à cause de l'incertitude de la descendance ». Extrait de texte tiré des *Practicae resolutiones lectissimorum casuum* de Antonio Diana, casuiste du 17^{ème} siècle, cité dans l'ouvrage de Bernard Quilliet (2007, p. 383).

²⁷³ « [L'exemple bien connu] est celui de l'aversion face à l'inceste. Les données empiriques montrent que les êtres humains sont généralement dégoûtés à l'idée même de l'inceste et ont une tendance à condamner moralement ce comportement même s'ils sont incapables de justifier rationnellement cette condamnation (Faucher, 2007 ; Haidt, 2001). Or le sentiment de dégoût face à l'inceste a une source biologique. Ce sentiment est évolutivement avantageux car il dissuade des rapports sexuels avec des individus parents et minimise ainsi les dangers liés à la consanguinité (progéniture plus susceptible d'être atteinte de maladies génétiques) » (CLAVIEN, 2012, p. 6).

²⁷⁴ Sharon Street présente d'autres exemples : « Il est clair, par exemple, qu'il serait fatal au succès reproductif de juger que le fait que quelque chose met en danger notre survie est une raison de le faire, ou que le fait que quelqu'un est de notre parenté est une raison de lui faire du tort. Une créature qui accepterait ces jugements évaluatifs sauterait elle-même d'une paroi, chercherait à rencontrer ses propres prédateurs, et elle attaquerait sa propre progéniture, ce qui résulterait sur la rapide élimination à la fois d'elle-même et de ses tendances évaluatives. Par contraste, il est évident qu'il serait bénéfique (en termes de succès reproductif) de juger que le fait que quelque chose est favorable à notre survie est une raison qui le favorise, et le fait que

Le problème est alors le suivant : la reconnaissance de l'influence des pressions de l'évolution sur nos croyances morales amène les partisans du réalisme moral à un dilemme : soit ils considèrent que nos croyances sont effectivement formées par les pressions évolutives, indépendamment des vérités morales. Cependant, dans ce cas, cela implique que, probablement, un bon nombre de nos croyances morales sont *fausses* ; ce n'est que *par accident* que nous pourrions avoir une croyance morale non seulement avantageuse, mais par ailleurs vraie.

A l'exception d'une telle coïncidence, la seule conclusion possible est que de nombreuses ou la plupart de nos jugements moraux sont sur la mauvaise voie. Il s'agit de l'important résultat sceptique qui attend n'importe quel réaliste qui affirme qu'il n'existe pas de relation entre les influences évolutives sur nos jugements évaluatifs et les vérités évaluatives indépendantes.²⁷⁵

De l'autre côté, si le réaliste admet qu'il existe une relation entre les pressions évolutives et la saisie de vérités morales, il doit soutenir qu'il était avantageux pour les êtres humains d'être capables de saisir des vérités morales. Selon cette approche, si les jugements moraux que nous exprimons sont avantageux pour la reproduction, c'est parce que, de manière générale, il est avantageux pour nous de saisir la vérité (morale).

Selon cette hypothèse, notre capacité à reconnaître les vérités évaluatives nous a conféré certains avantages et nous a aidés à nous épanouir et à nous reproduire [...] Cela s'explique du fait que les jugements évaluatifs qui ont été les plus

quelque chose favorise nos descendants est une raison de la faire. » [“It is clear, for instance, how fatal to reproductive success it would be to judge that the fact that something would endanger one's survival is a reason to do it, or that the fact that someone is kin is a reason to harm that individual. A creature who accepted such evaluative judgements would run itself off cliffs, seek out its predators, and assail its offspring, resulting in the speedy elimination of it and its evaluative tendencies from the world. In contrast, it is clear how beneficial (in terms of reproductive success) it would be to judge that the fact that something would promote one's survival is a reason in favor of it, or that the fact that something would assist one's offspring is a reason to do it.” (STREET, 2006, p. 114)]. La valeur morale moindre que nous accordons aux étrangers s'explique également par des pressions évolutives: en effet, « cette inclination naturelle à la méfiance face aux individus étrangers est porteuse d'avantages évolutifs si bien qu'il y a tout lieu de penser qu'elle ait été sélectionnée au cours de l'évolution. Il n'y a donc rien d'étonnant à ce que les êtres humains attribuent si souvent (mais pas toujours) une valeur morale moindre aux membres de groupes distants » (CLAVIEN, 2012, p. 6).

²⁷⁵ “Barring such a coincidence, the only conclusion remaining is that many or most of our evaluative judgements are off track. This is the far-fetched skeptical result that awaits any realist who takes the route of claiming that there is no relation between evolutionary influences on our evaluative judgements and independent evaluative truths.” (STREET, 2006, p. 122).

avantageux pour la sélection sont, en général, précisément les jugements évaluatifs qui sont vrais.²⁷⁶

Cependant, cette approche pose un ensemble de difficultés. La principale est qu'on ne perçoit pas quel rôle évolutif spécifique peut jouer la capacité de saisir les vérités morales. En effet, il semble que, pour jouer ce rôle évolutif, il n'y ait pas besoin d'autre chose que d'une capacité morale dont le contenu est *directement* déterminé par les pressions de l'évolution. Postuler l'existence de vérités morales indépendantes n'explique pas *mieux* le caractère avantageux de nos jugements moraux. Or, si nous pouvons expliquer nos jugements moraux en faisant l'économie d'une référence à des vérités morales objectives, alors nous devons préférer l'explication parcimonieuse (STREET, 2006, p. 129).²⁷⁷

Ensuite, il n'est pas évident que la référence à des *vérités morales* joue un rôle explicatif quelconque. En effet, il faut encore pouvoir expliquer pourquoi il est avantageux de saisir la vérité. Une solution serait d'affirmer qu'il est *toujours* avantageux de saisir la vérité – même en dehors d'un contexte moral. Toutefois, dans de nombreux cas, il semble au contraire que saisir des vérités constitue un *handicap* reproductif plutôt qu'un avantage.

Prenons, par exemple, les vérités au sujet de la présence ou de l'absence d'ondes électromagnétiques de la fréquence la plus basse. Pour la plupart des organismes, de telles vérités ne sont pas pertinentes pour la survie et la reproduction ; de ce fait, posséder cette capacité de les saisir ne leur conférerait aucun bénéfice. Et il faut aussi tenir compte des coûts liés au développement et au maintien d'une capacité aussi sophistiquée. Etant donné que pour la majorité des organismes, cela constituerait une dépense d'énergie et de ressources sans bénéfices en termes de succès reproductif, la possession d'une telle capacité serait en fait positivement désavantageuse.²⁷⁸

²⁷⁶ "According to this hypothesis, our ability to recognize evaluative truths conferred upon us certain advantages that helped us to flourish and reproduce [...]. For the evaluative judgements that it proved most selectively advantageous to make are, in general, precisely those evaluative judgements which are true." (STREET, 2006, p. 126).

²⁷⁷ Le même argument est donné par Joshua Greene : « Il est plus économe de supposer que lorsque nous ressentons la force des théories rétributives du châtimeur, nous gravitons simplement dans la direction de nos inclinations émotionnelles évoluées et non dans la direction d'une vérité morale indépendante ». "[I]t is more parsimonious to suppose that when we feel the pull of retributivist theories of punishment, we are merely gravitating toward our evolved emotional inclinations and not toward some independent moral truth" (GREENE, 2008, p. 72).]

²⁷⁸ "Take, for instance, truths about the presence or absence of electromagnetic wavelengths of the lowest frequencies. For most organisms, such truths are irrelevant to the undertakings of survival and reproduction; hence having an ability to grasp them would confer no benefit. And then one must also take into

Pour finir, les réalistes qui admettent l'influence de l'évolution dans la formation de nos jugements moraux fondamentaux ne peuvent pas expliquer pourquoi certains de ces jugements sont *faux* – après réflexion. Par exemple, les êtres humains ont tendance à penser que le fait qu'un individu n'appartient pas à son groupe est une raison de ne pas lui accorder la même importance qu'aux membres du groupe. Cette tendance s'explique par le fait qu'elle procure un avantage en termes de succès reproductif, puisque ceux qui la possèdent ont tendance à aider les individus ayant une parenté génétique forte ou qui sont les plus susceptibles d'aider à leur tour (STREET, 2006, p. 133). Pourtant, malgré l'avantage que confère cette tendance, de nombreux réalistes vont considérer qu'il est *faux* que l'appartenance ou la non-appartenance à un groupe est une raison de favoriser ou défavoriser quelqu'un. Le problème est que les réalistes ne peuvent pas expliquer cette disparité entre les attitudes évaluatives fondamentales et la vérité morale.

4.4.2. LES JUGEMENTS MORAUX QUI NE SONT PAS AVANTAGEUX POUR LA REPRODUCTION DES GÈNES

Il faut admettre qu'un grand nombre de nos croyances morales ont un contenu déterminé par l'évolution. Etant donné que ces croyances reflètent une pression évolutive plutôt qu'une vérité morale indépendante, plusieurs auteurs ont conclu que les croyances morales qui sont le résultat de l'évolution ne sont pas justifiées. En effet, puisque le processus par lequel les croyances morales sont formées n'a pas pour fonction d'obtenir une représentation vraie du monde, ces croyances sont *injustifiées* sur le plan épistémique (STREET, 2006, p. 122).

Selon certains auteurs (JAQUET, 2018; LAZARI-RADEK de & SINGER, 2012), il existe néanmoins certaines croyances morales qui ne sont pas expliquées par les pressions évolutives. Par exemple, l'intuition utilitariste, largement répandue, selon laquelle « il faut accorder au bien-être de chacun une considération égale » ne s'explique pas par la théorie de l'évolution.²⁷⁹ Au contraire, la croyance dans la vérité du principe utilitariste est néfaste pour la reproduction.

account the significant costs associated with developing and maintaining such a sophisticated ability. Since for most organisms, this would be energy and resources spent for no gain in terms of reproductive success, the possession of such an ability would actually be positively disadvantageous." (STREET, 2006, p. 130).

²⁷⁹ L'utilitariste Henry Sidgwick considérait que l'intuition utilitariste pouvait, elle aussi, être expliquée par l'évolution. Il est en effet avantageux pour l'espèce que les individus soient prêts à se sacrifier pour le bien

[En] vérité, évaluée en termes de réplcation de nos gènes, il s'agit d'une croyance qu'il est mauvais d'avoir ; ceux qui agissent selon celle-ci vont parfois sacrifier leurs intérêts les plus fondamentaux pour satisfaire les intérêts d'individus parfaitement étrangers. Selon ce point de vue [utilitariste], nous saisissons la vérité du [principe utilitariste] de la même manière que nous saisissons les vérités mathématiques ; en faisant usage de nos capacités rationnelles. De ce fait, contrairement à notre croyance que l'inceste est immoral, notre croyance dans [le principe utilitariste] est épistémiquement justifiée.²⁸⁰

Les partisans de l'utilitarisme disposent ainsi d'un excellent argument en faveur de leur théorie : l'intuition morale fondamentale sur laquelle repose l'utilitarisme ne peut pas être expliquée en termes de pression de l'évolution.²⁸¹ Il n'est donc pas possible de contester l'intuition utilitariste en montrant qu'elle est le produit de l'évolution.

du plus grand nombre. Cependant, la conception de l'évolution a changé depuis. Comme le font remarquer Katarzyna de Lazari-Radek et Peter Singer, il est admis aujourd'hui que l'évolution se produit à l'échelle soit de l'individu, soit de la communauté, mais pas à l'échelle de l'espèce : « [U]n Sidgwick contemporain serait plus proche de Street que le Sidgwick historique. Sidgwick pensait que ce qu'il appelait la « moralité du sens commun » – c'est-à-dire, l'ensemble des règles morales que nous acceptons intuitivement comme étant vraies – a tendance à produire des actions qui maximisent l'utilité. Mais s'il avait partagé notre compréhension scientifique moderne, selon laquelle les forces de l'évolution opèrent au niveau des gènes ou de l'individu, ou tout au plus au niveau de la communauté, plutôt qu'au niveau des espèces (et certainement pas au niveau des êtres sensibles), il aurait certainement été ouvert à la possibilité que les forces de l'évolution ont produit des attitudes évaluatives qui échouent à satisfaire les vérités morales ultimes telles que « Fais ce qui est le mieux pour le bien-être de tous ». "[A] contemporary Sidgwick might be closer to Street than the historical one. Sidgwick thought that what he calls "the morality of common sense"—that is, the set of moral rules that we intuitively assume to be true—tends to produce actions that maximize utility. But if he shared our modern scientific understanding that evolutionary forces operate at the level of the gene or the individual, or at most the community, rather than at the level of the species (and certainly not at the level of all sentient beings), he would surely have been open to the possibility that these evolutionary forces have produced evaluative attitudes that fail to conduce to ultimate moral truths such as "Do what is best for the well-being of all." (LAZARI-RADEK de & SINGER, 2012, p. 15)]

²⁸⁰ "[A]s a matter of fact, assessed in terms of the replication of our genes, this is a very bad belief to have; those who act on it will sometimes sacrifice their most vital interests for the interests of perfect strangers. On [this utilitarian view], we grasp the truth of [the Utilitarian principle] just as we access mathematical truths; by using our rational capacities. Hence, unlike our belief that incest is wrong, our belief in [the Utilitarian principle] is epistemically warranted." (JAQUET, 2018, p. 1152)

²⁸¹ Selon Lazari-Radek et Singer, l'évolution explique que nous ayons acquis les capacités rationnelles qui nous permettent de saisir les vérités morales. Ces capacités rationnelles sont les mêmes capacités qui nous permettent de saisir des vérités mathématiques ou physiques. Toutefois, du fait que l'acquisition de ces capacités de réflexion est avantageuse à la survie et la reproduction, cela n'implique pas que ces capacités ne permettent pas aussi de saisir des vérités morales et mathématiques indépendantes, qui ne sont pas nécessairement utiles à notre survie. La capacité de saisir des vérités morales n'a pas besoin d'être une capacité autonome, différente de notre capacité de raisonner – qui, elle, offre des avantages évolutifs. Comme le décrivent bien les auteurs : « La capacité de raisonner nous a permis de résoudre un ensemble de problèmes qui auraient autrement mis en péril notre survie, mais dès que nous sommes capables de raisonner, nous ne pouvons pas nous empêcher de reconnaître et découvrir certaines vérités qui ne favorisent pas notre survie ». ["We might have become reasoning beings because that enabled us to solve a variety of problems that would

Mais les partisans de l'utilitarisme vont encore plus loin ; ils soutiennent que l'ensemble des *objections morales* qui ont été formulées contre l'utilitarisme sont *quant à elles*, sujettes à l'explication évolutionniste – et doivent donc être rejetées.

La majorité des objections contre l'utilitarisme sont basées sur des croyances qui ont été formées par l'évolution. Ainsi, l'objection selon laquelle l'utilitarisme est trop exigeant repose sur la croyance que nous avons des devoirs particuliers envers nos proches, l'objection selon laquelle l'utilitarisme exige parfois de punir les innocents repose sur la croyance que la punition doit être rétributive, et l'objection selon laquelle l'utilitarisme ne délivre pas la bonne solution dans le problème du Trolley repose sur la croyance que l'action de pousser le gros individu [sur les rails] est immoral. En supposant que l'ensemble de ces croyances ne sont pas justifiées, du fait qu'elles sont le résultat d'un processus qui ne suit pas la vérité [*track-off process*], l'objection correspondante disparaît. L'utilitarisme peut tout à fait être très exigeant, permettre le châtiment des innocents, et il peut impliquer que nous devons pousser le gros individu sur les rails, mais ces caractéristiques ne peuvent pas être utilisées contre lui. ²⁸²

La question qui se pose pour nous consiste donc à déterminer si le principe DIP est le genre de principe qui tombe ou non sous la critique du dilemme darwinien de Street. S'il tombe sous sa critique, alors la croyance en la vérité du principe apparaît injustifiée. S'il ne tombe pas sous la critique – à l'instar du principe utilitariste – la croyance en la vérité du principe pourrait être justifiée, en montrant par exemple que certaines facultés cognitives avantageuses pour la reproduction nous ont permis, en plus, de saisir des vérités normatives nouvelles.

otherwise have hampered our survival, but once we are capable of reasoning, we may be unable to avoid recognizing and discovering some truths that do not aid our survival." (LAZARI-RADEK de & SINGER, 2012, p. 16)]

²⁸² "[M]ost objections to Utilitarianism are based on beliefs that were shaped by evolution. Thus, the objection that utilitarianism is too demanding rests on our belief that we have special duties to our relatives, the objection that utilitarianism sometimes requires punishing the innocent rests on our belief that punishment should be retributive, and the objection that utilitarianism does not issue the correct verdict in the trolley problem rests on our belief that pushing the large man is wrong. Assuming that these beliefs are unjustified because we owe them to an off-track process, the corresponding objections vanish. Utilitarianism may well be very demanding, allow for punishment of innocents, and entail that we should push large men from footbridges, but this cannot be turned against it (Greene 2008; Singer 2005: 348)." (JAQUET, 2018, p. 1152).

Nous pouvons voir quel est ainsi le problème auquel devra faire face l'interprétation *morale* du principe DIP : l'intuition morale en faveur du « principe d'humanité », selon lequel il faut des normes « adaptées » à l'être humain et à ses limitations motivationnelles – limitations qui sont elles-mêmes explicables par des pressions évolutives – semble être, elle aussi, explicable par l'évolution. Il semble en effet avantageux pour un individu de croire qu'il est immoral d'exiger de lui trop de sacrifices, ou d'exiger de lui des actions qui s'opposent à ses motivations fondamentales – par exemple, à sa tendance à favoriser ses proches, à son besoin irréprensible de se comparer à autrui, à son désir d'éprouver du plaisir et d'être heureux, etc. La croyance « il est immoral d'exiger l'impossible » ou « il est immoral d'exiger trop de sacrifices » s'explique facilement par son caractère avantageux en termes de reproduction des gènes.²⁸³ Elle n'est pas le résultat d'un processus de formation de croyances dont la fonction serait de « suivre la vérité » (*track the truth*) (STREET, 2006, p. 126). Cette croyance dans la vérité du « principe d'humanité », dont DIP serait dérivé, doit donc être rejetée comme injustifiée.²⁸⁴

En conclusion, considérer que DIP exprime une vérité morale a pour conséquence de soumettre ce principe au « dilemme du réalisme moral » formulé par Sharon Street. Il ne semble donc pas approprié d'opter pour cette interprétation. Cela signifie qu'aucune des trois interprétations alternatives à l'interprétation conceptuelle – i.e. l'interprétation par présupposition, pragmatique et morale – n'est satisfaisante. Le seul choix qui reste semble être le suivant : soit nous renonçons complètement à DIP, soit nous défendons l'interprétation conceptuelle de DIP contre ses critiques – ce que nous ferons dans le prochain chapitre. Dans le premier cas, l'attrait que ce principe exerce dans notre pensée et la manière dont nous faisons usage du langage normatif restent un mystère. Et il faut également renoncer à la logique déontique telle qu'elle existe aujourd'hui. Pour éviter ces conséquences fâcheuses, nous préférons défendre l'interprétation conceptuelle de DIP contre ses critiques – ce que nous ferons dans le prochain chapitre. Si nous y parvenons,

²⁸³ Il est intéressant qu'un auteur comme Joshua Greene, qui soutient que nos croyances morales s'expliquent par l'évolution, soutient simultanément qu'il serait « inhumain » d'attendre des individus qu'ils s'opposent à leurs tendances naturelles. « Il me semble que celui qui n'est pas prêt d'agir d'après les tendances humaines qui ont des causes évolutionnaires amORALES est quelqu'un qui n'est pas prêt à être humain. » [“It seems that one who is unwilling to act on human tendencies that have amoral evolutionary causes is ultimately unwilling to be human.” (GREENE, 2008, p. 76).]

²⁸⁴ Lazari-Radek et Singer offrent un argument similaire à l'encontre du principe de l'égoïsme rationnel, selon lequel nous avons une raison de faire toute action qui sert notre intérêt (LAZARI-RADEK de & SINGER, 2012, p. 28).

nous serons capables d'expliquer facilement l'attrait intuitif du principe ainsi que les usages que nous faisons du langage normatif. Toutefois, même si nous parvenons à défendre l'interprétation conceptuelle, cela ne garantit pas encore que DIP permette de défendre l'internalisme moral.²⁸⁵ En effet, comme nous le verrons dans la prochaine section, le principe DIP-motivation, qui sert de prémisse à la défense de l'internalisme moral, rencontre un problème de « vacuité », lorsqu'il est employé pour exclure des normes morales trop exigeantes sur le plan motivationnel.

5. LA VACUITÉ DU PRINCIPE « DEVOIR-IMPLIQUE-POUVOIR »

*Les questions intéressantes au sujet du « naturalisme en éthique » s'étendent au-delà de considérations purement logiques. Le naturalisme dans un sens plus large consiste dans la tentative de décrire certains aspects fondamentaux de ce qu'est une bonne vie humaine, sur la base de considérations au sujet de la nature humaine. Si ce projet doit échouer, ce ne sera pas pour des raisons purement logiques ; ce sera pour la raison plus intéressante qu'il n'existe pas de vérités appropriées au sujet de la nature humaine.*²⁸⁶

Jusqu'à présent, nous avons vu des objections contre le principe général DIP. Nous allons maintenant voir une objection à l'encontre, plus spécifiquement, du principe DIP-motivation. Selon cette objection, *même si nous acceptons* l'interprétation conceptuelle de DIP-motivation, il reste un problème : le principe DIP-motivation est inutile. En effet, le principe DIP-motivation souffre de *vacuité*, c'est-à-dire qu'il ne permet pas d'effectuer une partition entre les doctrines morales acceptables et inacceptables. Il ne le permet pas car il n'existe pas, en fait, de limites aux motivations humaines.

²⁸⁵ Pour rappel, l'internalisme moral consiste dans la thèse suivante: si S a le devoir moral de faire X, alors il existe, dans l'ensemble motivationnel de S, un désir que l'action X permet de satisfaire. Cette thèse est déduite du principe DIP-motivation, combiné à la conception humienne de la motivation, selon laquelle toute motivation requiert la présence d'un désir autonome. Rappelons aussi que l'internalisme moral peut également être défendu à l'aide d'un argument plus traditionnel, fondé sur l'internalisme des raisons d'agir, mais ce n'est pas la voie que nous avons adoptée.

²⁸⁶ "Interesting questions about 'naturalism' in ethics in fact go beyond these purely logical issues. Naturalism in a broader sense consists in the attempt to lay down certain fundamental aspects of a good human life on the basis of considerations of human nature. If this project fails, it is not for purely logical reasons; it will rather be for the more interesting reason that the right sort of truths do not exist about human nature." (WILLIAMS, 1995, p. 101).

L'objection s'attaque au principe DIP-motivation tel qu'il est utilisé par les philosophes que nous qualifierons d' « empiristes ». Selon ces philosophes, il est possible de dériver des énoncés normatifs sur la base de prémisses descriptives au sujet des motivations humaines. Ainsi, les empiristes font un usage *contraposé* du principe DIP-motivation ; du fait qu'un être humain ne peut pas être motivé à faire X, alors il n'a pas le devoir moral de faire X. Selon cette approche empiriste, l'antécédent de l'énoncé conditionnel, à savoir « un être humain ne peut pas être motivé à faire X » est un énoncé descriptif dont la vérité doit être établie par les sciences empiriques – notamment par la psychologie et la biologie. Sur la base d'un tel énoncé descriptif, et par l'intermédiaire de DIP-motivation, les empiristes obtiennent une conclusion normative, à savoir « il n'est pas le cas qu'un être humain a le devoir moral de faire X » :

(P1)	Un être humain ne peut pas être motivé à faire X	Prémisse descriptive
(P2)	Si un être humain ne peut pas être motivé à faire X, alors il n'a pas le devoir moral de faire X	Prémisse conceptuelle (DIP-motivation)
(C)	∴ un être humain n'a pas le devoir moral de faire X.	Conclusion normative

Cet usage doit être opposé à l'usage « moraliste » de DIP-motivation. En effet, les moralistes, de leur côté, font un usage non contraposé de DIP ; selon eux, le fait que les êtres humains ont le devoir moral de faire X permet de conclure que les êtres humains *peuvent* être motivés à faire X. Autrement dit, alors que les empiristes obtiennent une conclusion normative sur la base d'une prémisse descriptive empirique, les moralistes obtiennent une conclusion descriptive (non empiriste) sur la base d'une prémisse normative. Nous y reviendrons plus en détail par la suite.

Dans la présente section, nous allons clarifier l'objection de *vacuité* contre l'usage empiriste de DIP-motivation. Pour ce faire, nous reprendrons un argument de David Velleman. L'idée générale de l'argument est la suivante : certes, le principe DIP-motivation permet d'établir qu'il existe une liaison nécessaire entre certains énoncés normatifs et certains énoncés descriptifs au sujet des motivations humaines. Le problème provient du fait que ce principe *est vide*, car il n'existe pas de vérité au sujet de ce que l'être humain est *incapable d'être motivé à faire* – ou du moins, si une telle vérité existe, elle n'est pas

connaissable par des êtres humains. Autrement dit, même si nous acceptons que, s'il existait des choses que les êtres humains étaient incapables d'être motivés à faire, alors ils n'auraient pas le devoir moral de les faire, il s'avère qu'il n'existe *rien*, en fait, que les êtres humains sont incapables d'être motivés à faire. En d'autres termes, même si nous acceptons DIP-motivation, ce principe est inutile car toute prémisse descriptive au sujet de ce qu'un être humain ne peut pas être motivé à faire, est toujours fausse.

Dans les prochaines sections, nous allons élaborer cette objection de *vacuité* contre le principe DIP-motivation, tel qu'il est employé par les empiristes. Pour y parvenir, nous présenterons le problème de la vacuité tel qu'il a été formulé par David Velleman. L'auteur montre que le problème de la vacuité de DIP se pose pour *n'importe quelle* notion normative, même pour des notions normatives qui semblent, de prime abord, très étroitement liées aux motivations – contrairement aux notions morales. C'est le cas notamment de la notion de « bien pour » qui, à première vue, paraît étroitement liée aux motivations des personnes.

Dès lors que nous aurons, avec Velleman, montré que le principe DIP est *vide* lorsqu'il est utilisé pour déterminer ce qui est « bon pour » une personne humaine, nous étendrons l'argument pour l'appliquer aux notions morales, notamment à la notion d'obligation morale. Autrement dit, notre argument consistera en deux étapes : d'abord, montrer que DIP souffre de vacuité lorsqu'il est utilisé pour déterminer ce qui est « bon pour » une personne. Ensuite, soutenir que, si le problème se pose pour la notion de « bien pour », alors il se pose *d'autant plus* pour les notions morales – dont le lien avec les motivations est plus difficile à établir.

Ces deux étapes de l'argument occuperont ce cinquième chapitre. Dans le chapitre suivant, nous présenterons et répondrons à de possibles critiques que les empiristes pourraient formuler contre l'objection de la vacuité. Nous concluons, contre les empiristes, qu'il n'est pas possible de rejeter des normes morales à l'aide du principe DIP-motivation – et cela, même si DIP-motivation exprime une vérité conceptuelle – car il n'existe rien que les êtres humains sont incapables de désirer.

5.1. UNE RELATION ÉTROITE ENTRE LE BIEN NON MORAL ET LA MOTIVATION

Les empiristes font un usage de DIP-motivation dans le but de relier des vérités au sujet des désirs ou motivations des individus à des vérités morales. Cependant, la relation entre, d'un côté, des notions comme « moralement bon » ou « immoral », « juste » ou « injuste », et, de l'autre côté, les vérités au sujet de ce que les êtres humains sont motivés ou peuvent être motivés à faire, n'est pas évidente. Au contraire, il semble, de prime abord, que les notions morales *ne sont pas* reliées aux motivations des individus; les exigences morales sont considérées comme étant précisément des exigences qui valent *indépendamment* des désirs des individus. Ce qui est exigé moralement d'une personne est indépendant de ce que cette personne est motivée ou peut être motivée à faire. En conséquence, nous considérons en général qu'il existe un écart important entre les notions morales et les motivations des personnes.

Les normes de la morale et de la rationalité ne sont pas faites pour correspondre aux tempéraments individuels ; de ce fait, nous pouvons imaginer un individu qui exprime une indifférence totale aux normes morales, et peut-être même aux normes de rationalité, en disant qu'elles ne sont pas faites pour lui.²⁸⁷

Nous comprenons donc qu'il est particulièrement difficile, pour les théoriciens qui considèrent que les normes morales s'appliquent indépendamment des désirs des individus, d'expliquer pourquoi un individu se préoccuperait de respecter ces exigences. C'est pourquoi la question de la « motivation morale » constitue un véritable problème ; comment expliquer que des personnes se préoccupent ou soient motivées à respecter des exigences qui ne les concernent pas directement ? Des exigences qui sont valables indépendamment de leurs motivations individuelles ?

Il existe en revanche un autre concept normatif entretenant un lien beaucoup plus étroit avec les préoccupations et motivations individuelles. Il s'agit du concept normatif *non moral* de « bien pour (quelqu'un) ». Comme le fait remarquer Velleman :

Mais le bien propre à un individu est, de son côté effectivement adapté à cet individu, et on ne peut imaginer que cet individu affirme que son propre bien

²⁸⁷ "The norms of morality and rationality aren't tailored to suit individual tempers; and so we can imagine an agent expressing indifference to morality, and perhaps even to rationality, by saying they aren't for him." (VELLEMAN, 1998, p. 89)

n'est pas fait pour lui, puisque la caractéristique d'être fait pour lui semble essentielle au fait d'être spécifiquement son bien à lui.²⁸⁸

Contrairement aux notions morales (bon/juste/vertueux), qui ne semblent pas entretenir de relation directe avec ce dont les individus se préoccupent, la notion de « être bon pour quelqu'un » semble de son côté entretenir une relation directe avec ce dont les personnes se préoccupent. Il est beaucoup plus difficile de concevoir quelqu'un qui affirmerait « je ne me soucie pas de mon propre bien, car mon propre bien n'est pas fait pour moi ». A première vue, en conséquence, l'internalisme²⁸⁹ au sujet de la notion de « bon pour quelqu'un » possède une forte légitimité intuitive. En effet, il semble intuitivement vrai que :

Internalisme du bien personnel	Quelque chose est bon pour une personne S seulement si l'objet présente une certaine relation R à la motivation ou à l'état affectif de S.

Cette thèse internaliste au sujet du bien non moral des personnes consiste simplement à affirmer que le bien de toute personne entretient une relation avec les désirs de cette personne ; il est difficile d'envisager une personne indifférente à son propre bien.

Toutefois, comme nous le verrons, cette première impression est trompeuse : il n'existe en fait pas de lien *nécessaire* ou *conceptuel* entre le bien d'une personne et sa motivation. Ce lien ne peut être établi que par le biais du principe « devoir-implique-pouvoir ». Nous verrons ainsi que l'internalisme du « bien pour » n'est vrai que si DIP est vrai. Dit autrement, il est vrai que la notion de « être bon pour S » entretient une relation conceptuelle R avec la motivation de S uniquement si le principe DIP est vrai. L'internalisme du bien personnel dépend donc, comme l'internalisme moral, de l'acceptation de DIP.

²⁸⁸ "But a person's own good is indeed tailored to him, and we cannot imagine him saying that his good is not for him, since its being for him seem essential to its being specifically his good." (VELLEMAN, 1998, p. 89)

²⁸⁹ Une théorie est dite "internaliste", en métaéthique, lorsque les motivations de l'individu figurent parmi les conditions de vérité d'un énoncé au sujet de cet individu. Nous avons déjà mentionné l'internalisme moral, selon lequel un individu S a le devoir moral de faire X seulement s'il existe dans l'ensemble motivationnel de S un désir que X permet de satisfaire. Nous avons déjà abordé, également, l'internalisme du jugement moral, selon lequel l'énoncé « S juge que X est moralement bon » est vrai uniquement si S est motivé à faire X. Pour finir, nous avons vu également l'internalisme *des raisons d'agir*, selon lequel l'énoncée « S a une raison de faire X » est vrai uniquement si S est motivé à faire X. L'internalisme qui nous préoccupera dans cette section est l'internalisme du « bien personnel » ou internalisme de la valeur. Selon ce nouvel internalisme, l'énoncé « X est bon pour S » est vrai seulement si S est motivé à obtenir X ou accorde de l'importance au fait d'avoir X.

5.2. DEUX FORMES D'INTERNALISME AU SUJET DU BIEN PERSONNEL

Avant d'argumenter en faveur de la thèse selon laquelle il n'existe pas de relation directe entre le bien personnel et la motivation, mais seulement une relation *indirecte*, par le biais de DIP, il nous faut commencer par distinguer deux notions différentes du bien personnel.

5.2.1. DEUX CONCEPTIONS DU « BIEN POUR »

Velleman distingue deux conceptions du bien personnel et deux formes d'internalisme correspondantes. Il est possible ainsi de distinguer (1) le bien *pratique* qui concerne la motivation et (2) le bien-*être* qui concerne l'état affectif. La différence se marque principalement par la perspective adoptée par la personne lors de la délibération, selon qu'elle adopte une attitude active ou passive.

Dans le premier cas, établir quel est le bien (pratique) d'une personne consiste pour elle à déterminer ce qu'elle doit poursuivre dans son propre intérêt. Réfléchir à son bien consiste pour elle à déterminer l'objet approprié de son intérêt personnel, c'est-à-dire à déterminer ce qui mérite d'être l'objectif de sa propre action en vue de réaliser son intérêt personnel. Le bien de la personne sous cet angle se définit comme ce qui est digne d'être poursuivi (ou d'être voulu ou désiré) par cette personne *pour son propre intérêt*. Le bien est ainsi directement lié à l'action de la personne, puisqu'il guide les motivations intéressées de la personne vers leurs objets appropriés (VELLEMAN, 1998, p. 98). Ce bien est donc « pratique » en tant qu'il concerne l'individu en tant qu'acteur intéressé qui cherche à attribuer un objectif à son action.

Une telle définition du bien pour une personne exclut certaines choses auxquelles nous accordons pourtant de la valeur. Ainsi, les choses qui sont dites bonnes uniquement dans la mesure où elles ne *sont pas* poursuivies par la personne ne peuvent pas figurer parmi les biens pratiques de cette personne, puisque ces derniers biens se définissent précisément par le fait d'être dignes d'être poursuivis. Par exemple, la valeur d'une visite imprévue d'un ami réside pour beaucoup dans le fait justement d'être imprévue. Le fait de rechercher ou d'agir en vue d'obtenir la visite d'un ami ôte de la valeur à cette visite. Si je paie un ami pour qu'il me rende une visite de temps à autre sans prévenir, j'aurai obtenu ma

visite imprévue, mais sa valeur sera moindre du fait que je l'ai recherchée. La visite imprévue d'un ami n'est donc pas un bien pratique pour une personne, dans le sens où elle n'est pas digne d'être poursuivie par la personne dans son propre intérêt. Au contraire, en étant l'objectif d'une action, elle ne serait pas un bien pour la personne. Pourtant, nous voulons quand même pouvoir dire que la visite imprévue d'un ami *est un bien* pour la personne qui reçoit la visite. Or, s'il ne s'agit pas d'un bien dans ce premier sens, c'est-à-dire un bien dans le sens pratique, cela doit être un bien *dans un autre sens*. Nous parvenons ainsi au second sens de la notion de « bien pour une personne ».

Dans ce second sens, la personne ne se présente plus comme un acteur intéressé qui cherche à attribuer un objectif à son action, mais comme un *patient* que les événements ou les objets affectent positivement ou négativement. Le bien de la personne ne concerne plus l'intérêt personnel et donc la détermination de l'objet approprié de l'action, mais le bien-être et donc la détermination de l'effet positif ou négatif des événements et des objets sur l'état affectif, la joie ou le plaisir, de la personne. La personne est donc, dans ce cadre, un *patient* qui subit les événements et voit son bien-être affecté positivement ou négativement selon les événements en question. Ceux-ci affectent positivement ou négativement l'état de la personne, et son bien-être en est proportionnellement et passivement rehaussé ou réduit.

Dans ce second sens, la visite imprévue d'un ami est un bien pour la personne puisqu'elle affecte positivement son bien-être. Cette visite imprévue remplit la personne d'une joie qu'elle n'aurait pas ressentie si elle avait cherché elle-même à susciter cette visite. Ce second type de bien, le bien-être, englobe donc (entre autre) les choses dont la valeur dépend justement du fait de ne pas être poursuivies. Il ne s'agit donc plus ici de déterminer ce qui est digne d'être poursuivi par une personne, ou d'être l'objet de son vouloir et de son désir, mais ce qui est digne d'avoir de la valeur ou de l'importance pour la personne.

Ainsi, nous pouvons distinguer, d'un côté, le bien « pratique » de la personne, qui constitue le bien de la personne en tant qu'acteur, bien qui est déterminé par l'intérêt personnel, et, de l'autre, le bien « passif » ou « affectif » de la personne, son « bien-être », qui constitue le bien de la personne en tant que patient et qui est déterminé par la manière dont l'état émotionnel de la personne est affecté.

5.2.2. DEUX VARIÉTÉS D'INTERNALISME DU « BIEN POUR »

De ces deux conceptions du « bien pour », il est possible de dégager deux conceptions correspondantes de l'internalisme. Pour rappel, ce dernier est une doctrine qui affirme l'existence d'un lien nécessaire entre le « bien pour la personne » et sa motivation ou son état affectif. Quelque chose est un bien *pratique* pour une personne s'il est digne d'être l'objet de son vouloir ou l'objectif de son action, c'est-à-dire s'il est l'objet approprié de sa motivation. En outre, si quelque chose est un bien *affectif* pour une personne, dans le sens où il participe à son bien-être, cela signifie qu'il est un objet digne de se voir accorder de la valeur ou de l'importance par la personne. Alors que le bien pratique est l'objet approprié des motivations intéressées de la personne en tant qu'acteur, le bien affectif ou bien-être est l'objet approprié de ses *affections* intéressées, et en ce sens constitue un objet digne de valeur pour la personne.

Internalisme du bien « pratique » de S	Une entité X est « un bien pratique » pour S si et seulement si X est l'objet approprié de la motivation intéressée de S – étant donné que X favorise l'intérêt personnel de S.
Internalisme du bien « affectif » de S	Une entité X est « un bien passif » pour S si et seulement si il est approprié pour S de se soucier ou d'accorder de la valeur à X – étant donné que X présente un effet positif sur l'état affectif de S.

Du fait qu'être un bien pour la personne se définit comme « être l'objet approprié de sa motivation » ou « être l'objet approprié de son souci », cela signifie que la personne, si elle est rationnelle, *doit* poursuivre cet objet ou lui accorder de l'importance. Bien sûr, il est possible que d'autres éléments viennent peser dans la balance et qu'ainsi, toutes choses considérées, la personne ne doive plus rationnellement poursuivre X ou lui accorder de l'importance. Mais en l'absence de telles considérations supplémentaires, le fait d'être l'objet approprié de la motivation ou du souci de S – autrement dit, le fait d'être un objet digne de poursuite ou de valeur pour S – crée une exigence rationnelle pour S de poursuivre ou d'accorder de la valeur au dit objet. Ainsi, si quelque chose est un bien pour une personne, alors cette personne *doit* rationnellement – ou « a une raison de » – poursuivre ou accorder de la valeur à cette chose, toutes choses égales par ailleurs.

Nous avons donc vu comment peut s'établir une connexion nécessaire entre le bien d'une personne et la motivation ou l'état affectif de cette personne. Cependant, ce lien ne constitue pas un lien très fort. En effet, tout ce qui a été établi pour l'instant, c'est uniquement que le bien personnel est un concept *normatif* qui permet d'évaluer la motivation et l'état affectif d'une personne (VELLEMAN, 1998, p. 100). Quelque chose est un bien pour une personne s'il doit rationnellement être l'objet de sa motivation ou de son affection. Le rapport entre le bien de la personne et sa motivation ou son état affectif est ici un *rapport normatif* plutôt que descriptif. « X est un bien pour S » signifie que S a l'obligation rationnelle de poursuivre, de vouloir ou de désirer X – dans le cas de la motivation – ou de lui accorder de l'importance et de s'en réjouir – dans le cas de l'affection. Il ne signifie *pas* que S *désire* actuellement X ou qu'il *accorde* de l'importance à X.

Un internalisme plus fort ne se contenterait pas d'établir un lien *normatif* entre le bien pour la personne et sa motivation ou son état affectif, mais chercherait à établir un lien *descriptif*. Un internalisme plus fort soutiendrait que « X est un bien pour S » signifie que « X est l'objet du vouloir ou du désir de S » ou que « X a de l'importance pour S ». Il ferait ainsi partie de la signification du concept de « bien pour une personne » de ne s'appliquer qu'à des objets qui, *dans les faits*, sont l'objet du vouloir ou du désir de S. En d'autres termes, quelque chose ne serait pas un bien pour S, si dans les faits, S ne le désirait pas ou ne lui accordait pas de valeur.

Selon un tel internalisme *fort*, « être un bien pour S » signifie « être désiré/valorisé par S ». ²⁹⁰ Toutefois, cette interprétation exclut qu'il existe des choses qui *sont désirées ou valorisées* par S mais qui, néanmoins, ne sont pas *bonnes* pour S et, réciproquement, qu'il existe des choses qui sont *bonnes* pour S mais qui ne sont pas désirées ou valorisées par S. Or, il existe de nombreux exemples pour ces deux situations ; par exemple, une personne peut désirer fumer une cigarette, alors que fumer une cigarette n'est pas bon pour elle. De

²⁹⁰ Il est important pour notre propos que le lien entre la notion de « bien pour » et la motivation soit un lien *conceptuel*. C'est pour cela que nous considérons la possibilité d'une *synonymie* entre les deux notions. Bien entendu, il est possible de soutenir, comme l'a fait Spinoza, que les deux notions ne sont pas *synonymes*, mais que, néanmoins, le fait que quelque chose est *désiré* par une personne explique que cette chose soit bonne pour elle : « Il ressort donc de tout cela que, quand nous nous efforçons à une chose, quand nous la voulons, ou aspirons à elle, ou la désirons, ce n'est jamais parce que nous jugeons qu'elle est bonne ; mais au contraire, si nous jugeons qu'une chose est bonne, c'est précisément parce que nous nous y efforçons, nous la voulons, ou aspirons à elle, ou la désirons. » (SPINOZA, 1988 (1677), p. 221). Cependant, même si nous adhérons à la théorie de Spinoza, plutôt qu'à la thèse de la *synonymie*, l'objection exposée dans ce paragraphe s'appliquerait encore.

la même manière, une personne peut *ne pas désirer* faire du sport, alors qu'il serait bon pour elle de faire du sport. Dans les deux cas, il existe un écart entre les désirs et états affectifs *actuels* de la personne et ce qui est bon pour elle. En conséquence, « être bon pour S » ne signifie pas « être désiré/valorisé par S ».

Nous pouvons modifier la définition du bien personnel pour éviter de considérer que les désirs et états affectifs *actuels* de la personne définissent son bien. Pour y parvenir, il est nécessaire de spécifier les circonstances (idéales) particulières dans lesquelles l'internalisme est valable, c'est-à-dire dans lesquelles l'individu désire ce qui est effectivement bon pour lui. Par exemple, en désignant le bien pour une personne comme étant ce que la personne désirerait en situation d'information vraie et complète et en l'absence d'erreur de raisonnement – ce qui n'est pas forcément le cas actuellement – alors il n'existerait plus de cas où les désirs de la personne ne correspondent pas à ce qui est bon pour elle. En effet, en supposant que les cas de conflit entre les désirs effectifs et le bien de la personne soient tous générés par un manque d'information, une croyance fausse ou une erreur de raisonnement, tout ce qu'une personne désire en situation d'information vraie et complète et en l'absence d'erreur de raisonnement, est bon pour elle. Le bien d'une personne se définit par ce que la personne désirerait si elle était pleinement rationnelle.²⁹¹

Toutefois, une telle conception du bien personnel nous ramène à notre définition initiale, dans laquelle le bien d'une personne est défini comme ce qui est l'objet approprié de son intérêt personnel, c'est-à-dire l'objectif qu'il serait rationnel pour la personne de poursuivre, c'est-à-dire ce que cette personne désirerait si elle était pleinement rationnelle. Définir des conditions de rationalité du désir revient au même que de déterminer un objet rationnel de désir pour la personne.

Ce que poursuivrait une personne en vue de son propre intérêt dans des conditions d'information complète peut bien être tout ce que cette personne devrait poursuivre en vue de son propre intérêt.²⁹²

Ainsi, nous revenons à la conception du bien pour une personne comme signifiant ce que la personne doit rationnellement désirer et poursuivre ou ce à quoi elle doit

²⁹¹ Il s'agit de la position défendue par Peter Railton (RAILTON, 2016).

²⁹² "What a person would pursue in a spirit of self-interest under conditions of full information may well be all that he rationally ought to pursue in that spirit [...]" (VELLEMAN, 1998, p. 95).

rationnellement accorder de l'importance. Nous revenons de ce fait à une connexion entre le bien d'une personne et son état motivationnel ou affectif qui est *normative*. Le lien nécessaire entre la notion de « bien pour une personne » et la notion de ce que « la personne doit rationnellement poursuivre/désirer/valoriser » est établi, mais sans qu'aucune liaison ne puisse être directement établie entre le bien de la personne et son état (motivationnel ou affectif) actuel – comme chercherait justement à l'établir un internalisme plus fort.

5.3. LE BIEN PERSONNEL COMME OBJET POSSIBLE DE SA MOTIVATION OU DE SON SOUCI

Nous avons vu jusqu'à présent qu'il existe bien une relation *nécessaire* ou *conceptuelle* entre le bien d'une personne et sa motivation ou son état affectif, mais nous avons vu que cette relation est purement *normative*. Elle ne concerne pas les motivations actuelles des personnes. Nous allons voir, maintenant, qu'il est nécessaire de faire appel à DIP pour établir une relation entre la notion normative de « bien pour » et des énoncés *descriptifs* au sujet des motivations des personnes.

5.3.1. LE BIEN PERSONNEL MIS EN RELATION AVEC LA MOTIVATION PAR LE BIAIS DE DIP

Nous disposons jusqu'à présent d'un internalisme très faible au sujet de la notion de « bien pour », en ce sens qu'il n'établit aucun lien entre le bien de la personne et la motivation ou l'état affectif *actuel* de cette personne. Le lien concerne uniquement ce que la personne *doit* ou *devrait* désirer/valoriser. Il est cependant possible d'établir une autre relation conceptuelle entre le bien de la personne et ses désirs et états affectifs. Dans cette nouvelle conception, un objet est un bien pour une personne si la personne *a la capacité* de le prendre comme objet de sa motivation ou de son souci. En d'autres termes, si la personne n'a *pas* la capacité de faire de l'objet un objet de sa motivation ou si elle n'a pas la capacité de lui accorder de la valeur, alors cet objet ne peut pas être un bien pour cette personne. Nous trouvons donc un lien entre la notion de « bien pour la personne » et une description de « ce que la personne a la capacité de poursuivre » ou « ce à quoi la personne a la capacité d'accorder de l'importance ».

Cependant, ce lien ne peut être établi que par l'intermédiaire du principe DIP. En effet, le seul lien *direct* qui existe entre la notion de « bien pour S » et la motivation/état affectif

de S est celle d'un lien normatif ; « X est un bien pour S » implique que « il est approprié pour S de poursuivre X ou de lui accorder de la valeur », ce qui implique que S a l'obligation rationnelle – toutes choses égales par ailleurs – de désirer ou de valoriser X. Pour mettre en relation des notions normatives et descriptives, il faut ajouter que la notion d' « obligation » implique celle de « possibilité ». Il est donc nécessaire d'accepter le principe DIP pour établir une relation *conceptuelle* entre la notion normative de « bien pour S » et la notion descriptive de « ce que S est capable de désirer/valoriser ». Ce n'est qu'en admettant le principe DIP que la notion de « bon pour » implique nécessairement celle de « pouvoir désirer/valoriser ».

Naturellement, ce qui *doit* être l'objet d'une attitude doit être quelque chose qui *peut* être l'objet de l'attitude ; il ne peut pas être le cas qu'un individu *doit* avoir une attitude à l'égard d'un objet s'il est incapable d'avoir cette attitude. Donc le bien pratique d'un individu doit consister dans des choses qu'il est capable de poursuivre pour elles-mêmes ; et son bien-être doit consister en des choses auxquelles il est capable d'accorder de la valeur pour elles-mêmes.²⁹³

Un individu ne peut donc avoir l'obligation rationnelle d'avoir une certaine attitude (motivation ou état affectif) envers un objet que s'il est capable d'adopter ladite attitude envers l'objet en question. Cette conclusion dérive de l'application du principe « devoir-implique-pouvoir » à la définition d'un « bien pour la personne », que ce bien soit considéré comme bien pratique ou comme bien affectif. Puisque « bien pour une personne » est une notion normative exigeant une certaine attitude de la part de la personne – en tant qu'elle est rationnelle – alors la notion de « bien pour une personne » implique que la personne possède la capacité d'adopter l'attitude en question. Si la personne est incapable d'adopter l'attitude requise, alors elle ne peut pas avoir l'obligation d'adopter cette attitude à l'égard de l'objet – et cet objet ne peut pas être un bien pour elle.

Le lien conceptuel entre les notions normatives non morales de « bien pour » et la motivation s'établit donc selon le raisonnement suivant :

²⁹³ “Naturally, what ought to be the object of an attitude must be something that can be the object of that attitude; it cannot be the case that a person ought to have some attitude toward an object if he's constitutionally incapable of having it. Hence a person's practical good must consist in things that he is capable of being moved to pursue for his own sake; and his well-being must consist in things that he is capable of caring about at all for his own sake.” (VELLEMAN, 1998, p. 97)

- (1) X est bon pour S si et seulement si S doit rationnellement désirer/valoriser X – en l’absence de raisons plus fortes de ne pas désirer/valoriser X.
- (2) S doit rationnellement désirer/valoriser X seulement si S peut désirer/valoriser X.
- (3) $\therefore$ X est bon pour S seulement si S peut désirer/valoriser X.

Comment déterminer si une personne est capable ou incapable d’adopter une certaine attitude ? Comme nous l’avons mentionné préalablement, la capacité pertinente est déterminée par les capacités *de l’espèce* à laquelle cette personne appartient – plutôt que par ses capacités individuelles. En ce sens, une personne est considérée comme essentiellement capable ou incapable d’adopter une certaine attitude vis-à-vis d’un objet s’il fait partie de la nature des créatures de l’ensemble auquel elle appartient d’avoir la capacité ou l’incapacité d’adopter cette attitude – même si aucune des créatures de l’ensemble ne fait usage *dans les faits* de cette capacité.

A l’aide de cette conception des capacités motivationnelles d’un individu humain, et sur la base de DIP, nous obtenons la thèse suivante au sujet du bien d’une personne : si une personne est *essentiellement* incapable de vouloir un objet, c’est-à-dire s’il fait partie de la nature des créatures de l’ensemble auquel elle appartient d’être incapables de vouloir cet objet, alors cet objet ne peut pas être un bien pour cette personne.

Tout ce que [l’internalisme au sujet du « bien pour »] implique, alors, c’est que quelque chose ne peut pas être bon pour quelqu’un si cet individu est incapable de s’en soucier – si s’en soucier dépasse les capacités affectives de créatures de son espèce. Et cette conclusion permet qu’une chose puisse être bonne pour lui aussi longtemps que la possibilité de s’en soucier n’est pas exclue par sa nature même.²⁹⁴

Cette conclusion internaliste implique que quelque chose ne peut pas être bon pour une personne si la personne ne peut pas désirer ou accorder de la valeur à cette chose. Ce qui est bon pour une personne est limité par les possibilités motivationnelles et affectives de la personne. Et puisque les possibilités motivationnelles et affectives sont déterminées par ce que les êtres humains sont capables de désirer ou de valoriser, ce qui est bon pour une

²⁹⁴ “All that [internalisme about value] implies, then, may be that something can’t be good for someone if he’s constitutionally incapable of caring about it – if caring about it is beyond the affective capacities of a creature like him. And this conclusion allows that the thing can be good for him so long as the possibility of his caring about it isn’t ruled out by his very nature.” (VELLEMAN, 1998, p. 91).

personne humaine doit être quelque chose que des créatures humaines sont capables de désirer ou valoriser.

En conclusion, nous avons vu jusqu'à présent que l'internalisme du bien personnel dépend lui aussi, comme l'internalisme moral, de l'acceptation du principe conceptuel DIP. Nous allons voir maintenant pourquoi cette dépendance à DIP pose des difficultés pour n'importe quelle forme d'internalisme.

5.3.2. LA FAIBLESSE DE L'INTERNALISME DU BIEN PERSONNEL

Comme le souligne David Velleman, l'internalisme du bien personnel que nous avons présenté dans la précédente section est faible pour trois raisons : en premier lieu, cet internalisme stipule seulement que la personne doit être *capable* de désirer un objet ou de le valoriser, mais il ne stipule pas que cette capacité soit effectivement réalisée par la personne. Un objet peut être un bien pour une personne sans que la personne n'ait jamais effectivement adopté l'attitude rationnelle requise envers cet objet. Il suffit que ne soit pas exclue la possibilité pour cette personne de prendre cette attitude envers l'objet.

En second lieu, cet internalisme est faible car il concerne l'espèce (humaine) plutôt que les individus particuliers membres de l'espèce. Il faut que ce que l'individu *ne peut pas désirer ou valoriser* soit quelque chose que *l'espèce humaine* est incapable de désirer ou valoriser. Troisièmement, l'internalisme concerne exclusivement ce que l'espèce *est capable ou incapable* de désirer/valoriser, mais il n'affirme pas qu'il est *dans la nature* de l'espèce de désirer l'objet. Il est tout à fait possible qu'aucun membre de l'espèce n'ait jamais désiré un objet, mais que pourtant il soit possible pour cette espèce de désirer l'objet en question. Autrement dit, l'internalisme ne porte pas sur ce que les êtres humains désirent *dans les faits* ou au sujet de ce qui leur importe *dans les faits*, mais seulement ce que les êtres humains *sont capables* de désirer ou de valoriser.

Le problème soulevé par le fait que le principe DIP concerne ce que les membres d'une espèce *sont incapables* de désirer ou valoriser est le suivant : en ce qui concerne les êtres humains, il n'existe pas d'objets que les êtres humains sont *essentiellement* incapables de désirer/valoriser. De ce fait, même si nous admettons que le principe DIP est vrai, celui-ci est inutile pour partager ce qui est bon de ce qui n'est pas bon pour une personne humaine (VELLEMAN, 1998, p. 92).

Le raisonnement est le suivant :

- (1) Il n'existe rien que les êtres humains sont incapables de désirer/valoriser.²⁹⁵
- (2) Ce que les êtres humains sont incapables désirer/valoriser ne peut pas être un bien pour un être humain.
- (3) ∴ Il n'existe rien qui ne peut pas être un bien pour un être humain, en vertu du fait que l'être humain est incapable de le désirer/valoriser.

Si l'argument est concluant, nous parvenons à la conclusion suivante : le principe DIP est *vide* lorsque nous l'utilisons pour identifier ce qui est bien et ce qui n'est pas un bien pour l'être humain. Il s'agit d'un principe *inutile*, parce que *vide de contenu*, lorsqu'il est appliqué à la notion de « bien pour une personne humaine ».

5.3.3. LE BIEN DES CHATS ET DES PIERRES

Il est important de noter que David Velleman – et nous le suivons sur ce point – ne rejette pas le principe DIP ; selon lui, s'il existait effectivement des choses que les êtres humains s'avéraient incapables de désirer/valoriser, alors ces choses ne pourraient pas constituer des « biens pour » les êtres humains. Le problème provient du fait que, dans le cas de l'être humain, il n'existe pas de choses que ceux-ci sont incapables par nature de désirer/valoriser. Selon le philosophe, le problème de *vacuité* de DIP concerne donc *spécifiquement la nature humaine* ; si l'usage de DIP est *inutile* dans le cas de l'être humain, cela ne signifie pas qu'il soit inutile pour déterminer le bien d'autres espèces d'animaux ou le bien d'objets inanimés.

(i) Les chats et les pierres

Prenons le cas d'un objet inanimé, telle qu'une pierre ; dans son cas, *rien* ne peut être bon pour la pierre, car un objet inanimé est *par nature* incapable de se soucier, de désirer ou de valoriser quoi que ce soit. Aucun objet ne peut donc être un bien pour une pierre. Pour ce qui concerne les animaux, qui eux sont capables d'avoir des états mentaux conatifs et affectifs, nous pouvons déterminer un ensemble de choses que les animaux sont *incapables* par nature de désirer/valoriser. Par exemple, un chat est *essentiellement incapable* de se soucier de la notoriété ou de la gloire. Etant donné que la nature des chats les rend

²⁹⁵ Cette thèse rappelle celle de Descartes concernant l'absence de limites de la volonté humaine – par opposition aux limites de son entendement. (DESCARTES, 1897, p. 628).

incapables de se soucier ou de valoriser la notoriété et la gloire, alors la notoriété et la gloire ne peuvent pas être des *biens* pour les chats.²⁹⁶

Certaines créatures sensibles sont constitutivement incapables de se soucier de certaines choses, parce que ces choses dépassent leur capacité de compréhension. Par exemple, un chat ne peut pas se soucier d'être célèbre ; aucune chose capable de se soucier d'être célèbre ne peut être adéquatement qualifiée de chat (plutôt que d'être qualifiée de personne de forme féline). La présente conclusion implique donc que la gloire n'est pas intrinsèquement bonne pour un chat – une conséquence qui peut avoir de l'importance pour, par exemple, le traitement humain des animaux de foire.²⁹⁷

D'autres auteurs utilisent un raisonnement similaire – quoi que sous une formulation imprécise – concernant le fait qu'il existe des choses que les animaux sont *incapables* de désirer, pour conclure que ces choses ne peuvent pas constituer des *biens* ou des *torts* pour les animaux. Par exemple, il est en général admis que nous pouvons faire du tort à un être humain sans pourtant affecter son bien-être – par exemple en se moquant de lui (sans qu'il n'en ait conscience), en l'observant en cachette (sans qu'il n'en ait conscience), en ne respectant pas sa volonté (sans qu'il n'en ait conscience). Le problème qui se pose en éthique animale est de savoir s'il est également possible de faire du tort à un animal sans que son bien-être ne soit affecté – par exemple, en se moquant de lui, en teignant sa fourrure en couleur, ou en le déguisant d'une façon humiliante, etc. Or, il n'est pas possible de faire du tort à un animal lorsque son bien-être n'est pas affecté, par exemple en l'humiliant, car un animal – contrairement à un humain – est *incapable* de se soucier du fait d'être humilié ou d'être l'objet de moqueries.

A ces animaux non humains qui sont dépourvus d'une conscience de soi et d'une conception de leur statut moral, similaires à celles que possèdent les êtres humains, il n'est pas possible de manquer respect ou de les discriminer par des

²⁹⁶ Réciproquement, l'absence de notoriété et de gloire, ou la disgrâce, la diffamation et la baisse de popularité, ne peuvent pas être des maux pour les chats. Nous y reviendrons.

²⁹⁷ "Some sentient creatures are constitutionally incapable of caring about some things, because those things are necessarily beyond their grasp. For example, a cat cannot care about his being famous; nothing capable of caring about fame would qualify as a cat (rather than as a person in feline form). The present conclusion therefore implies that fame is not intrinsically good for a cat – a consequence that might bear significantly, say, on the humane treatment of show animals." (VELLEMAN, 1998, p. 92).

pratiques que nous considérerions comme humiliantes si elles concernaient des êtres humains. Etant donné que les animaux sont dépourvus de tout intérêt ou de tout désir d'être traités avec respect et de ne pas être humiliés [...], nous ne devons pas aux animaux le même niveau de respect à l'égard d'un traitement non dégradant que ce que nous devons à des êtres humains paradigmatiques.²⁹⁸

Ou encore :

La plupart des animaux sont dépourvus des capacités nécessaires pour [se sentir honteux et humiliés]. Ils ne comprennent pas ce qu'est le ridicule et, en conséquence, ne peuvent pas se sentir honteux ou humiliés. Donc, faire subir aux animaux un traitement que les êtres humains considéreraient comme étant ridiculisant s'ils en étaient eux-mêmes la cible, n'affecte pas le type d'êtres que les animaux sont et cela n'est pas en opposition avec leur statut moral, puisque les animaux ne disposent pas des désirs concernant la manière dont ils souhaitent être traités par les humains.²⁹⁹

Bien que ces citations fassent référence au fait que les animaux *ne se soucient pas* en fait de respect, de réputation ou d'humiliation, ce qui importe véritablement est que les animaux sont *incapables* de se soucier du respect, de réputation ou d'humiliation. De ce fait, étant par nature incapables de se soucier de ces choses, le respect, la réputation ou l'humiliation ne peuvent pas constituer des biens ou des maux pour les animaux – contrairement à ce qu'il en est pour les êtres humains.

(ii) Les êtres humains

Nous voyons donc que le principe DIP, appliqué à la notion normative de « bien pour » n'est pas complètement dépourvu de signification et d'implications normatives ; pour d'autres espèces que l'être humain, il permet de déterminer quelles choses peuvent être

²⁹⁸ “[T]hose nonhuman animals who lack both self-awareness and a conception of social standing similar to that of humans cannot be disrespected or discriminated against by practices that would be considered humiliating in the case of paradigmatic humans. Given that animals lack an interest or desire in being treated respectfully and in not being humiliated, [...], we do not owe animals the same level of respect concerning non-degrading treatment that we owe to paradigmatic humans.” (MARTIN, 2019, pp. 86-87).

²⁹⁹ “Most animals lack the prerequisite cognitive capacities for [feeling ashamed and humiliated]. They do not understand what ridicule is and, as a consequence, cannot feel ashamed or degraded. Thus, being exposed to what humans would consider to be ridicule if they were the targets of the action does not affect the kind of beings animals are and is not at odds with their moral status, since they lack corresponding desires about how they wish to be treated by humans.” (MARTIN, 2019, pp. 86-87).

bonnes ou non pour une espèce. La difficulté provient du fait que, dans le cas de l'humain, il n'existe rien qui « dépasse la compréhension humaine » telle que l'être humain serait incapable de le désirer. Comme le soutient David Velleman :

Malheureusement, cette conclusion [au sujet des autres espèces] ne nous permet pas de spécifier des conséquences analogues pour les personnes.³⁰⁰

De ce fait, si Velleman a raison, bien que le principe DIP-motivation soit analytiquement vrai, et que l'internalisme du bien personnel soit, en conséquence, aussi vrai, le principe ne s'applique à aucune situation de l'être humain. En ce qui concerne l'être humain, DIP s'avère donc *vide* et *inutile*.

Or, si le principe DIP-motivation est vide lorsqu'il concerne le bien personnel des êtres humains, le principe DIP-motivation est tout aussi *vide* lorsqu'il concerne le bien moral. En effet, s'il n'existe rien que l'être humain est incapable de désirer/valoriser, alors il n'existe pas de limites motivationnelles à ce qui peut être exigé d'un être humain. La prémisse de l'argument suivant est toujours fausse, ce qui a pour conséquence que le raisonnement ne sera jamais concluant :

(P1)	Un être humain ne peut pas être motivé à faire X	Prémisse descriptive
(P2)	Si un être humain ne peut pas être motivé à faire X, alors il n'a pas le devoir moral de faire X	Prémisse conceptuelle (DIP-motivation)
(C)	∴ un être humain n'a pas le devoir moral de faire X.	Conclusion normative

Si nous avons raison, alors l'usage que font les empiristes du principe DIP-motivation est illégitime ; tout raisonnement empiriste fondé sur ce principe repose sur une prémisse fausse.

5.4. PREMIÈRE CRITIQUE : COMMENT DÉTERMINER L'INCAPACITÉ ?

Nous allons maintenant présenter et répondre à un ensemble de critiques qui peuvent être formulées à l'encontre de l'objection de la vacuité que nous avons soulevée contre le principe DIP-motivation, lorsque ce principe porte sur les motivations humaines. Nous allons

³⁰⁰ "Unfortunately, we cannot similarly specify the analogous consequences that this conclusion might lead to for people" (VELLEMAN, 1998, p. 92)

étudier trois types d'objections : la première critique (5.4) met en cause le statut de l'affirmation selon laquelle il n'existe rien que l'être humain soit incapable de désirer – contrairement à ce qui est le cas pour chats et les pierres. Selon cette critique, l'affirmation ne concerne qu'un problème *épistémique* et pas du tout un problème relatif à la *nature humaine*. La seconde critique (5.5) conteste l'affirmation selon laquelle il n'existe rien que l'être humain soit incapable de désirer/valoriser. Il existerait *au moins un désir* général que l'être humain est incapable d'avoir, à savoir le désir général « de vivre une vie entière qui ne contient rien de ce qui est bon pour moi ». Nous répondrons longuement à cette critique. La troisième critique (5.6) conteste également l'affirmation selon laquelle il n'existerait pas de limites motivationnelles connaissables concernant les êtres humains. Selon cette critique, les sciences empiriques, notamment la biologie, permettent de déterminer un ensemble de choses que les êtres humains sont incapables de désirer. Nous verrons que cette dernière critique n'est pas non plus légitime. Nous concluons donc que, même si le principe DIP exprime une vérité conceptuelle – ce que nous croyons, il ne peut pas être employé *par les empiristes* pour rejeter des théories morales trop exigeantes.

Nous nous sommes contentés, jusqu'à présent, d'affirmer qu'il n'existe pas, pour l'être humain, le même genre de limites motivationnelles ou affectives que celles qu'il existe, par exemple, pour les pierres ou les chats. Il n'existe rien, disions-nous, que les êtres humains sont, en vertu de leur nature, incapables de désirer ou de valoriser. Comment pouvons-nous savoir qu'il n'existe pas de telles limites ? Nous ne l'avons, jusqu'à présent, pas expliqué.

Nous avons déterminé les limites motivationnelles des pierres par le simple fait que les pierres sont incapables de désirer ou valoriser *quoi que ce soit*. Il n'existe donc rien qui soit un bien ou un mal pour les pierres. Pour les chats, nous avons déterminé leurs limites motivationnelles en référence à leur capacité de *compréhension* : les chats ne peuvent pas comprendre l'humiliation et le respect, et de ce fait l'humiliation et le respect ne peuvent pas être, respectivement, un bien ou un mal pour eux. Si nous étendons ce critère à l'être humain, il faudrait pouvoir identifier un objet dont la représentation *dépasse* ses capacités de compréhension, de sorte qu'il serait impossible pour celui-ci de désirer l'objet. Affirmer qu'il n'existe rien que l'être humain est incapable de désirer revient donc à affirmer qu'il n'existe rien qui dépasse les capacités de compréhension de l'être humain. Mais comment savons-nous cela ?

En fait, nous n'en savons rien et, en tant qu'êtres humains, nous ne *pouvons pas le savoir*. En tant qu'êtres humains, nous ne pouvons pas identifier les objets qui, potentiellement, dépassent nos capacités de compréhension. En effet, pour identifier un tel objet – dont la représentation dépasserait nos capacités de compréhension – il faudrait déjà avoir la capacité de se le représenter. Ainsi, il est impossible, pour un être humain, de savoir quelles sont les objets qui outrepassent sa capacité de désirer. Or, étant donné que nous ne pouvons pas identifier nos propres limites motivationnelles, nous ne pourrions jamais savoir quelles sont les choses qui ne peuvent pas être des biens pour nous, êtres humains.

Si certaines choses sont dépourvues de valeur pour l'être humain parce qu'elles dépassent la capacité de compréhension de l'être humain, nous, les êtres humains, ne sommes pas en position de dire ce qu'elles sont.³⁰¹

Bien entendu, il est possible qu'il existe une espèce d'êtres ayant des capacités de compréhension de nature complètement différente que les nôtres, et qui, de leur perspective, peuvent identifier des objets que les êtres humains sont incapables de saisir – et donc de désirer. Mais, de notre perspective d'être humain, une telle identification reste impossible. En conclusion, il n'existe rien de *connaissable* que les êtres humains sont incapables de désirer.

Cet argument fait subir un infléchissement à la force du raisonnement initial ; alors que nous prétendions montrer qu'il n'existe rien que les êtres humains sont incapables de désirer, nous affirmons désormais seulement que l'être humain est incapable de *connaître* ses propres limites motivationnelles ; il est incapable d'identifier ce qui dépasse ses capacités de compréhension. Nous sommes donc passés d'un argument ontologique – i.e. au sujet de ce que les êtres humains *sont incapables* de désirer – à un argument seulement épistémique – i.e. au sujet de ce que les êtres humains peuvent connaître à propos de leur propres limites motivationnelles.

Ainsi, contrairement à ce que nous prétendions, il existe peut-être quelque chose les êtres humains sont incapables de désirer. Nous ne pouvons pas l'exclure. Tout ce que nous pouvons dire, c'est que nous ne pouvons pas *savoir*, en tant qu'être humain, s'il existe quelque chose que les êtres humains sont incapables de désirer.

³⁰¹ "If some things lack human value because they lie beyond human grasp, we humans are in no position to say what they are. [...]" (VELLEMAN, 1998, pp. 91-92)

Bien que la conclusion présente soit moins forte que la précédente, elle nous semble néanmoins suffisante pour affirmer qu'aucune référence à des limitations motivationnelles de l'être humain ne peut être avancée – par un être humain – pour contester qu'un objet puisse être un bien pour une personne ou pour critiquer des exigences morales qui outrepasseraient les capacités motivationnelles de l'homme. Personne ne peut prétendre que « il n'est pas le cas que X est bon pour S » ou que « il n'est pas le cas que S a le devoir moral de faire X » sur la base d'une prétendue, mais inconnaissable, limitation motivationnelle humaine. Dans les deux cas – i.e. concernant le bien d'une personne et concernant ses devoirs moraux – aucun être humain ne peut savoir quelles sont les choses qu'il nous est impossible, en tant que créature humaine, de désirer/valoriser.

5.5. SECONDE CRITIQUE : L'IMPOSSIBILITÉ DE VOULOIR RENONCER À SON PROPRE BIEN

Un critique pourrait néanmoins ajouter : certes, il est vrai qu'il n'existe aucune chose *particulière* que les êtres humains sont incapables de désirer. Il est même possible, pour un être humain, de désirer des choses qui ne sont pas bonnes pour lui – contre les exigences de rationalité – et même de désirer des choses qu'il *sait* être mauvaises pour lui (par exemple, fumer, devenir gros, vivre en isolation totale, ...). Cependant, il est impossible, pour un être humain, de désirer vivre une existence entière qui ne contiendrait rien de ce qui est bon pour lui. Même si, de fait, il a certains désirs qui portent sur des choses qui sont contraires à son propre bien, il ne peut pas avoir le désir général de « vivre une vie dépourvue de toutes les choses qui sont bonnes pour moi ».

Ainsi, même s'il n'existe pas de désir spécifique dont on peut dire que l'être humain est *incapable* d'avoir un tel désir, il reste néanmoins possible d'identifier quelque chose que l'être humain est incapable de désirer, à savoir « vivre une vie qui ne contient rien de ce qui est bon pour soi ».

Notre critique continuerait son objection en disant que ce désir nouvellement identifié, que l'être humain est incapable d'avoir, est insignifiant en ce qui concerne la détermination de ce qui est « bon pour » une personne humaine. En effet, l'identification de ce désir nous permet uniquement de d'obtenir la conclusion *triviale* suivante :

(P1)	Un être humain est incapable de désirer vivre une vie qui ne contient rien de ce qui est bon pour lui.	Prémisse descriptive
(P2)	Ce que l'être humain est incapable de désirer ne peut pas être un bien pour lui	Internalisme du bien personnel (basé sur DIP)
(C1)	∴ Vivre une vie qui ne contient rien de ce qui est bon pour lui ne peut pas être un bien pour l'être humain.	Conclusion normative

Or, il semble trivialement vrai qu'une vie qui ne contient rien de bon pour soi ne peut pas être un bien pour soi. Ainsi, bien que le principe « devoir-implique-pouvoir » ne soit désormais plus *vide*, puisque nous avons identifié un désir que l'être humain est incapable d'avoir, il reste que l'usage de DIP dans la détermination du contenu de la notion de « bien pour » permet de tirer uniquement une conclusion triviale.

En revanche, et notre critique tient un point important cette fois, DIP *cesse d'être vide* et ne donne pas lieu à une thèse *triviale* lorsque nous l'appliquons à des concepts moraux, plutôt qu'au concept de « bien pour ». En effet, dans le contexte moral, l'identification d'un désir que l'être humain est incapable d'avoir, combiné à l'internalisme *moral*, permet de construire le raisonnement suivant :

(P1)	Un être humain est incapable de désirer vivre une vie qui ne contient rien de ce qui est bon pour lui.	Prémisse descriptive
(P2)	Un être humain n'a pas le devoir moral de faire une action qu'il est incapable de désirer. ³⁰²	Internalisme moral (basé sur DIP-motivation)
(C)	∴ Un être humain n'a pas le devoir moral de vivre une vie qui ne contient rien de ce qui est bon pour lui.	Conclusion normative

Nous voyons que, dans le cas de l'internalisme moral, le principe DIP-motivation *n'est pas vide* et ne donne pas lieu à une conclusion *triviale*. En effet, la conclusion obtenue

³⁰² Ou à laquelle il est incapable d'accorder de la valeur.

permet d'exclure les doctrines morales qui exigeraient d'un individu de sacrifier l'intégralité ou une trop grande part de ce qui est bon pour lui. L'internalisme moral – soutenu par DIP-motivation – permet donc de critiquer les doctrines morales qui sont trop exigeantes, en vertu du fait qu'elles exigent des individus de sacrifier une trop grande quantité de leur propre bien.³⁰³

5.5.1. DEUXIÈME VERSION DE LA CRITIQUE : DÉSIRER CE QUE L'ON CROIT ÊTRE SON PROPRE BIEN

Il semble difficile de contester qu'un être humain ne peut pas désirer vivre une existence qui ne contient rien de bon pour lui. Cependant, il est toujours possible qu'un individu *ait une conception erronée* de son propre bien. Ainsi, nous pouvons tout à fait envisager une personne qui a une conception complètement fausse de ce à quoi ressemble une existence qui est bonne pour elle ou une existence qui contient des choses bonnes pour elle. De ce fait, il est inadapté de décrire le contenu du désir en termes *objectifs*. Il est en effet tout à fait possible qu'une personne désire avoir une vie qui ne contient rien de ce qui est (objectivement) bon pour elle. Ce qui est impossible, en revanche, c'est qu'une personne désire avoir une vie qui ne contient rien de ce qu'elle *croit* (subjectivement) être bon pour elle – mais au sujet duquel elle peut se tromper.

En général, c'est d'ailleurs en ces termes (subjectifs) qu'est décrit, dans la littérature, le désir que les êtres humains sont incapables d'avoir. Par exemple, nous trouvons dans un texte d'Elisabeth Anscombe l'idée qu'il est impossible pour un être humain de ne jamais rien désirer de ce qu'il *croit* être bon pour lui.

Il est possible de ne pas vouloir quelque chose dont on estime avoir besoin. Mais [...] il n'est pas possible que jamais on ne désire quoi que ce soit dont on estime avoir besoin.³⁰⁴

De la même manière, lorsque Bernard Williams soutient qu'on ne peut exiger d'une personne trop de sacrifices, il fait référence à la conception *subjective* que la personne se fait de son propre bien :

³⁰³ C'est par cet argument que nous sommes amenés à considérer que l'exigence de ne pas formuler des principes qui requièrent trop de sacrifices constitue un cas particulier de l'exigence générale de ne pas formuler des principes qui soient inadaptés aux motivations humaines.

³⁰⁴ "It is possible not to want something that you judge you need. But [...] it is not possible never to want anything you judge you need." (ANSCOMBE E. , 1958, p. 7).

Nous pouvons parvenir à un point auquel il devient déraisonnable pour un homme d'abandonner, au nom de l'ordre impartial du bien du monde des agents moraux, quelque chose qui est une condition pour qu'il soit intéressé à simplement vivre encore dans ce monde.³⁰⁵

Ainsi, l'être humain est incapable de désirer vivre une vie qui ne contient rien de ce qu'il *croit* être bon pour lui, bien qu'il soit tout à fait capable de désirer vivre une vie qui ne contient rien de ce qui est objectivement bon pour lui – puisqu'il peut complètement se tromper au sujet de son propre bien. Suite à cette reformulation du contenu du désir qu'un être humain est incapable d'avoir, nous parvenons à la conclusion morale suivante : « il ne faut pas exiger d'un individu de sacrifier l'ensemble de ce qu'il *croit* être bon pour lui ».

Il est possible de résumer le raisonnement comme suit :

(P1)	Un être humain est incapable de désirer vivre une vie qui ne contient rien de ce qu'il <i>croit</i> être bon pour lui.	Prémisse descriptive
(P2)	Si S ne peut pas désirer faire X, alors S n'a pas le devoir moral de faire X.	Internalisme moral (basé sur DIP-motivation)
(C)	∴ Un être humain n'a pas le devoir moral de vivre une vie qui ne contient rien de ce qu'il <i>croit</i> être bon pour lui.	Conclusion normative

Selon le présent raisonnement, il existe bien quelque chose que l'être humain est incapable de désirer – et qui ne peut donc pas être l'objet d'un devoir moral. De ce fait, l'objection de la *vacuité* n'est pas concluante.

Pour répondre à cette objection, notons que le raisonnement repose à nouveau sur DIP-motivation. Or, nous pouvons interpréter DIP-motivation et, avec lui, la conclusion à laquelle il mène ci-dessus, de trois manières différentes, qui toutes posent un problème³⁰⁶ :

³⁰⁵ “[T]here can come a point at which it is quite unreasonable for a man to give up, in the name of the impartial good ordering of the world of moral agents, something which is a condition of his having any interest in being around in the world at all.” (WILLIAMS, 1981, p. 14).

³⁰⁶ Il est à nouveau possible d'interpréter DIP de trois manières différentes car nous discutons de nouveau de notions morales et non plus de la notion de « bien pour ». Or, dans le cas des notions morales,

soit nous interprétons DIP-motivation comme un principe conceptuel (5.5.2), soit comme un principe pragmatique, selon lequel instituer des normes trop exigeantes met en danger la stabilité sociale (5.5.3), soit nous interprétons DIP-motivation comme un principe moral, selon lequel il est cruel d'exiger de quelqu'un de sacrifier tout ce qui a de la valeur à ses yeux (5.5.4). Dans ces trois cas, toutefois, la solution proposée fait face à des difficultés.

Notons que nous ne reconsidérons pas, dans ces lignes, les mêmes arguments que ceux que nous avons étudiés lorsque nous traitons du principe DIP dans sa forme générale. Nous nous intéressons uniquement aux arguments qui pourraient être avancés pour justifier *la conclusion particulière* selon laquelle une théorie morale est déficiente – soit du point de vue conceptuel, soit d'un point de vue pragmatique, soit d'un point de vue moral – si elle exige des êtres humains de sacrifier *tout ce qu'ils croient être bons pour eux*. Or, comme la signification de cette conclusion varie selon la signification de DIP-motivation, nous nous intéresserons aux différentes manières dont peut être interprété DIP-motivation dans ce contexte.

5.5.2. INTERPRÉTATION CONCEPTUELLE DE DIP-MOTIVATION

Dans le premier cas, nous adoptons l'interprétation *conceptuelle* de DIP-motivation ; selon cette interprétation, il est *contradictoire* d'exiger des êtres humains de faire quelque chose qu'ils sont incapables de désirer – et donc d'exiger d'eux qu'ils sacrifient l'intégralité de ce qu'ils *croient* être leur propre bien. Le problème, avec cette interprétation, est le suivant ; il est possible que *chaque personne* ait une conception différente de ce qu'est son propre bien. De ce fait, les exigences morales deviennent relatives à la conception que chaque personne se fait de son propre bien. Les exigences morales peuvent donc varier entre deux personnes. Elles peuvent varier même pour une même personne au cours de sa vie. Nous sommes donc contraints d'adopter une forme extrême de subjectivisme moral : le devoir de chacun est relatif à la conception que chacun se fait de son propre bien, tout au long de sa vie.

En outre, étant donné que chacun peut avoir une *croyance fausse* au sujet de ce qui est bon pour lui, les normes morales peuvent dépendre de conceptions outrageuses ou

nous l'avons vu, il n'est pas évident – contrairement au cas de la notion de « bon pour » – qu'elles entretiennent une relation *conceptuelle* quelconque avec la motivation de la personne.

déraisonnables du bien personnel. Par exemple, un individu peut croire qu'il est bon pour lui de ne pas être en contact avec des noirs et des arabes, qu'il est bon pour lui que sa femme lui soit soumise et obéisse à l'ensemble de ses ordres, qu'il est bon pour lui d'avoir des ouvriers sous-payés, qu'il est bon pour lui de manger exclusivement de la nourriture industrielle tous les jours et de fumer. Si nous admettons qu'un tel homme existe, alors, en vertu de DIP, notre homme ne peut pas avoir le devoir moral à la fois d'accepter de vivre avec des noirs et des arabes, de ne pas traiter sa femme comme une esclave, de payer ses ouvriers décemment, et il ne peut pas avoir le devoir de prudence de ne pas manger exclusivement de la nourriture industrielle et d'arrêter de fumer. Respecter de telles exigences morales et prudentielles impliquerait, pour notre homme, de renoncer à tout ce qu'il croit être bon pour lui – ce qu'interdit DIP-motivation.

Pour éviter le subjectivisme moral, nous pourrions considérer qu'il ne faut tenir compte que des conceptions du bien personnel que *tous* les êtres humains partagent – santé, richesse, amitié, plaisir, l'honneur, ... Les normes morales ne peuvent pas exiger de sacrifier l'ensemble de ce que *tous les êtres humains*, ou *la majorité des êtres humains*, croient être bons pour eux. Mais, dans ce cas, il devient nécessaire d'identifier des objets *spécifiques* qui sont considérés par tous les êtres humains comme étant constitutifs de leur propre bien. La liste des conceptions subjectives susceptibles de constituer des limites aux sacrifices qui peuvent être exigés d'un être humain deviendra plus courte et plus précise.

Par exemple, il est probable que tous les êtres humains considèrent qu'il est bon pour eux d'être en bonne santé, d'avoir des amis, d'éprouver du plaisir et de ne pas être objet de déshonneur, d'être riche et d'éprouver de l'amour. Or, nous avons vu que, pour tout objet *spécifique* constitutif de la conception du bien personnel, il est possible pour un être humain de ne pas le désirer. Autrement dit, un être humain peut toujours ne pas désirer une chose *spécifique* qu'il considère pourtant comme étant bonne pour lui. Par exemple, un être humain peut croire que l'amitié est bonne pour lui, mais pourtant ne pas désirer avoir d'amis. Cette position peut être condamnée comme *déraisonnable*, mais il n'est pas impossible pour un être humain d'avoir une telle attitude. Il est même possible qu'un être humain ne désire *aucun* des éléments de la liste courte et précise de ce que tous les êtres humains considèrent comme étant bons pour eux. En conséquence, il n'est pas vrai qu'un

être humain est incapable de désirer vivre une vie qui ne contient rien de la liste courte et précise de ce que tous les êtres humains considèrent comme étant bon pour eux.

Ajoutons encore que, parmi les conceptions du bien qui sont partagées par l'ensemble des êtres humains, il est possible, à nouveau, que figurent des conceptions outrageuses et déraisonnables. Il est possible que tous les êtres humains croient, par exemple, qu'il est bon pour eux d'avoir des esclaves qui travaillent à leur compte et que tous les êtres humains croient que le fait d'amasser de la richesse sans limite est bon pour eux. Or, il ne semble pas que nous voudrions que les normes morales soient restreintes par de telles conceptions du bien personnel.

Si nous défendons la thèse selon laquelle une théorie morale est fausse si elle exige d'un être humain de sacrifier tout ce qu'il croit être bon pour lui, il n'est pas possible de le faire en s'appuyant sur une interprétation *conceptuelle* de DIP-motivation, car cela impliquerait que les normes morales varieraient en fonction des conceptions subjectives du bien personnel.

5.5.3. INTERPRÉTATION PRAGMATIQUE DE DIP-MOTIVATION

La seconde solution consiste à interpréter DIP-motivation, appliqué aux conceptions subjectives du bien personnel, comme un principe pragmatique : il ne faut pas exiger trop de sacrifices aux personnes en termes de leur conception de leur propre bien, car cela risque de mener à des turbulences sociales.³⁰⁷ Si une société est régie par des normes sociales trop exigeantes pour les êtres humains, cela risque d'avoir pour conséquence que...

...la norme sera souvent violée, et son respect donnera lieu à une bonne dose d'anxiété, ceux qui respecteront la norme sans éprouver d'anxiété seront étranges sous d'autres aspects, et ainsi de suite. Cela suffirait pour fournir un

³⁰⁷ Notons que nous renonçons ici à l'interprétation conceptuelle de DIP ; l'énoncé « S a le devoir moral de sacrifier tout ce qu'il croit être bon pour lui » n'est pas rendu faux par le fait qu'il est impossible pour S de désirer sacrifier tout ce qu'il croit être bon pour lui. Cependant, il socialement risqué d'exiger des personnes de renoncer à ce qu'elles croient être leur propre bien, car cela risque de générer une insatisfaction et une instabilité sociale. En effet, si trop de personnes n'obtiennent pas une existence qu'elles considèrent comme étant bonne pour elles, alors elles éprouveront de l'insatisfaction, ce qui risque de dégénérer en mouvements sociaux déstabilisateurs.

argument puissant contre le fait d'ériger un tel comportement en norme pour une société humaine.³⁰⁸

En réponse à cette interprétation pragmatique de DIP, il suffit de noter qu'il est tout simplement faux que tenir compte des croyances individuelles au sujet de ce qui est bon pour soi, c'est-à-dire tenir compte des conceptions *subjectives* du bien personnel, constitue une bonne manière d'éviter les turbulences et instabilités sociales. Au contraire, il y a de fortes chances que, pour éviter les tensions sociales, il faille considérer le bien personnel *objectif* des personnes, plutôt que le bien personnel *subjectif*. Eviter d'exiger trop de sacrifices *subjectivement* conçus n'empêche aucunement l'apparition de turbulences sociales. Pour empêcher l'émergence de turbulences sociales, il faut considérer les intérêts *objectifs* de tous les citoyens.³⁰⁹

En effet, du point de vue social, la négligence des intérêts objectifs de certains groupes d'individus peut amener à l'apparition de troubles sociaux *sans que les individus n'aient la*

³⁰⁸ "[...] the norm will often be broken, its observance will give rise to a good deal of anxiety, those who comply without anxiety to the norm will be unusual in other respects, and so forth. This [would be] enough to provide a strong argument against the behaviour being made into a norm for a human society [...]" (WILLIAMS, 1995, p. 101).

³⁰⁹ Rien dans notre propos n'interdit que la frustration soit *également* générée par la négligence d'intérêts subjectivement conçus, lorsque la conception *subjective* des intérêts correspond à la conception *objective*. En effet, il est possible que, ayant une conception adéquate de leurs intérêts, les individus dont les intérêts *objectifs* sont négligés expérimentent une frustration supplémentaire du fait de la négligence de leurs intérêts *subjectivement* conçus.

Pour obtenir un contre-exemple à notre propos, il faudrait que des turbulences sociales puissent résulter de la frustration d'intérêts exclusivement *subjectifs*, qui ne correspondent pas à des intérêts objectifs. Par exemple, une interdiction générale de fumer pourrait bien susciter de violentes turbulences sociales, à cause de la frustration des intérêts conçus *subjectivement*, sans pourtant que cette interdiction ne porte atteinte aux intérêts *objectifs* des personnes. Et cela risque d'être *d'autant plus le cas* avec une théorie qui exigerait de sacrifier *tout* ce que les personnes considèrent (subjectivement) comme étant bons pour elles. Autrement dit, une doctrine morale qui exigerait de sacrifier l'ensemble des intérêts *subjectifs* risque tout autant de susciter des turbulences sociales qu'une doctrine morale qui exige de sacrifier l'ensemble des intérêts objectifs. Il est donc faux que les préoccupations pragmatiques au sujet des turbulences sociales devraient concerner les intérêts objectifs exclusivement. Elles devraient concerner *aussi bien* les intérêts objectifs que subjectifs.

Pour répondre à cette objection, notons que plus les intérêts subjectivement conçus sont moralement problématiques, plus il devient difficile de considérer que l'objectif d'éviter les turbulences justifie la prise en compte de tels intérêts subjectifs. Par exemple, il se peut que, dans une communauté, la majorité développe l'idée que la société se porterait mieux si les individus en surpoids – i.e. IMC de plus de 25 – étaient tous liquidés. Si le gouvernement décide de négliger de tels intérêts subjectifs de la majorité, il se peut bien que cela crée des turbulences sociales également. Néanmoins, il semble difficile de dire qu'une théorie politique appropriée doive tenir compte de l'opinion moralement douteuse de ces citoyens. Comme nous le verrons plus tard, la sélection des intérêts subjectifs *dont il faut tenir compte* et *dont il ne faut pas tenir compte* repose elle-même sur des considérations morales et ne peut donc pas être utilisées pour identifier les exigences morales.

croyance que leurs intérêts sont négligés (et sans la croyance que l'état social dans lequel ils vivent est injuste). Comme le montre Peter Railton :

De la même manière qu'un individu qui néglige ses intérêts objectifs fera l'expérience de formes d'insatisfaction, un arrangement social – par exemple, une forme de production, de hiérarchie sociale ou politique, etc. – qui s'écarte de la rationalité sociale en négligeant les intérêts objectifs d'un groupe particulier donnera potentiellement lieu à de l'insatisfaction et à du désordre [...].³¹⁰

Une telle insatisfaction et un tel désordre social se produiront même si les citoyens n'ont pas la croyance que leurs intérêts sont négligés. Il suffit que leurs intérêts objectifs soient *effectivement négligés* pour que surgisse du désordre.

Si dans les faits les intérêts objectifs de certains groupes sont négligés, il y aura potentiellement du désordre qui se manifestera de différentes manières – par l'aliénation, la perte de moral, un déclin de l'efficacité de l'autorité, et cela bien avant que ne s'opère un quelconque changement dans les *croyances* des individus au sujet du caractère injuste de la société [i.e. au sujet de la prise en compte de leurs intérêts] [...].³¹¹

Selon Railton, il est même possible d'identifier les sociétés dans lesquelles les intérêts d'une partie de la population sont négligés par l'apparition des symptômes sociaux – et non, par exemple, en conduisant une enquête sur la manière dont les gens se représentent leur condition.

En plus du potentiel de désordre social, les sociétés qui échouent à se rapprocher de la rationalité sociale [i.e. qui échouent à prendre en compte des intérêts objectifs de chacun] peuvent partager aussi d'autres caractéristiques : elles peuvent exhiber une tendance à l'apparition de doctrines religieuses ou idéologiques, ou à l'apparition d'appareil répressifs de certains types ; elles

³¹⁰ "Just as an individual who significantly discounts some of his interests will be liable to certain sorts of dissatisfaction, so will a social arrangement-for example, a form of production, a social or political hierarchy, etc.-that departs from social rationality by significantly discounting the interests of a particular group have a potential for dissatisfaction and unrest." (RAILTON, 1986, p. 191).

³¹¹ "If in fact the interests of certain groups are being discounted, there will be a potential for unrest that may manifest itself in various ways-in alienation, loss of morale, decline in the effectiveness of authority, and so on-well before any changes in belief about the society's justness occur, and that will help explain why members of certain groups come to believe it to be unjust, if in fact they do." (RAILTON, 1986, pp. 191-192)

peuvent devenir moins productives sous un certain angle (par exemple, en ne parvenant pas à développer certaines ressources humaines) et plus productives sous d'autres aspects (par exemple, en parvenant à extraire plus de travail d'un certain groupe de travailleurs, à moindre coût), et peuvent donc présenter une réussite économique différente, selon les conditions de production auxquelles elles font face, etc. [...].³¹²

En résumé, du point de vue social, ce n'est pas tant la négligence des intérêts *subjectifs* qui compte pour déterminer le potentiel de désordre social que contient un arrangement social, mais plutôt la négligence des intérêts objectifs des personnes – i.e. de ce qui est objectivement bon pour elles, à savoir ce qu'elles désireraient pour elles-mêmes si elles étaient pleinement rationnelles.

Puisque ce n'est pas la négligence des intérêts *subjectifs* qui compte, mais la négligence des intérêts *objectifs*, l'idée selon laquelle il est socialement dangereux d'exiger des individus de sacrifier tout ce qu'ils *croient* être bon pour eux est fausse. L'interprétation *pragmatique* de DIP-motivation échoue donc. Il n'existe pas de raison pragmatique de ne pas exiger des individus de sacrifier tout ce qu'ils croient être bon pour eux.

5.5.4. INTERPRÉTATION MORALE DE DIP-MOTIVATION

Finalement, si nous interprétons plutôt DIP-motivation comme étant un principe *moral*, de sorte qu'il serait *immoral* d'exiger des individus de sacrifier l'ensemble de ce qu'ils pensent être bon pour eux, alors cette interprétation fait face à l'objection de l'explication évolutionniste que nous avons vue précédemment. Notre croyance qu'il est immoral d'exiger des individus trop de sacrifices de leur propre bien – subjectivement conçu – s'explique par le fait qu'il est avantageux d'avoir cette croyance en termes de réplication des gènes. Cette croyance n'est donc pas justifiée.

³¹² "In addition to possessing a certain sort of potential for unrest, societies that fail to approximate social rationality may share other features as well: they may exhibit a tendency toward certain religious or ideological doctrines, or toward certain sorts of repressive apparatus; they may be less productive in some ways (for example, by failing to develop certain human resources) and more productive in others (for example, by extracting greater labor from some groups at less cost), and thus may be differentially economically successful depending upon the conditions of production they face, and so on." (RAILTON, 1986, p. 192).

En conséquence, l'interprétation morale de DIP-motivation doit également être rejetée : il n'est pas *immoral* d'exiger des personnes de sacrifier tout ce qu'elles *croient* être bon pour elles.

Nous avons jusqu'à présent discuté deux critiques de l'objection de « vacuité » formulée à l'encontre de l'usage empiriste du principe DIP. La première reprochait à l'objection de ne concerner que ce que nous *pouvons connaître* au sujet des limitations motivationnelles, plutôt que ce que nous *connaissions* à leur sujet. La seconde critique reprochait à l'objection d'ignorer un désir spécifique qu'un être humain est incapable d'avoir, à savoir celui de « vivre une vie entière qui ne contient rien de ce qui est bon pour soi ». Nous avons vu que nous pouvions répondre à ces deux critiques. Nous allons maintenant en considérer une troisième et nous lui répondrons dans la dernière section de ce chapitre.

5.6. TROISIÈME CRITIQUE : LES LIMITES BIOLOGIQUES AUX EXIGENCES SOCIALES

Une critique additionnelle – proche de la critique précédente – de l'argument de Velleman, selon lequel DIP est vrai mais *vide* – consiste simplement à contester que le principe soit vide. Certes, selon la critique, il faut admettre que nous ne pouvons pas identifier les incapacités motivationnelles de l'être humain qui sont dues à des incapacités *de compréhension*. Il n'est pas possible d'identifier des désirs qui outrepassent les capacités de compréhension de l'être humain, à la manière dont nous identifions des désirs (ex. la gloire) qui outrepassent les capacités de compréhension des chats. Néanmoins, il existe des limites motivationnelles issues de notre nature *biologique*. Il s'agit donc de limites *biologiques*, plutôt que de limites résultant d'une incapacité de *compréhension*.

Les contraintes biologiques sont telles que *les êtres humains en général* sont incapables de désirer un certain objet. Cela n'exclut pas qu'il existe *des cas anormaux* d'êtres humains qui désirent précisément ces choses que les êtres humains sont en général (en raison de leur nature biologique) incapables de désirer. Néanmoins, la nature biologique des êtres humains restreint ce que l'être humain *normalement constitué* peut désirer. Et cette limitation biologique générale permet de formuler des limites à ce qui peut être exigé des êtres humains. Bernard Williams propose un argument dans ce sens :

Il peut exister des formes de comportement que les êtres humains sont tout à fait capables de vouloir et, même, à une échelle individuelle et limitée, de réaliser, mais qui, pour des raisons biologiques, sont condamnées à être psychologiquement coûteuses, ou à être confinées à un petit groupe d'individus peu communs sous d'autres rapports, ou qui sont condamnées à échouer en tant qu'institutions sociales *générales*.³¹³

Nous trouvons également ce type de position chez Clavier, qui préconise de rejeter les théories morales qui ne sont pas informées par ce que les sciences nous disent au sujet de la nature humaine.³¹⁴

Inversement, les données empiriques peuvent aider à cerner le champ des systèmes moraux viables. Par exemple, si l'on dispose de connaissances précises relatives aux limites de certaines capacités humaines, il est possible de poser des restrictions aux obligations morales imposées aux êtres humains (à condition bien sûr d'admettre que tout système moral crédible se doit d'être praticable) (CLAVIER, 2007, p. 5).

Quelles sont ces limites ? Clavier suggère, par exemple, que nous pourrions rejeter une doctrine morale attribuant aux individus de lourdes responsabilités morales, du fait qu'une telle doctrine s'opposerait à la constitution psychologique humaine.

Imaginez que des données psychologiques permettent de montrer l'effet dévastateur du sentiment de responsabilité illimitée sur la santé psychique des gens ; dans ce cas, une morale prônant ce principe serait à rejeter (CLAVIER, 2007, p. 5).³¹⁵

³¹³ "[T]here could be patterns of behaviour which human beings are entirely capable of wanting and indeed, on an individual or limited scale, of achieving, but which for biological reasons are bound to be psychologically costly, or confined to a small group of otherwise unusual individuals, or otherwise bound to fail as general social institutions" (WILLIAMS, 1995, p. 108)

³¹⁴ Notons que Christine Clavier ne défend pas une théorie morale réaliste. Selon elle, il faut renoncer aux doctrines morales qui ne tiennent pas compte de la nature humaine, non parce que cela violerait un principe moral d'humanité, mais parce qu'une telle doctrine ne pourrait pas être implémentée de façon durable. Ce sont donc des considérations pratiques qui motivent sa position. Les exemples que Clavier propose constituent néanmoins d'excellentes sources d'inspiration pour notre propos.

³¹⁵ Un tel argument contre une responsabilisation individuelle exacerbée, pourrait servir à redéfinir la "division du travail" moral entre l'Etat et les citoyens. En effet, pour éviter d'attribuer une responsabilité morale trop lourde aux citoyens, il faudrait déléguer à l'Etat, au nom de la « contrainte humaniste », la prise en charge des objectifs moraux. Nous retrouvons ainsi l'idée qu'il revient à l'Etat de garantir la réalisation des objectifs moraux, en rendant compatible l'adoption d'une vie morale et l'accomplissement du bien individuel ;

De la même manière, Clavien soutient que, s'il est prouvé que les êtres humains *ont besoin* ou *sont naturellement* enclins à donner à autrui sans rien attendre en retour, alors les doctrines morales *de l'égoïsme* doivent être rejetées, car elles sont contraires à la nature humaine.

Ou alors, imaginez qu'une grande masse de données empiriques laissent penser que les êtres humains possèdent une tendance génétique à aider non seulement leurs proches parents mais également des personnes inconnues sans attente de retour de service (c'est-à-dire sans motif égoïste sous-jacent), alors les théories morales qui reposent sur une conception de l'être humain comme un calculateur égoïste [...] sont à rejeter (CLAVIEN, 2007, p. 224).

Tout comme il est psychologiquement insupportable, pour un être humain, de se voir attribuer une responsabilité trop grande, ou de poursuivre une éthique purement égoïste, il est également, contre nature pour l'être humain et pour une grande classe d'animaux, d'être constamment observé par autrui – et notamment par des individus inconnus. Comme le fait remarquer la philosophe anglaise Mary Midgley :

Être constamment observé produit de l'horreur non seulement chez l'homme, mais chez une grande variété d'espèces animales. Pour la plupart des créatures sociales, un regard direct constitue une menace ouverte.³¹⁶

Ce constat permet, selon l'auteure, de tirer la conclusion qu'il faut toujours accorder à l'individu humain une sphère d'intimité, à l'intérieur de laquelle il est protégé du regard extérieur. Il faut garantir à un être humain la possibilité de n'être pas observé. Une telle thèse pourrait permettre de formuler, par exemple, des critiques contre les politiques d'accueil des migrants, lorsque ces politiques préconisent de réunir – pour ne pas dire entasser – les migrants dans le même espace – sans sphère d'intimité – pour de longues durées. Elle pourrait permettre aussi de critiquer des doctrines politiques préconisant une surveillance constante des comportements individuels, notamment par le biais de vidéos de surveillance.

cependant, l'argument n'est plus désormais un simple argument de faisabilité, mais un argument moral. Il serait immoral d'attribuer aux citoyens une trop grande responsabilité morale.

³¹⁶ "Being stared at produces horror not only in man, but in a great range of animal species. In most social creatures, a direct stare constitute an open threat." (MIDGLEY, 1995, chap. 1).

D'autres exemples peuvent encore être proposés ; le besoin d'amour et d'honneur constituent également des besoins essentiels de l'être humain. En conséquence, il faut établir des institutions qui n'empêchent pas la satisfaction de ces besoins.

Lorsque nous demandons « comment modifier cette institution ? », nous pouvons répondre à cette question seulement à l'aide d'une autre question, « qu'est-ce que demande la nature humaine ? ». A la fois l'amour et l'honneur sont partout des éléments essentiels de la vie. Il n'existe pas de société où le fait que quelqu'un soit méprisé, ignoré et détesté ne constitue pas une explication claire de sa misère. Qu'il en soit ainsi dépend de notre constitution émotionnelle. Nous avons inévitablement besoin d'amour et d'honneur, et nous leur accordons de la valeur en eux-mêmes, et non pas seulement en tant que moyens pour atteindre un bien plus éloigné et plus simple.³¹⁷

Il faut, selon Midgley, adapter nos sociétés pour que les besoins *fondamentaux* de l'être humain, c'est-à-dire les besoins relevant de sa nature biologique, ne soient pas contredits, notamment en n'adoptant pas de normes qui requièrent de renoncer à la satisfaction de ces besoins. Un modèle de société fondé sur des normes qui ne tiennent pas compte de ces besoins fondamentaux est trop exigeant et doit donc être rejeté.

Une autre conception de la nature humaine souvent évoquée, par les philosophes, est celle de l'être humain comme ayant besoin d'avoir « une personnalité unifiée ». La néo-kantienne Christine Korsgaard, par exemple, considère que seule une vie menée en conformité avec l'impératif catégorique de Kant permettrait de satisfaire ce besoin humain de se construire une identité personnelle unifiée dans le temps (KORSGAARD, 1996). Le philosophe français René Sève considère également que l'observation empirique du besoin d'avoir une image cohérente de soi constitue le fondement de la morale.

Nous considérons donc qu'il faut comprendre l'éthique à partir du besoin de l'homme d'une représentation globale de son existence, représentation qui n'est pas forcément théorique et conceptuelle mais qui cependant rassemble,

³¹⁷ "When we ask, "how should this institution be altered?" we can approach an answer only by asking the further question, "what does human nature demand?" Both love and honor are essential elements in life everywhere. There is no society where it is not a clear explanation of someone's misery to say that he is universally despised, ignored, or hated. That this is so depends on our emotional constitution. We inevitably need and value love and honor in themselves, not just as means to any further and simpler good." (MIDGLEY, 1995, chap. 9.)

hiérarchise et rend cohérents un maximum possible d'éléments de vie passés, présents ou futurs et ce faisant, selon l'expression consacrée, leur donne un sens (1993, p. 106).³¹⁸

La même référence au besoin d'avoir une image cohérente de soi se retrouve encore dans les textes de Mary Midgley.

Les personnes ont un souhait naturel et une capacité d'avoir une personnalité intégrée, et ont en horreur le fait d'être complètement fragmentés.³¹⁹

Selon les auteurs susmentionnés, étant donné que l'être humain cherche à se construire une identité cohérente, toute doctrine morale doit lui permettre de satisfaire ce besoin – ou, du moins, ne doit pas exiger la frustration de ce besoin. Une théorie normative qui serait incompatible avec la satisfaction de ce besoin devrait être rejetée – ce qui inclut, notamment, la doctrine de l'égoïsme ou de l'amoralisme.

Quelles sont les conséquences normatives de l'existence de besoins fondamentaux humains biologiquement déterminés, combinée à la « contrainte motivationnelle » susmentionnée ? Le fait que l'être humain a des besoins fondamentaux, et qu'il est (biologiquement) impossible pour un être humain de ne pas désirer les satisfaire, a pour conséquence qu'il ne peut pas être exigé moralement d'un être humain de renoncer à la satisfaction de l'intégralité de ses besoins fondamentaux. Ainsi, par exemple, une doctrine morale telle que le stoïcisme doit être rejetée.

Si nous acceptons l'ensemble des thèses descriptives que nous venons de voir, au sujet de la nature humaine, et que nous acceptons, par ailleurs, que ces thèses permettent de formuler des *limites* aux exigences morales ou sociales imposées aux êtres humains, alors nous obtenons le résultat suivant : une société humaine est une société dans laquelle les membres possèdent une sphère d'intimité (aversion de l'observation par autrui), dans laquelle les règles ne changent pas continuellement (désir de continuité), dans laquelle les individus ne se voient pas attribuer une responsabilité trop grande (atteinte à la santé

³¹⁸ Selon René Sève, la satisfaction du besoin de posséder une image cohérente de soi donne lieu ultimement à un sentiment de « dépassement de soi » qui trouve son expression dans la religion, les sciences et l'art (SÈVE, 1993).

³¹⁹ "People have a natural wish and capacity to integrate themselves, a natural horror of being totally fragmented [...]" (MIDGLEY, 1995, chap. 9.).

psychique), une société qui ne rend pas impossible l'amour sincère et dans laquelle l'honneur des personnes est préservé. Il semble donc évident, par ces exemples, que la description de la nature humaine permet de formuler des *limites* à ce qui peut être exigé d'un individu et au type de société qui est *réalisable*, c'est-à-dire qui ne suscitera pas de ressentiment et d'insatisfaction, sources d'instabilité.

Nous allons tenter, dans les prochaines sous-sections, de répondre à cette critique de l'argument de la vacuité. Nous allons y de la manière suivante : d'abord (5.6.1), si la critique est correcte, elle implique que les êtres humains qui *ont* le désir en question ne sont pas de véritables êtres humains – sur le plan biologique ou psychologique. Or, il est possible de considérer que cette conséquence est absurde. En second lieu (5.6.2), les arguments avancés ne concernent pas, selon nous, ce que les êtres humains *ne peuvent pas* désirer ou valoriser, mais ce que les êtres humains *devraient* désirer ou valoriser, étant donné leur constitution. Autrement dit, la critique porte sur ce qui *est bon pour* les êtres humains, et non sur ce que les êtres humains *ne peuvent pas désirer* ou vouloir. Cependant, si tel est le cas, nous retombons dans la discussion au sujet de ce qui est *bon pour* une personne. Or, nous avons vu dans les précédentes sections (5.5) que le principe DIP-motivation ne permet pas d'exclure certaines normes par le biais d'une référence à ce qui est *bon pour* une personne. Pour finir, nous montrerons que les partisans de la contrainte humaniste devraient accepter que *n'importe quelle* motivation humaine, pourvu qu'elle soit essentielle, pourrait rendre fausse une théorie morale. Cela impliquerait d'accepter même les motivations les plus terribles et outrageuses. Or, personne n'accepte que les déficiences motivationnelles qui sont le résultat d'une déficience morale puissent servir à *justifier* un comportement (5.6.3) ou même que de telles déficiences motivationnelles puissent servir à *excuser* un comportement (5.6.4). En conclusion, étant donné que, même s'il existait des limites motivationnelles, seules celles qui sont *moralement acceptables* pourraient servir à rejeter une norme, cela montre que la morale est prioritaire par rapport aux considérations motivationnelles. Il n'existe donc pas de limites motivationnelles à ce qui peut être exigé, d'un point de vue moral, des êtres humains.

5.6.1. DES INDIVIDUS NON HUMAINS

Pour débiter, il nous faut d'abord apporter quelques précisions. Comme nous l'avons déjà mentionné, l'incapacité de se motiver ne doit pas être une incapacité individuelle. En

effet, si un individu est incapable de se motiver à se conformer à son devoir – par simple faiblesse de la volonté – nous avons admis que cela ne supprime pas le devoir. L'exemple suivant, proposé par David Estlund, illustre notre intuition à cet égard:

Supposons que Bill soutient qu'il n'a pas le devoir de se retenir de déverser ses déchets (*dumping*) parce qu'il est incapable de se motiver à le faire. Il ne souffre d'aucune sorte de phobie spéciale, de compulsion ou de maladie. Il est simplement profondément égoïste et, de ce fait, est incapable de complètement désirer obéir. Déverser ses déchets au bord de la route est plus simple que de les emballer proprement et de les mettre dans une corbeille ou de les amener à la déchetterie. Il souhaiterait avoir plus de force de volonté, mais il ne la possède pas. Se retenir de jeter ses déchets est quelque chose que, sous tous les autres aspects, il lui est facilement possible de faire. Il ne parvient simplement pas à se motiver à le faire. Il serait absurde pour Bill de suggérer que son incapacité supprime son devoir. Cette incapacité motivationnelle n'a manifestement pas le pouvoir de bloquer une exigence dans une situation individuelle.³²⁰

Si nous nous déplaçons maintenant de l'incapacité motivationnelle *individuelle* à l'incapacité motivationnelle *typique* d'un être humain, est-ce que cela permet d'imposer des limites au contenu des exigences morales? En suivant Estlund, nous pensons que le caractère *typique*, ou *statistiquement répandu*, d'une motivation humaine ne suffit pas non plus pour faire disparaître un devoir moral. En effet, si nous refusons qu'une incapacité individuelle supprime le devoir dans *un cas* particulier, alors nous devons refuser que cette incapacité supprime le devoir dans *tous les cas* particuliers.

Supposons que des personnes font la queue pour obtenir une évaluation morale de leur comportement. Il est dit à Bill que son égoïsme est effectivement une incapacité motivationnelle, mais qu'elle ne l'exempte pas de son devoir d'être moins égoïste. A la suite de Bill arrive Nina, avec la même demande [i.e. être exemptée de son devoir à cause de son égoïsme insurmontable]. A nouveau, sa

³²⁰ "Suppose Bill pleads that he is not required to refrain from dumping because he is motivationally unable to bring himself to do it. There is no special phobia, compulsion, or illness involved. He is simply deeply selfish and so cannot thoroughly will to comply. Dumping his trash by the road is easier than wrapping it properly and putting it by the curb or taking it to the dump. He wishes he had more willpower, and yet he doesn't have it. Refraining is something he could, in all other respects, easily do, except that he can't thoroughly will to do it. It would be silly for Bill to propose this as requirement-blocking. This motivational incapacity is patently powerless to block the requirement in the individual case." (ESTLUND, 2011, p. 220).

demande est rejetée, pour les mêmes raisons que pour Bill. A la suite de Nina arrive alors Kim, puis ainsi de suite. Puisque chaque cas est identique, notre jugement est le même. La queue peut bien contenir tous les êtres humains, mais ce fait n'ajoute rien aux cas individuels.³²¹

Pour servir de critère de *réfutation* d'un principe moral, il faudrait que l'incapacité humaine ne reflète pas seulement une donnée statistique, mais qu'elle constitue un *trait essentiel* de la nature humaine. Or, si une incapacité représente un *trait essentiel* de la nature humaine, cela signifie qu'une personne qui ne possède pas cette incapacité n'appartient pas pleinement à l'espèce humaine.

Admettons, par exemple, que l'être humain est *essentiellement* incapable de ne pas se soucier de son propre bien, il est incapable de ne pas être un minimum égoïste. Dans ce cas, le fait que Mère Theresa est capable d'un désintéressement *total*, *c'est-à-dire qu'elle est capable de ne pas se soucier de son propre bien*, a pour conséquence que Mère Theresa ne fait pas pleinement partie de l'espèce humaine. Elle a quelque chose de *surhumain*. De la même manière, si nous considérons qu'un être humain est *essentiellement* incapable de désirer la mort de ses propres enfants, alors nous considérons qu'une personne capable de se motiver à tuer son propre enfant, tel que Abraham, n'est pas vraiment un être humain – sous cet angle. Or, selon Estlund, de tels jugements sont absurdes ; Mère Theresa et Abraham sont bel et bien des êtres humains.

Qu'est-ce que cela signifierait si l'égoïsme n'était pas seulement typique mais *essentiellement* typique des êtres humains ? Une possibilité serait qu'un agent n'est pas vraiment un humain si ses motivations sont différentes. Il existe des exemples historiques de personnes qui ont été extrêmement non égoïstes, telle que Mère Theresa. L'objection que nous considérons est la suivante : une théorie qui exigerait des personnes d'être aussi peu égoïste que Mère Theresa est fausse,

³²¹ "Suppose people line up to get your moral opinion on their behavior. Bill is told that his selfishness is indeed a motivational incapacity, but that it does not exempt him from the requirement to be less selfish. But behind Bill comes Nina with the same query. Again, we dispatch her, on the same grounds as Bill. Behind Nina is Kim, and so on. Since each poses the same case, our judgment is the same. The line might contain all humans, but that fact adds nothing to any individual's case." (ESTLUND, 2011, pp. 220-221).

parce qu'un individu aussi peu égoïste n'est pas vraiment un être humain. Mais cela est absurde.³²²

Du point de vue qui nous intéresse, à savoir du point de vue biologique, il est en effet absurde de considérer Mère Theresa et Abraham comme des individus non humains.

5.6.2. CE QUE LES ÊTRES HUMAINS DEVRAIENT DÉSIRER, MAIS NE DÉSIRENT PAS NÉCESSAIREMENT

Les références à des désirs que les êtres humains *ne peuvent pas s'empêcher d'avoir*, en vertu de leur nature, pose un problème encore plus essentiel. En effet, chacun des exemples proposés porte sur un désir qu'un être humain *devrait* avoir, étant donné sa nature biologique, plutôt que sur un désir qu'un être humain *possède nécessairement*, étant donné sa nature biologique. En effet, à chaque fois qu'un désir particulier est décrit comme étant *essentiel* à l'être humain, cela ne signifie pas que tout être humain *possède* le désir en question – et qu'une créature qui n'aurait pas ce désir *ne compterait pas* comme un être humain – mais cela indique seulement que tout être humain *devrait* avoir le désir en question, que cela serait *bon pour lui*, en tant qu'il est un être humain, d'avoir un tel désir. Or, nous l'avons déjà dit, il est tout à fait possible, pour un être humain, de ne pas désirer ce qu'il sait, pourtant, être bon pour lui.

Selon nous, toutes les références à ce qu'un être humain *désire nécessairement* sont en fait des références à ce qu'un être humain *devrait désirer*, à ce qui est bon pour lui. Par exemple, rien dans la nature biologique de l'homme n'exclut qu'il puisse désirer vivre en isolation totale, qu'il puisse désirer tuer ses propres enfants, ou qu'il puisse désirer vivre une vie fragmentée. Le problème est plutôt que, si notre individu satisfait de tels désirs, il risque bien d'éprouver de l'insatisfaction, des frustrations et d'être malheureux. Il risque peut-être même, dans le cas de l'isolation, de perdre l'esprit, de devenir fou. De tout cela, cependant, nous pouvons seulement conclure que les désirs de notre homme *ne sont pas bons pour lui*, mais nous ne pouvons pas conclure que notre homme *est incapable* d'avoir de tels désirs en vertu de sa nature humaine.

³²² "What would it mean for selfishness to be not only typical but essentially typical of humans? One possibility is that an agent does not count as human if her motivations are otherwise. There are historical examples of people who were extremely unselfish, such as Mother Theresa. The objection we are considering is this: a theory that requires people to be as unselfish as Mother Theresa is false because anyone who is that unselfish does not count as human. But this is absurd." (ESTLUND, 2011, p. 221).

De la même manière, Mary Midgley soutient, par exemple, que le parricide est contre-nature pour l'être humain parce que nous sommes des créatures constituées de manière à être incapables de désirer tuer notre propre père ; nous formons dès l'enfance des liens affectifs très forts avec nos parents et ces liens se perpétuent à l'âge adulte, de sorte que nous sommes, selon elle, essentiellement incapables de *désirer* la mort de nos propres parents. Pourtant, du fait que nous sommes le genre de créatures qui forme certains liens affectifs avec ses parents dès l'enfance, il ne s'ensuit pas que nous sommes *incapables* de désirer commettre un parricide. Il s'ensuit uniquement que, étant donnée notre constitution affective, le parricide *n'est pas bon pour nous*, en tant qu'être humain. Un parricide susciterait chez un être humain une immense souffrance ou un désordre psychologique néfaste. Le parricide est ainsi *mauvais* pour un être humain.

Or, nous l'avons vu, le seul lien direct qu'entretient la notion de « bien pour » ou de « mauvais pour » avec la motivation est un lien *normatif*, et le seul lien *indirect* que la notion de « bien pour » entretient avec la motivation est celle de « pouvoir être motivé ». Il n'existe aucun lien direct entre ce qui est bon/mauvais pour une personne et ce qu'elle serait *incapable* de désirer.

Il est donc *approprié* à la nature affective d'un être humain de ne pas désirer tuer son propre père. Toutefois, comme nous l'avons déjà dit, la reconnaissance de ce fait implique uniquement qu'un être humain *ne devrait pas* (rationnellement) désirer tuer notre propre père. Et ce fait lui-même implique, certes, qu'un être humain est capable de ne pas désirer tuer son propre père ; mais cela n'implique *pas* qu'il est biologiquement impossible pour l'être humain de désirer tuer son propre père.

Prenons un exemple positif, plus facile à formuler que l'exemple *néгатif* du parricide. Selon Mary Midgley toujours, il est approprié à la nature humaine de vivre en communauté. Que peut-on en conclure ? Le seul type de raisonnement que ce constat nous autorise à construire est le suivant:

(P1)	Il est bon pour une créature comme l'être humain de vivre en communauté	Prémisse normative non morale.
(P2)	Tout être humain doit (rationnellement) désirer ce qui est bon pour lui	Prémisse normative non morale (rationalité)
(C1)	Un être humain doit (rationnellement) désirer vivre en communauté	Conclusion normative
(P3)	Si un être humain doit faire X, alors il peut faire X	DIP
(C2)	Un être humain peut désirer vivre en communauté	Conclusion descriptive

Ce raisonnement nous autorise uniquement à conclure que l'être humain *a la capacité* de désirer vivre en communauté. Or, ce que cherche à obtenir notre critique est une conclusion plus forte ; elle cherche à prouver qu'*un être humain est incapable de désirer vivre en dehors* d'une communauté – ou incapable de ne pas désirer vivre en communauté. Cependant, cette conclusion ne lui est pas accessible. Elle peut seulement obtenir que, du fait qu'il est *bon* pour une créature humaine de vivre en communauté, alors elle est *capable* de désirer vivre en communauté. Elle ne peut pas obtenir la conclusion selon laquelle l'être humain est *incapable* de désirer vivre en dehors de toute communauté.

Si notre critique souhaite rétorquer que l'être humain est incapable de *ne jamais désirer ce qui est bon pour lui*, ou incapable de désirer vivre une vie entière qui ne contient rien de ce qu'il croit être bon pour lui, alors nous nous retrouvons dans la discussion de la section précédente (5.5). Et nous avons déjà répondu à cette critique.

5.6.3. LA VALEUR MORALE DES LIMITES MOTIVATIONNELLES POUR LA JUSTIFICATION

Les théories qui font appel à la *nature humaine*, ou aux traits essentiels de la nature humaine, soit dans le but d'élaborer un système moral positif, soit dans le but de critiquer un système moral concurrent, rencontrent encore un autre problème ; elles sont contraintes d'accepter, comme contrainte motivationnelle sur les normes morales, des traits de la nature humaine moralement problématiques.

L'approche de ces théories est ainsi la suivante : elles ont pour objectif de développer une description *moralement neutre* de la nature humaine et de ses traits essentiels. Elles

prévoient ensuite d'utiliser cette description pour rejeter les théories morales qui n'en tiennent pas compte dans la formulation des principes moraux. Le problème, si nous suivons la voie proposée, est que *n'importe quel trait essentiel* à la nature humaine, indépendamment de sa valeur axiologique ou morale, représente une raison de rejeter une théorie morale qui n'en tient pas compte. La difficulté provient du fait que, si nous acceptons ce raisonnement, alors il faudrait admettre que même des motivations moralement problématiques, mais constitutives de la nature humaine, permettent de séparer les bonnes et les mauvaises théories morales.

En effet, si l'argument pour des devoirs moins exigeants devait procéder de l'observation d'incapacités motivationnelles des êtres humains, il s'agirait d'un argument profondément troublant, puisqu'il aurait la forme suivante : la tendance à la partialité est caractéristique des personnes, puisqu'elle fait partie de la nature humaine. *En conséquence*, les exigences de se comporter différemment sont fallacieuses et fausses. Si l'argument précédent est un bon argument, alors l'argument suivant est également un bon argument : les personnes présentent une certaine tendance à la cruauté, et cette tendance fait partie de la constitution motivationnelle de la nature humaine. (On suppose qu'il en est ainsi). En conséquence, les exigences d'agir différemment [i.e. de ne pas être cruel] sont fallacieuses et fausses.³²³

Bien entendu, personne n'admet cette conclusion ; personne n'admet que la cruauté est permise parce que la cruauté est un trait essentiel de la nature humaine. Personne n'admet que, pour tenir compte de l'envie des moins bien lotis d'une société, les envieux ont le droit de prendre la richesse légitimement acquise des mieux lotis. Personne n'admet que, pour tenir compte du désir naturel de violence de l'homme, celui-ci a le droit de se défouler sur d'autres personnes, ou que l'État a le devoir moral d'organiser des guerres permettant l'expression de ce désir profond. Il est donc faux que tout trait essentiel de la nature

³²³ "Indeed, if the argument for a less demanding duty were to proceed from the observed motivational incapacities of humans, it would be a deeply troubling argument, having the following form: This tendency to partiality is what people are motivationally like, as a matter of human nature. Therefore, requirements to be otherwise are specious and false. If that's a good form of argument, then the following is a good argument: People tend to a certain degree of cruelty, and this is part of what they are motivationally like as a matter of human nature. (Suppose this is so.) Therefore, requirements to be otherwise are specious and false. That is an absurd argument. The form of argument is bad, and so no prerogative in favor of partiality can be inferred from the fact, if it is one, that humans are naturally partial or selfish." (ESTLUND, 2011, pp. 223-224).

humaine impose des limites aux normes morales et politiques. Seuls les traits essentiels *moralement acceptables* peuvent servir à formuler des limites aux exigences morales.

Bernard Williams formule une critique dans ce sens. Il reproche aux théories morales fondées sur une prétendue *nature humaine*, de choisir quels sont les traits essentiels de l'être humain en fonction d'une doctrine morale *préalable*.³²⁴ Selon lui, en effet, toute tentative d'identifier une *marque distinctive* ou des *traits essentiels* de l'être humain, dans le but de formuler des normes morales, est voué à rencontrer au moins deux difficultés. En premier lieu, la sélection dépendra certainement d'une évaluation morale préalable.

Il existe des objections générales contre la procédure consistant à vouloir identifier des objectifs et idéaux moraux sur la base de la marque distinctive de la nature humaine. [...] En premier lieu, un degré considérable d'évaluation a déjà joué un rôle dans la sélection de cette marque distinctive, comme c'est le cas pour la rationalité ou la créativité. Si quelqu'un approchait cette thématique sans aucun préjugé le problème consistant à découvrir la marque distinctive de l'homme par rapport aux autres animaux, il pourrait tout autant [...] parvenir à une morale qui exige des hommes de passer autant de temps que possible à faire du feu ; à développer des caractéristiques physiques spécifiques à l'être humain ; ou à avoir des rapports sexuels indépendamment de la saison ; ou de détruire l'environnement et de renverser l'équilibre de la nature ; ou de tuer pour le plaisir.³²⁵

En second lieu, même s'il était possible d'identifier un trait distinctif ou essentiel de la nature humaine, ce trait ne caractérise pas nécessairement une personne vertueuse ; il peut constituer une caractéristique louable chez une personne, et une caractéristique moralement détestable chez une autre.

³²⁴ L'auteur discute des philosophes qui cherchent à fonder une théorie morale sur les traits distinctifs de la nature humaine, mais son argument s'appliquerait tout aussi bien aux philosophes qui cherchent à fonder leur théorie morale sur les traits *essentiels* de la nature humaine.

³²⁵ "[...] First, a palpable degree of evaluation has already gone into the selection of the distinguishing mark which is given this role, such as rationality or creativity. If one approached without preconceptions the question of finding characteristics which differentiate men from other animals, one could as well, on these principles, end up with a morality which exhorted men to spend as much time as possible in making fire; or developing peculiarly human physical characteristics; or having sexual intercourse without regard to season; or despoiling the environment and upsetting the balance of nature; of killing things for fun." (WILLIAMS, 1972, p. 59).

En second lieu, cette approche confirme le caractère moralement *ambigu* des caractéristiques distinctives de l'être humain. [...] Puisque s'il est distinctif de l'homme d'employer son intelligence et des outils pour modifier son environnement, c'est aussi une marque distinctive de l'homme d'employer son intelligence pour n'en faire qu'à sa tête et des outils pour détruire d'autres personnes. S'il s'agit bien d'une marque distinctive de l'être humain d'avoir parfaitement conscience de lui-même en tant qu'individu parmi d'autres, d'avoir conscience que les autres ont des sentiments comme lui, c'est une précondition non seulement de la bienveillance mais aussi (comme Nietzsche l'a montré) de la cruauté : l'homme sadique sophistiqué ne ressemble pas plus aux autres animaux que l'homme ayant des affections naturelles pour autrui. Au contraire, il leur ressemble moins. Si nous considérons comme impératif moral suprême ce vieux cri, « sois un homme ! », il est terrifiant de s'imaginer les différentes manières dont il peut être lu de façon littérale.³²⁶

Dans les théories morales fondées sur la notion de *nature* humaine, les traits considérés comme *essentiels* sont précisément ceux qui sont compatibles avec la théorie morale, et ceux qui sont considérés comme *inessentiels* ou même *inhumains* sont ceux qui ne sont pas compatibles avec la théorie morale défendue. En d'autres termes, c'est la théorie morale (censée former la conclusion du raisonnement) qui détermine quels sont les traits essentiels de la nature humaine. Il est donc faux que nous pouvons décrire de façon moralement neutre les traits essentiels de la nature humaine, pour ensuite les intégrer dans notre meilleure théorie morale. Nous devons d'abord savoir quels sont les traits de la nature humaine qui sont moralement acceptables.

En conclusion, le seul moyen de faire converger, d'une façon acceptable, une description de la nature humaine avec une théorie morale, c'est d'adapter la description de la nature humaine à la théorie morale – et non l'inverse. Or, une telle démarche consiste

³²⁶ "There are more general objections to the procedure of trying to elicit unquestionable moral ends or ideals from the distinguishing marks of man's nature. Second, and very basically, this approach bears out the moral ambiguity of distinctive human characteristics [...] For if it is a mark of man to employ intelligence and tools in modifying his environment, it is equally a mark of him to employ intelligence and tools in getting his own way and tools in destroying others. If it is a mark of man to have conceptualized and fully conscious awareness of himself as one among others, aware that others have feeling like himself, this is a precondition not only of benevolence but (as Nietzsche pointed out) of cruelty as well: the man of sadistic sophistication is not more like other animals than the man of natural affections, but less so. If we offer as the supreme moral imperative that old cry, "be a man!", it is terrible to think of many of the ways in which it could be taken literally." (WILLIAMS, 1972, pp. 59-60).

précisément à commettre le sophisme moraliste dont nous traiterons dans le chapitre suivant. Avant de nous y intéresser, il nous faut encore répondre à une dernière critique potentielle.

5.6.4. LA LÉGITIMITÉ DU BLÂME

Nous avons affirmé que le caractère *essentiel* d'une motivation humaine ne permet pas de justifier le rejet d'une norme ou théorie morale qui s'opposerait à cette motivation. Autrement dit, un énoncé descriptif au sujet du caractère essentiel d'une motivation humaine ne *rend pas fausse* une théorie morale dont les principes contrediraient cette motivation. Nous pouvons toutefois nous demander : même si nous acceptons que les vérités au sujet des motivations humaines ne peuvent pas *falsifier* une théorie morale, ne peuvent-elles pas quand même rendre le blâme *inadéquat*? Si une personne est incapable de se motiver à faire quelque chose – en vertu d'une incapacité essentielle à la nature humaine – n'est-il pas inapproprié de lui reprocher son manquement à son devoir moral ? Par exemple, si un être humain est essentiellement incapable (pour des raisons biologiques) d'attribuer la même valeur aux enfants des autres personnes qu'aux siens, n'est-il pas inapproprié de le blâmer parce qu'il ne parvient pas à attribuer la même valeur aux enfants des autres personnes? Cela même si cette incapacité ne *supprime pas* le devoir en question ?

A des fins de clarté du propos, il nous faut commencer par introduire une distinction devenue classique dans les discussions philosophiques, à savoir la distinction entre, d'un côté, la justification et, de l'autre, l'excuse (AUSTIN, 1956 - 1957). Les circonstances qui appellent la justification ou l'excuse sont identiques : une personne est accusée d'avoir mal agi et cherche à se défendre. Cependant, la défense ne procède pas dans les deux cas de la même manière. Dans le cas de la justification, l'individu admet qu'il a fait l'action dont on l'accuse, mais il conteste que cette action était mauvaise. Dans le cas de l'*excuse*, à l'inverse, il admet que l'action est mauvaise, mais il *conteste* être (pleinement) responsable de l'action. Pour résumer :

Dans l'une des deux défenses, brièvement, nous acceptons la responsabilité mais nions que l'action était mauvaise: dans l'autre, nous admettons que l'action était

mauvaise, mais nous refusons la responsabilité totale, voire toute la responsabilité.³²⁷

Dans la thématique qui nous intéresse, si quelqu'un soutient que l'incapacité motivationnelle *fait disparaître le devoir moral*, alors il s'agit d'un cas de *justification* ; le fait que l'action contredit des motivations humaines essentielles rend *faux* l'énoncé normatif selon lequel l'action est *obligatoire*. Par contre, si quelqu'un soutient que l'incapacité motivationnelle ne modifie pas les caractéristiques morales de l'action, mais modifie l'attribution de responsabilité, alors il utilise une défense par *l'excuse*.

Nous avons vu que l'incapacité motivationnelle ne peut pas servir de *justification* ; l'incapacité motivationnelle, indépendamment de tout jugement évaluatif concernant son caractère moralement approprié ou non, ne suffit pas pour falsifier un énoncé moral déontique. Mais qu'en est-il pour *l'excuse* ? Est-ce que l'incapacité motivationnelle (humaine) supprime la responsabilité ?

Selon Estlund, lorsqu'il s'agit de déterminer la *responsabilité* – et non pas le contenu du devoir moral – l'incapacité motivationnelle est pertinente, même dans les cas où cette incapacité est *purement individuelle* – et ne constitue pas une incapacité de l'espèce humaine. Imaginons, par exemple, un touriste qui, à la suite d'un manque d'attention, est sur le point de tomber du pont de glace du Grand Canyon. A l'entrée du pont se trouve un serveur du restaurant adjacent. Notre serveur a la capacité physique de venir en aide au touriste en mauvaise posture. Le problème provient du fait que le pont (en forme de fer à cheval) que le serveur doit traverser est un pont de verre transparent, suspendu à 400 m au-dessus du vide. Notre serveur, malgré sa volonté de venir en aide au montagnard, est incapable de se motiver à traverser ledit pont, à cause de sa peur du vide. Dans un tel cas, le serveur a toujours le devoir moral de venir en aide au touriste – et donc de traverser le pont. Cependant, le serveur n'est pas *blâmable* s'il ne réussit pas à accomplir son devoir à cause du vertige qui le paralyse. Il n'est pas *responsable* de son incapacité.

Si nous considérons que le serveur n'est pas *blâmable* dans le contexte du pont de glace suspendu, alors nous pouvons admettre que les incapacités motivationnelles

³²⁷ "In the one defence, briefly, we accept responsibility but deny that it was bad: in the other, we admit that it was bad but don't accept full, or even any, responsibility" (AUSTIN, 1956 - 1957, p. 2).

inhérentes à la nature humaine constituent, elles aussi, des circonstances qui suppriment (au moins partiellement) la responsabilité.

Il est particulièrement clair que la peur de la hauteur réduit dans une certaine mesure la susceptibilité de l'agent au blâme, ce qui signifie que des réductions similaires de responsabilité peuvent exister dans les contextes d'injustice résultant de la nature humaine.³²⁸

Selon le philosophe, cependant, le fait que l'incapacité motivationnelle constitue un trait individuel ou un trait répandu chez les êtres humains ne fait pas de différence concernant l'évaluation de la responsabilité – i.e. concernant la validité de l'excuse. Si le seul fait qu'une incapacité motivationnelle est un trait de la nature humaine suffisait pour légitimer l'excuse, alors n'importe quelle caractéristique répandue constituerait une raison légitime pour une excuse – même les cas de motivations déraisonnables ou immorales. Or, nous ne considérons pas que n'importe quelle incapacité motivationnelle – pour autant qu'elle fasse partie de la nature humaine – constitue une excuse légitime pour ne pas accomplir son devoir moral.

Si une certaine peur de la hauteur excuse le fait qu'une personne désobéisse à son devoir de marcher sur le pont de glace, il en est ainsi indépendamment du fait qu'une autre personne (ou même que tous les êtres humains) souffre de la même peur. La nature humaine, dans le sens discuté dans cet article, n'est pas pertinente. De la même manière que ce que j'ai soutenu dans le cas de la justification, le fait de laisser les excuses être déterminées par la nature humaine rendrait légitime le soutien à des principes odieux qui excuseraient la cruauté ou la bigoterie, du moment qu'elles constituent des traits caractéristiques ou naturels de l'être humain.³²⁹

³²⁸ "What is relatively clear is that the fear of heights mitigates the agent's liability to blame to some extent, which means that there might be similar mitigation in contexts of injustice due to human nature." (ESTLUND, 2011, p. 231).

³²⁹ "If a certain fear of heights excuses the failure to step onto the glass bridge as required, this is so whether or not anyone else (or all other humans) suffer from the same fear. Human nature, in the sense under discussion in this article, is not germane. Just as I argued in the case of justification, letting excuses be set by human nature would rely on an abhorrent principle that would excuse cruelty or bigotry if only they are characteristic or natural." (ESTLUND, 2011, p. 232).

Les incapacités motivationnelles qui sont susceptibles d'offrir une excuse pour ne pas avoir accompli un devoir doivent donc avoir des caractéristiques spécifiques – et le fait d'être une caractéristique essentielle de la nature humaine ne fait pas partie de ces caractéristiques. En premier lieu, il faut que l'incapacité motivationnelle soit *raisonnable*. En effet, si une personne est incapable de se motiver, par peur du vide, à descendre un trottoir de 10 cm de hauteur pour aider un cycliste qui a été touché par une voiture sur la route, il serait difficile de considérer qu'il s'agit là d'une bonne excuse. Il n'est pas raisonnable d'avoir peur du vide sur un trottoir, alors qu'il semble raisonnable d'avoir peur du vide à 400m de hauteur.³³⁰

Une autre condition doit encore être satisfaite pour que *l'excuse* soit appropriée, c'est-à-dire pour que la responsabilité de la personne soit levée. Il faut que l'incapacité motivationnelle soit *moralement acceptable*, que la personne soit *moralement bien disposée*. Ainsi, lorsque le serveur se trouve dans l'incapacité de se motiver à sauver le touriste suspendu au pont de glace du Grand Canyon, il reste vrai que, si le serveur n'avait pas peur du vide, alors il serait motivé à sauver le touriste en danger – et serait blâmable s'il ne le faisait pas. L'échec de motivation ne provient pas d'une incapacité à saisir les raisons *morales* de sauver le touriste. Imaginons, par contre, que le serveur, s'il n'avait pas eu peur du vide, parviendrait à se motiver à sauver le touriste uniquement si celui-ci est un homme blanc. Si le touriste est noir, le serveur, à cause d'une éducation raciste, est incapable de se motiver à lui venir en aide. Dans un tel cas, l'incapacité motivationnelle, symptôme d'une incapacité morale, ne constitue pas une excuse appropriée pour lever la responsabilité.

Supposons que nous ne parvenons pas à nous motiver à sauver une personne innocente, mais que nous parvenons à nous motiver à sauver n'importe quel chat.
Ou que nous ne parvenons à nous motiver à traverser la passerelle, mais uniquement si c'est pour sauver quelqu'un de la même race que la nôtre. Ces

³³⁰ Quoique, selon nous, le serveur n'ayant pas de raison de craindre le vide, n'a pas non véritablement d'excuse valable de ne pas surmonter sa peur. Par contre, si sa peur était due au fait qu'il doit traverser un pont scabreux, en mauvais état, fragile, etc. pour sauver une autre personne, alors sa peur est appropriée et son incapacité à se motiver raisonnable.

deux situations indiquent des motivations moralement déficientes, alors que la peur plus générale de la hauteur n'indique pas une déficience morale.³³¹

A nouveau, si nous admettons que, dans le cas de la passerelle, seules les incapacités motivationnelles *qui ne sont pas moralement problématiques*, constituent des excuses légitimes, alors nous devons admettre que, dans le cas des incapacités à faire son devoir social ou moral, seules les incapacités motivationnelles qui ne sont pas moralement problématiques – i.e. qui ne sont pas le symptôme d'une déficience morale – constituent des excuses légitimes. Si l'incapacité motivationnelle des individus, dans une société donnée, est une incapacité moralement infondée, alors elle ne constitue pas une condition qui rend le blâme illégitime.

Si les personnes ne peuvent pas se motiver à agir de la manière dont l'exige la justice sociale, il est possible que ces personnes soient motivées par des raisons de justice sociales mais excusées par une incapacité similaire à celle de la peur de la hauteur, mais il est aussi possible que cela ne soit pas le cas. Il est possible que leur incapacité motivationnelle soit due à une indifférence moralement inappropriée aux intérêts légitimes d'autrui.³³²

Par exemple, si les hommes sont (constitutivement) incapables de supporter la présence de femmes dans une réunion, est-ce que cela excuse leur refus d'intégrer des femmes dans les postes clés d'une entreprise ? Cela ne semble pas être le cas. Les hommes ont le devoir moral d'accepter la présence de femmes dans les postes de direction, même s'ils sont incapables de supporter leur présence.

Le propos d'Estlund est ainsi le suivant : il existe des types de motivations (ou d'incapacités motivationnelles) qui rendent impossible tout type de justice sociale, parce que ces incapacités motivationnelles reflètent une incapacité à saisir des raisons morales

³³¹ "Suppose that you cannot bring yourself to cross to save an innocent person, but you can do so to save any cat. Or you can bring yourself to cross but only to save someone of your own race. Either of these might indicate morally deficient motivations in a way that a more general fear of heights does not." (ESTLUND, 2011, p. 234)

³³² "If people can't bring themselves to do as social justice would require, then, while it might be that they are properly motivated by reasons of social justice but excused by something akin to a general fear of heights, it might instead be the case that they are not. It might be that their motivational incapacity is owed to a morally improper indifference to the legitimate interests of others." (ESTLUND, 2011, p. 233).

d'agir. Or, dans ce cas, ces incapacités motivationnelles n'offrent ni une justification, ni une excuse pour ne pas faire son devoir moral.

Même si certaines incapacités motivationnelles permettent d'excuser la violation d'une exigence morale, elles ne permettent pas de la justifier. En ce sens, les standards moraux restent intacts. Et rien de cela ne serait affecté par le fait que ces incapacités sont largement partagées ou naturelles. D'autres incapacités motivationnelles n'offrent même pas d'excuses, et encore moins de justification, puisqu'elles reflètent une structure moralement déficiente dans les intérêts. Et, de nouveau, même si elles étaient répandues et naturelles, elles n'offriraient pas d'excuses, et encore moins de justification.³³³

Le critère de sélection des incapacités motivationnelles susceptibles d'excuser un échec moral n'est donc pas lui-même moralement neutre. Ce qui prouve que les énoncés normatifs concernant l'absence de devoir moral, ou l'absence de responsabilité morale, ne dérivent pas d'énoncés *purement descriptifs*.

Mon argument a pour but de montrer non seulement que les faits statistiques au sujet de la manière dont les personnes agissent ou pourraient agir, mais aussi les références à des incapacités motivationnelles naturelles, ne sont pas décisifs. La question est de savoir si ces incapacités reflètent des formes de déficiences morales dans la volonté ou dans les intérêts.³³⁴

Ainsi, par exemple, si nous admettons que l'envie n'est *jamais* moralement appropriée – contrairement au *ressentiment* par exemple – alors l'envie ne permet pas d'excuser une personne qui s'est montrée incapable, par envie, de respecter les exigences de la justice sociale. Si une personne ne supporte pas le fait qu'une autre soit plus riche qu'elle – même si cet écart de richesse est moralement légitime – et qu'en conséquence elle porte atteinte au bien de la personne qu'elle envie – au détriment possible de son propre intérêt – son

³³³ "Even if some motivational incapacities are excusing, this does not make them justifying, and so there are moral standards that survive intact. None of this would be affected by how widespread or natural the incapacities might be. Other motivational incapacities are not even excusing, much less justifying, since they reflect a morally deficient structure of concerns. Again, they would not become excusing, much less justifying, simply by being widespread or even natural." (ESTLUND, 2011, pp. 234-235).

³³⁴ "My argument is meant to show not only that there would be no support in mere statistical facts about how people do or might act, but that even appeals to natural motivational incapacity would not be decisive. The question is whether or not those incapacities are, for all their naturalness, forms of morally deficient will or concern." (ESTLUND, 2011, p. 235).

incapacité motivationnelle à respecter la répartition (juste) des richesses ne constitue pas une condition rendant le blâme inapproprié. La personne *reste blâmable* pour ne pas avoir respecté le droit d'autrui à avoir plus de richesse, car son incapacité résulte d'une incapacité à percevoir les raisons morales de l'écart de richesse.

6. CONCLUSION

Nous avons commencé par proposer une défense de DIP fondée sur la sémantique de la notion de « devoir ». Nous avons vu que cette défense soulève plusieurs difficultés. L'interprétation de DIP comme exprimant une vérité conceptuelle aurait notamment pour implication, dit-on, que les dilemmes moraux et la faiblesse de la volonté sont impossibles. Nous avons alors tenté de déterminer s'il existait une autre interprétation de DIP, qui ne rencontre pas de problèmes fondamentaux. Nous avons toutefois vu que chacune des interprétations rencontre de graves difficultés. Nous sommes alors parvenus à la conclusion (faible) que l'interprétation conceptuelle de DIP n'occupe pas une position moins confortable que les autres interprétations.

Dans un second temps, nous avons montré que les empiristes ne peuvent pas faire un usage adéquat du principe DIP-motivation. En effet, nous avons soutenu que DIP-motivation souffre d'un problème de *vacuité* lorsqu'il est employé dans un raisonnement empiriste. Nous avons également vu que toutes les tentatives pour résoudre cette difficulté sont vouées à l'échec aussi longtemps que nous restons dans une perspective empiriste. De ce fait, nous concluons qu'il n'est pas possible de rejeter une théorie morale en l'accusant d'être trop exigeante sur le plan motivationnel. La contrainte humaniste doit être rejetée.

Nous allons voir, dans le prochain chapitre, que l'interprétation conceptuelle de DIP est finalement l'interprétation correcte du principe. Pour ce faire, nous répondrons aux objections à l'analyse conceptuelles que nous avons étudiées précédemment. Nous verrons également quel est l'usage adéquat du principe DIP-motivation ; nous soutiendrons que le recours à DIP-motivation n'est légitime que dans le cadre d'un raisonnement *moraliste*. En ce sens, l'usage approprié de DIP-motivation est sous sa forme *non contraposée* : « si tu dois, tu peux ! ».

QUATRIÈME PARTIE

Le raisonnement moraliste réhabilité

UNE DÉFENSE DU MORALISME

Le fait que l'homme soit conscient qu'il peut le faire parce qu'il le doit ouvre en lui un abîme de dispositions divines qui lui fait éprouver une sorte de frisson sacré face à la grandeur et la sublimité de sa véritable destination (KANT, 1986 (1793)).

1. LE SOPHISME MORALISTE

Les chapitres précédents nous ont amenés aux deux conclusions (négatives) suivantes :

- (A) Il est dangereux et utopiste de vouloir organiser la société en faisant l'hypothèse que les individus sont exclusivement motivés à agir par intérêt personnel.
- (B) Les insuffisances motivationnelles liées à la nature humaine ne permettent pas de fixer des *limites* à ce qui peut être exigé d'un individu du point de vue moral.

Dans le présent chapitre, nous souhaitons proposer l'ébauche d'un argument positif en faveur du raisonnement *moraliste*, c'est-à-dire un raisonnement consistant à dériver un énoncé descriptif à partir d'un énoncé normatif. Pour ce faire, nous allons commencer par caractériser la doctrine moraliste par rapport à d'autres doctrines. Dans un second temps, nous présenterons l'accusation de sophisme moraliste sous son aspect formel et général et présenterons quelques exemples de sophismes moralistes qui affectent le discours ordinaire et scientifique. Nous nous intéresserons ensuite plus spécifiquement au sophisme moraliste *motivationnel* et montrerons que plusieurs doctrines morales commettent ce sophisme. Nous tenterons alors de *réhabiliter* le raisonnement moraliste, en montrant que, lorsqu'il repose sur le principe analytique DIP, il ne constitue plus un véritable sophisme. Sur la base d'un tel raisonnement, nous développerons l'esquisse d'un argument *positif* en faveur de la possibilité d'une motivation morale *catégorique* et *dominante*, c'est-à-dire d'une motivation morale *indépendante* des désirs, et capables même de *surmonter* les désirs. Pour finir, nous répondrons à diverses objections ou craintes qui peuvent être soulevées contre notre démarche. En premier lieu, nous apporterons des réponses aux objections qui ont été soulevées contre l'interprétation conceptuelle de DIP, notamment (1.3.1) l'incompatibilité de DIP avec les dilemmes moraux insolubles, (1.3.2) l'incompatibilité de DIP avec la faiblesse

de la volonté et (1.3.3) l'incompatibilité de DIP avec la théorie morale de Luther. En second lieu, nous répondrons à deux objections qui peuvent être soulevées contre la variante plus spécifique du raisonnement moraliste que nous souhaitons développer, à savoir celle qui vise à soutenir l'existence d'une motivation morale *catégorique et dominante*. L'une des objections soutient (2) qu'il n'existe pas de devoirs moraux catégoriques – et donc qu'il n'existe pas non plus de motivation morale catégorique – et l'autre soutient (3) qu'une motivation sans désir est conceptuellement impossible. Nous parviendrons alors à la conclusion finale de notre travail, à savoir qu'il n'existe pas de raison de douter que, s'il existe des devoirs moraux catégoriques, alors les êtres humains sont capables d'une motivation morale catégorique correspondante. Nous terminerons (4) en clarifiant notre position concernant le rôle des considérations motivationnelles en philosophie morale et politique.

1.1. QUEL LIEN ENTRE LES ÉNONCÉS DESCRIPTIFS AU SUJET DES MOTIVATIONS HUMAINES ET LES ÉNONCÉS MORAUX ?

Jusqu'à présent, nous avons étudié deux manières de concevoir la relation entre, d'un côté, les énoncés descriptifs au sujet des motivations humaines et, de l'autre, les énoncés normatifs au sujet des devoirs moraux de l'être humain.

(A) L'approche économique

La première manière de concevoir la relation est celle adoptée par les économistes ; selon cette approche, les énoncés descriptifs (au sujet des motivations humaines) et les énoncés normatifs (au sujet du bien moral, de la justice et de la vertu) sont indépendants les uns des autres. En conformité avec la loi de Hume, les économistes admettent – pour des raisons épistémiques, principalement – l'existence d'un fossé entre la recherche normative et la recherche descriptive.³³⁵ C'est pour cette raison que, selon eux, le seul moyen d'obtenir

³³⁵ La thèse selon laquelle il existe un fossé infranchissable entre, d'un côté, la recherche descriptive et, de l'autre, la recherche normative, ne constitue que l'un des engagements « métaéthiques » de la science économique. En effet, les économistes, plus ou moins explicitement, et avec un degré variable de cohérence, défendent d'autres thèses métaéthiques. Par exemple, Jean Tirole semble soutenir, alternativement, que le voile d'ignorance de Rawls offre un bon critère pour partager le juste de l'injuste (constructivisme) (TIROLE, 2016, pp. 14, 16), que la vérité des énoncés moraux varie d'une communauté à l'autre (relativisme) (2006, p. 1666) et que les énoncés moraux expriment des attitudes subjectives des individus (subjectivisme) (TIROLE, 2016, p. 72).

un comportement moralement approprié consiste à mettre en place des incitations garantissant artificiellement une convergence entre les motivations humaines et le contenu des normes morales ; il n'existe pas de convergence *a priori* entre les motivations humaines et le contenu des normes morales.

(B) L'approche naturaliste

La seconde manière de concevoir la relation est celle adoptée par les métaéthiciens naturalistes : selon eux, il est possible de dériver des énoncés normatifs (moraux) sur la base d'énoncés descriptifs au sujet des motivations humaines. A l'aide du principe DIP, notamment, il est possible de dériver l'énoncé « les êtres humains n'ont pas le devoir de venir en aide aux pauvres » à partir des énoncés « les êtres humains poursuivent exclusivement leur intérêt personnel » et « aider autrui ne sert pas l'intérêt personnel ». Selon cette approche, la recherche en philosophie morale (normative) est, au moins partiellement, dérivée de la recherche empirique au sujet des motivations humaines.

De la même manière, dans le manuel de l'économiste suisse Philippe Paoly (2012), on peut voir se succéder en quelques pages des positions très différentes : (1) Il soutient explicitement le non-cognitivism, selon lequel les jugements de valeur sont irrationnels et non scientifiques (p. 42). (2) Il soutient implicitement une forme de réalisme des valeurs lorsqu'il affirme que l'amélioration du bien-être, qui est l'objet de la science économique, est une bonne chose (p. 42). (3) Il soutient également, explicitement, un relativisme moral, selon lequel les vérités morales varient selon les communautés (p. 44). (4) Pour finir, il soutient implicitement un réalisme au sujet des normes. En effet, selon Paoly, il est juste d'être récompensé pour l'innovation et la prise de risque ; « [L]e profit normal du marché [est] le profit en tant que juste rémunération de l'innovation et du risque, de l'esprit d'entreprise » (p. 97) .

Il n'existe à vrai dire que peu de travaux concernant les prises de position métaéthiques des économistes – à quelques exceptions, bien sûr (MONGIN, 2018; MONGIN, 1999; MANIQUET, 2018). Il est néanmoins possible de dégager les engagements métaéthiques suivants : les économistes considèrent en général que les énoncés normatifs ne possèdent pas de valeur de vérité. Cela a pour conséquence qu'il n'existe pas, selon eux, de critère objectif permettant de faire des comparaisons entre les préférences individuelles (i.e. les fameuses « comparaisons interpersonnelles d'utilité »). Ensuite, étant donné leur prise de position non-cognitivistique initiale, il est difficile pour les économistes de justifier le recours au principe d'optimalité de Pareto. S'agit-il d'un critère normatif (minimal) ? Si tel est le cas, comment le justifier ? La plupart du temps, la défense du principe d'optimalité consiste à affirmer que ce critère reflète uniquement un engagement normatif *minimal* – sur lequel tout le monde est d'accord. Pour les différentes manières de défendre le principe de Pareto, se référer à l'ouvrage de Marc Fleurbaey, *Théories économiques de la justice* (FLEURBAEY, 1996).

D'après nous, les économistes devraient plutôt défendre le principe d'optimalité de Pareto à l'aide d'une théorie cognitiviste, par exemple la théorie contractualiste de Thomas Scanlon : la raison d'accepter le principe d'optimalité de Pareto est le fait que « personne n'a de raison de s'opposer au principe » (SCANLON, 1998). Si nous avons raison, alors cela exigerait de la part des économistes une révision de leurs engagements métaéthiques – du moins, de leurs engagements métaéthiques explicites. Malheureusement, il me semble que peu d'économistes, à l'exception, par exemple, de Mongin, ne se sont intéressés aux engagements métaéthiques de leur discipline. Réciproquement, nous n'avons pas connaissance de philosophes qui se seraient intéressés à cette question – contrairement à la question des engagements *normatifs* ou des engagements *ontologiques* de la discipline. Cela constituerait une excellente question de recherche pour le futur.

Nous avons étudié ces deux premières approches et nous avons avancé des raisons de penser qu'elles posent chacune des difficultés. Il existe, cependant, une troisième manière de concevoir la relation entre les énoncés descriptifs (au sujet des motivations humaines) et les énoncés moraux.

(C) L'approche moraliste

Selon l'approche moraliste, il est possible de dériver des énoncés *descriptifs* au sujet des motivations humaines sur la base d'énoncés *normatifs* portant sur les devoirs moraux des êtres humains. Dans cette approche, la recherche en philosophie morale est prioritaire par rapport à la description des motivations humaines, et notamment de la motivation morale. Non seulement les énoncés normatifs ont des conditions de vérité indépendantes des faits au sujet des motivations humaines, mais, plus encore, les énoncés normatifs permettent de dériver des énoncés descriptifs au sujet de la motivation morale des êtres humains.

Cette approche semble, de prime abord, véritablement contre-intuitive. En effet, comment peut-on, sur la base d'un énoncé sur *ce qui devrait être*, dériver un énoncé descriptif au sujet de *ce qui est*? Il s'agit néanmoins de la relation entre énoncés descriptifs et normatifs que nous allons essayer de défendre dans le présent chapitre. Il s'agit de l'approche dite *moraliste*.

Le représentant emblématique de l'approche moraliste est, évidemment, Emmanuel Kant. Selon lui, du fait que les êtres humains ont le *devoir* de faire une action, alors les êtres humains *peuvent* se motiver à accomplir leur devoir – peu importe l'importance des obstacles motivationnels que l'individu doit surmonter, et peu importe la quantité de sacrifices exigés. Selon le philosophe, l'être humain est toujours capable de trouver en lui les forces motivationnelles suffisantes pour accomplir son devoir et devenir une personne vertueuse. Comme l'illustre les passages suivants, Kant adhère au principe « devoir-implique-pouvoir » – dans sa forme non contraposée et non empiriste :

Mais si l'homme est corrompu dans le fond de ses maximes, comment est-il possible qu'il effectue par ses propres forces la révolution dont nous parlons et devienne de lui-même un homme de bien ? Le devoir, d'un autre côté, ordonne d'être tel, et il n'ordonne rien que ce qu'il nous est possible d'accomplir. C'est que la révolution dans

les pensées, et la réforme successive des sentiments [...] doivent par conséquent aussi être possibles à l'homme. En d'autres termes [...] il n'a qu'à travailler avec persistance à devenir un homme vertueux [...]. Car, du moment que la loi morale déclare que nous devons être actuellement des hommes meilleurs, il s'ensuit que nous devons aussi le pouvoir (KANT, 1841 (1793)).

Emmanuel Kant s'opposait, par sa théorie, aux théologiens qui considéraient que seule la grâce divine devait permettre à l'homme, enfant du péché originel, de faire le bien – ou pire encore, qui considéraient que l'homme était prédestiné soit au bien, soit au mal, en fonction d'une volonté divine indéchiffrable. Selon Kant, au contraire, l'homme possède en lui-même les ressources motivationnelles suffisantes pour accomplir son devoir moral. Il n'a pas besoin du secours divin.

Nous étudierons plus loin dans le chapitre des auteurs qui reprennent, sous une forme renouvelée, l'argument moraliste kantien en faveur de la thèse selon laquelle l'être humain est capable de se motiver à faire le bien, à faire ce qui est juste ou à devenir une personne vertueuse, même si cela ne satisfait aucun de ses désirs. Nous essaierons de défendre cet argument. Nous nous montrerons donc *kantiens* dans notre propos, mais nous ne serons kantiens que dans la mesure où nous jugeons légitime de dériver des énoncés motivationnels *descriptifs* à partir de prémisses morales *normatives* et pour aucune autre raison.³³⁶

³³⁶ Il existe dans la littérature métaéthique actuelle, selon nous, des conceptions très diverses de ce qui constitue une métaéthique « kantienne ». Cela rend très difficile l'identification d'un courant spécifiquement « kantien ». Pour répondre à cette difficulté, nous proposons de distinguer quatre courants de pensée métaéthiques dont les partisans se déclarent « kantiens » : (1) Pour les uns, il suffit de considérer que les énoncés moraux sont des *énoncés synthétiques a priori* pour que la métaéthique proposée soit considérée comme « kantienne » : « L'interprétation [...] intuitionniste rationaliste [...] affirme que Kant décrit l'impératif catégorique comme synthétique a priori parce qu'il exprime une saisie directe de la vérité d'un principe moral substantiel par le biais de nos facultés rationnelles. Cette interprétation place Kant à l'intérieur de la tradition rationaliste Platoniste. En philosophie morale, le rationalisme intuitionniste plus récent inclut H. A. Prichard et W. D. Ross. » [“The [...] “rational intuitionist” interpretation [...] says that Kant describes the categorical imperative as synthetic a priori because it expresses a direct insight by our rational faculties into the truth of a substantive moral principle. This interpretation places Kant within the Platonic rationalist tradition. In moral philosophy more recent rational intuitionist include H. A. Prichard and W. D. Ross.” (POTTER, 1997, p. 438)]. Dans ce cas, tous les partisans de l'intuitionnisme devraient être considérés comme « kantiens », des plus anciens (PRICHARD, 1912; ROSS, 1930; MOORE G. E., 1993 (1903); REINACH A., 2004 (1913)) aux plus récents (AUDI, 2001; HUEMER, 2005; SHAFER-LANDAU, 2005; WEDGWOOD, 2007) – malgré l'opposition fervente de Kant à l'intuitionnisme. (2) Pour d'autres philosophes, une métaéthique n'est véritablement « kantienne » que si elle repose sur l'idée d'« auto-législation ». Dans ce courant, deux positions différentes méritent d'être distinguées ; (a) La première est la théorie de l'« auto-législation du sujet » défendue par Christine Korsgaard (1996) et Jürgen Habermas (1999). Dans cette approche, la métaéthique kantienne consiste à localiser l'origine

1.2. QU'EST-CE QUE LE SOPHISME MORALISTE ?

Dans sa structure générale, un *sophisme moraliste* consiste à dériver fallacieusement un énoncé *descriptif* sur la base de prémisses *morales*. Ce sophisme représente donc l'exact opposé du *sophisme naturaliste* que nous avons vu précédemment. En effet, alors que le sophisme naturaliste consiste à dériver fallacieusement un « devoir-être » à partir de l'« être », le sophisme moraliste consiste à dériver fallacieusement un « être » à partir d'un « devoir-être ».

Avant d'étudier les cas de véritables *sophismes moralistes*, il convient de noter que toute déduction allant du normatif au descriptif n'est pas forcément fallacieuse. En effet, il existe des situations dans lesquelles la logique (seule) permet de dériver un énoncé descriptif sur la base d'un énoncé normatif. Une telle situation pourrait être la suivante :

(1) Cet homme porte un chapeau melon vert et a le devoir moral de subvenir aux besoins de ses enfants.	Prémisse normative	$\exists x(Nx \wedge Mx)$ <i>N</i> désigne une propriété naturelle et <i>M</i> une propriété morale.
(2) ∴ Cet homme porte un chapeau melon vert.	Conclusion descriptive	$\exists x(Nx)$

Dans cette déduction, il semble que nous avons, de façon tout à fait légitime, c'est-à-dire sans commettre d'erreur logique, dérivé une proposition *descriptive* à partir d'une prémisse *normative*. Il n'est donc pas le cas que toute déduction qui obtient une conclusion descriptive sur la base d'une prémisse normative soit fallacieux.

Toutefois, comme l'illustre bien la déduction précédente, la présence de la notion *normative* dans la prémisse ne joue aucun rôle dans la déduction de la conclusion. Ni la signification *morale* du prédicat *M*, ni même la présence de ce prédicat dans la prémisse, ne joue de rôle dans la déduction de la conclusion. Autrement dit, le fait que *x* soit *M* n'affecte

de la morale dans une loi que le sujet *se donne à lui-même*. (b) La seconde est la théorie de l'« auto-législation de la loi morale » défendue par Onora O'Neill (2015). Selon la philosophe, Kant considère que la « loi morale » est une loi « auto-légiférante », c'est-à-dire une loi qui se justifie d'elle-même. En outre, cette « loi morale », à savoir l'impératif catégorique, est la source *non seulement* des règles morales, mais également des règles régissant l'activité *théorique* – tel que le principe de non-contradiction. Pour finir, (3) certains philosophes, dont nous faisons partie, retiennent de Kant son argument visant à dériver la connaissance de la liberté humaine sur la base de la connaissance de la loi morale, c'est-à-dire son argument visant à dériver une connaissance descriptive sur la base d'une connaissance normative (POTTER, 1997; TIMMONS, 1985; NAGEL, 1970).

aucunement la validité de l'inférence. Même si x n'avait pas été M , l'inférence aurait été valide – par exemple, si l'homme au chapeau melon vert, au lieu d'avoir la propriété de devoir subvenir aux besoins de ses enfants, avait plutôt eu la propriété d'avoir des chaussettes rouges. De ce fait, nous ne sommes pas en présence d'un cas dans lequel une prémisse descriptive est dérivée d'une prémisse *normative*, puisque le contenu normatif ne joue aucun rôle (ni sémantique, ni logique). De notre côté, nous considérerons qu'une déduction constitue un exemple de *sophisme moraliste* uniquement dans les cas où la conclusion descriptive dérive du contenu normatif de la prémisse. Ce serait le cas dans une déduction du type suivant :

(1) a est M

(2) $\therefore a$ est N

où M désigne une propriété morale et N une propriété naturelle.

Bien entendu, en l'état, la déduction est logiquement inacceptable. La question consistera précisément, dans la suite du travail, à déterminer s'il existe des circonstances dans lesquelles une telle déduction (logiquement invalide) peut être justifiée, notamment par le recours à l'analyse sémantique de la notion de « devoir ».

1.2.1. QUELQUES EXEMPLES

Nous pouvons commencer par présenter quelques exemples de sophismes moralistes assez récurrents dans le langage ordinaire, mais aussi dans la science et la philosophie. Cela nous permettra de constater que le sophisme moraliste constitue une forme récurrente d'argument fallacieux.

(i) Les différences d'intelligence entre les noirs et les blancs

(1) Il n'existe pas de différence de statut moral entre les noirs et les blancs – ils doivent être traités de la même manière, avoir les mêmes droits politiques, économiques et sociaux.	Prémisse morale normative
(2) $\therefore$ Il n'existe pas de différence naturelle de QI entre les blancs et les noirs – i.e. les différences de QI observées entre les blancs et les noirs ne sont pas explicables par des facteurs naturels, telles que la	Conclusion descriptive

taille du cerveau, mais sont explicables par des facteurs socio-économiques.³³⁷

(ii) La répartition de tâches éducatives au sein du foyer

(1) Il ne doit pas exister de différence entre les hommes et les femmes dans la répartition des tâches liées au soin des enfants.	Prémisse normative
(2) ∴ Il n'existe pas de différences naturelles entre les hommes et les femmes concernant leur capacité à s'occuper de leur progéniture.	Conclusion descriptive

(iii) Le caractère non naturel des vices

(1) La guerre, la maltraitance et la violence gratuite sont immorales.	Prémisse normative
(2) ∴ Le désir de guerre, de maltraitance et de violence gratuite ne sont pas des traits naturels de l'être humain – mais sont plutôt le résultat de facteurs socio-économiques. ³³⁸	Conclusion descriptive

³³⁷ Un article publié en 2005 reproche à la communauté scientifique de commettre constamment ce sophisme moraliste. Nous avons choisi cet exemple précisément pour son caractère controversé, sans forcément adhérer aux thèses soutenues. Voici le paragraphe introductif de l'article en question : « Malgré des affirmations répétées du contraire, il ne s'est pas produit de réduction de la différence de 15 à 18 points de QI entre les noirs et les blancs (1.1. de déviations standards) ; les différences sont aussi importantes aujourd'hui qu'elles l'étaient lorsqu'elles ont été mesurées pour la première fois il y a presque 100 ans. Ces différences, ainsi que les différences concomitantes dans les niveaux de vie, les niveaux d'éducation et autres phénomènes relatifs, sont dues à des facteurs hérédités, plutôt qu'à des facteurs culturels. Les différences de QI sont attribuables à des différences de taille du cerveau plus qu'à du racisme, à des dangers de stéréotypes, à la sélection des éléments du test, ou à tout autre facteur suggéré par les commentateurs. Il est temps d'accepter la réalité. Il n'est plus temps de commettre "le sophisme moraliste" selon lequel la bonne science doit se conformer à des résultats qui sont approuvés. » [“Despite repeated claims to the contrary, there has been no narrowing of the 15- to 18-point average IQ difference between Black and Whites (1.1 standard deviations); the differences are as large today as they were when first measured nearly 100 years ago. They, and the concomitant difference in standard of living, level of education, and related phenomena, lie in factors that are largely heritable, not cultural. The IQ differences are attributable to differences in brain size more than to racism, stereotype threat, item selections on test, and all other suggestions given by commentators. It is time to meet reality. It is time to stop committing the “moralistic fallacy” that good science must conform to approved outcomes.” (JENSEN & RUSHTON, 2005, p. 328)].

³³⁸ Steven Pinker (1997) est connu pour avoir critiqué notamment la prise de décision de l'UNESCO dans le « Manifeste de Séville sur la violence », dans lequel l'UNESCO affirme que la violence ne peut pas être attribuée à des causes naturelles, notamment biologiques, du fait que la violence est immorale. Dans ce manifeste, en effet, sous prétexte de condamner le fait que certains résultats scientifiques ont été utilisés pour défendre une certaine position morale, par exemple la légitimité de la violence, les signataires du manifeste de Séville commettent l'erreur inverse : afin de défendre leur position morale – condamner la violence – ils

(iv) Le caractère naturel des vertus

- (3) Les êtres humains ont le devoir moral de coopérer, de partager leurs ressources et de produire des créations artistiques. Prémisse normative
- (4) ∴ les êtres humains ont une tendance naturelle à coopérer, à partager leurs ressources et à produire des créations artistiques.³³⁹ Conclusion descriptive

rejettent tout simplement la validité des résultats scientifiques. Le passage d'une condamnation de certains *usages* qui ont été faits de la science, au rejet des *thèses scientifiques* elles-mêmes, est évident dans le passage suivant : « Le mauvais usage de faits et théories scientifiques dans le but de légitimer la violence et la guerre, sans être un phénomène nouveau, est étroitement associé à l'avènement de la science moderne. Par exemple, la théorie de l'évolution a ainsi été « utilisée » pour justifier non seulement la guerre, mais aussi le génocide, le colonialisme et l'élimination du plus faible. [Donc] (1) IL EST SCIENTIFIQUEMENT INCORRECT que nous ayons hérité de nos ancêtres les animaux une propension à faire la guerre. Bien que le combat soit un phénomène largement répandu au sein des espèces animales, on ne connaît que quelques cas au sein des espèces vivantes de luttes destructrices intra-espèces entre des groupes organisés. En aucun cas, elles n'impliquent le recours à des outils utilisés comme armes. Le comportement prédateur s'exerçant à l'égard d'autres espèces, comportement normal, ne peut être considéré comme équivalent de la violence intra-espèces. La guerre est un phénomène spécifiquement humain qui ne se rencontre pas chez d'autres animaux. [Donc (2)] IL EST SCIENTIFIQUEMENT INCORRECT de dire que la guerre ou toute autre forme de comportement violent soit génétiquement programmée dans la nature humaine. [Donc (3)] IL EST SCIENTIFIQUEMENT INCORRECT de dire qu'au cours de l'évolution humaine une sélection s'est opérée en faveur du comportement agressif par rapport à d'autres types, etc.. ». (UNESCO, 1986).

³³⁹ Noam Chomsky est accusé, par Steven Pinker (2002, p. 301), d'adopter une telle conception la nature humaine dans le but de défendre sa position morale, à savoir l'anarchisme socialiste. Selon Pinker, la conception de la nature humaine décrite par Chomsky *dérive* de l'engagement moral préalable de ce dernier. « Le sophisme moraliste [...] sous-tend la croyance romantique que les humains ne peuvent pas posséder un désir de tuer, de violer ou de voler, parce que cela serait trop déprimant et réactionnaire. [...] L'[exemple] le plus fameux est celui de mon collègue du MIT Noam Chomsky, qui croit que les personnes ont des tendances innées pour la coopération, le partage et le travail créatif, ce qui justifierait une forme d'anarchisme socialiste. Il s'agit d'une vue plutôt romantique de la nature humaine, qui ignore complètement le Darwinisme moderne – vous ne pouvez pas être anarchistes si vous n'êtes pas romantiques, et vous ne pouvez pas être romantiques si vous êtes Darwinistes. » [“The moralistic fallacy [...] lies behind the romantic belief that humans cannot harbor desires to kill, rape, lie, or steal because that would be too depressing or reactionary [...]. The most famous [example] is my MIT colleague Noam Chomsky who believes that people have innate tendencies to cooperate, share, and produce creative works, justifying a kind of socialist anarchism. This is a rather romantic view of human nature that is innocent of modern Darwinism -- you can't be an anarchist unless you're a romantic, and you cannot be a romantic if you're a Darwinian.”] Propos tirés d'un entretien (SAILER, 2002). Bien entendu, Noam Chomsky soutient qu'il raisonne dans l'autre sens ; selon lui, sa doctrine politique de l'anarchisme socialiste dérive de l'observation scientifique de la nature humaine comme étant naturellement bienveillante, coopératrice et créative. Il s'agit là, nous semble-t-il, d'une caractéristique récurrente des philosophes qui commettent un sophisme moraliste ; ils croient dériver leur position normative sur la base d'énoncés scientifiques descriptifs, mais, en fait, ils dérivent plutôt leur conception de la nature humaine à partir d'un certain point de vue moral.

(v) L'envie, les blagues racistes, la jalousie et la tristesse

Des cas de « sophisme moraliste » apparaissent également dans l'analyse des émotions en philosophie. En effet, Justin d'Arms et Daniel Jacobson observent que plusieurs philosophes, partant du constat qu'il est toujours moralement inapproprié, par exemple, de ressentir de l'envie ou d'être amusé par une blague raciste, concluent qu'il n'est jamais correct ou adapté de ressentir de l'envie et qu'une blague raciste ne peut jamais être drôle. Cependant, selon d'Arms et Jacobson, ces philosophes confondent deux sens du terme « approprié » ; dans le premier sens, une émotion peut *être adaptée (fitting)* à son objet, dans le sens qu'elle le représente adéquatement (D'ARMS & JACOBSON, 2000, p. 66). Ainsi, il peut être *adapté ou correct (fitting)* d'être amusé par une blague raciste du seul fait que la blague en question possède les caractéristiques d'une blague *drôle*. Il peut être *adapté (fitting)* d'éprouver de l'envie lorsqu'un objet est effectivement enviable – notamment, lorsqu'un concurrent possède des caractéristiques favorables qui nous font défaut.

Dans un second sens, les termes *approprié et inapproprié* sont utilisés dans un sens proprement *moral* : il ne s'agit plus de savoir si l'émotion représente adéquatement l'objet, mais s'il est moralement acceptable ou inacceptable pour une personne d'éprouver cette émotion. Par exemple, il est possible qu'il soit *toujours immoral (wrong)* d'être amusé par une blague raciste et qu'il soit toujours immoral d'éprouver de l'envie. Cependant, le fait que rire à une blague raciste et éprouver de l'envie soient toujours moralement condamnables ne permet pas de déduire que les blagues racistes ne sont jamais drôles et que nos concurrents privilégiés ne sont pas véritablement enviables. Au contraire, il est tout à fait possible que la situation d'autrui soit *enviable* et qu'une blague raciste soit effectivement *drôle*, même s'il est immoral d'envier son concurrent et d'être amusé par une blague raciste. Conclure qu'il n'existe rien d'enviable et qu'aucune blague raciste n'est drôle, du fait que l'envie et l'amusement à l'égard des blagues racistes sont des émotions immorales, cela revient à commettre un « sophisme moraliste ».

Commettre un sophisme moraliste consiste à inférer, de l'affirmation qu'il est immoral ou vicieux de ressentir une émotion, que cette émotion doit, en conséquence, être inappropriée (*unfitting*). Nous soutiendrons, à l'inverse, qu'une

émotion peut être *appropriée* (*fitting*) malgré le fait qu'il est immoral (ou inopportun) de la ressentir.³⁴⁰

La manière la plus manifeste de commettre un sophisme moraliste est simplement d'inférer, de l'affirmation qu'il serait moralement condamnable d'éprouver F à l'égard de X, qu'en conséquence F n'est pas une réponse appropriée (*fitting*) à X. Cette inférence est fallacieuse ; elle est réfutée par notre maxime selon laquelle une émotion peut être appropriée bien qu'il soit immoral de la ressentir.³⁴¹

Voici deux autres exemples de "sophisme moraliste" proposés par d'Arms et Jacobson, qui illustrent le caractère fallacieux – mais pourtant séduisant – du sophisme moraliste.

(vi) La jalousie

(1) Il est moralement inapproprié de ressentir de la jalousie vis-à-vis de l'être aimé.	Prémisse normative
(2) ∴ la jalousie n'est jamais appropriée (<i>fitting</i>). Elle représente l'être aimé de manière inadaptée, c'est-à-dire d'une manière qui ne correspond pas à la réalité. Ainsi, dans la jalousie, l'être aimé est représenté comme étant une <i>possession</i> , alors que les êtres humains ne sont jamais la possession de personne.	Conclusion descriptive

(vii) Le chagrin

(1) Il est moralement inapproprié, pour une jeune femme devenue veuve récemment, de ressentir trop de chagrin – car elle doit s'assurer que ses enfants continuent de vivre, sans trop de traumatisme.	Prémisse normative

³⁴⁰ "[T]o commit the moralistic fallacy is to infer, from the claim that it would be wrong or vicious to feel an emotion, that it is therefore unfitting. We shall contend, to the contrary, that an emotion can be fitting despite being wrong (or inexpedient) to feel." (D'ARMS & JACOBSON, 2000, p. 69).

³⁴¹ "The most blatant way to commit the moralistic fallacy is simply to infer, from the claim that it would be morally objectionable to feel F toward X, that therefore F is not a fitting response to X. This inference is fallacious; it is belied by our dictum that an emotion can be fitting despite being wrong to feel." (D'ARMS & JACOBSON, 2000, pp. 75-76).

- (2) ∴ il n'est pas triste de perdre son époux – la Conclusion descriptive
tristesse n'est pas une réponse appropriée à la
perte de son conjoint

Dans le premier exemple, celui de la jalousie, le caractère fallacieux de l'argument n'apparaît pas immédiatement. Au contraire, il peut être tentant de considérer que, étant donné que la jalousie est moralement condamnable, alors la jalousie est une émotion qui est toujours *incorrecte* ou *inadaptée* ; elle attribue à l'objet jaloué une propriété que l'objet ne possède pas en fait.

Lorsqu'une émotion semble moralement incorrigible, les philosophes sont tentés de l'accuser de quelque grossière erreur [...]. Par exemple, certains philosophes ont attribué à la jalousie l'affirmation que l'être aimé est une chose susceptible d'être possédée ; ou que n'importe quelle quantité de l'affection de l'être aimé accordée à une autre personne fait diminuer la part d'affection qui revient à l'amant. Bien qu'il s'agisse de deux erreurs interprétatives contraires, elles trahissent la même impulsion. La tentation d'embellir certaines émotions et d'enlaidir d'autres émotions sont toutes les deux des symptômes d'une tendance psychologique que nous nommerons le *moralisme*.³⁴²

Par contre, dans le second exemple, celui du chagrin, il apparaît immédiatement que la conclusion ne suit pas la prémisse : les raisons morales de ressentir ou non une émotion telle que le chagrin ne permettent pas d'évaluer si l'émotion est *adaptée* à son objet. La perte d'un être aimé est effectivement un événement *triste*, c'est-à-dire un événement à l'égard duquel il est adapté de ressentir de la tristesse. Le fait que, dans certaines circonstances, nous avons le devoir moral de ne pas nous laisser aller à notre chagrin n'affecte en rien le fait que notre chagrin représente adéquatement l'objet.

Les cas de sophisme moraliste sont donc suffisamment répandus pour que nous leur accordions de l'intérêt. La question qui nous préoccupe désormais consiste à déterminer s'il

³⁴² "[W]hen an emotion seems morally incorrigible, philosophers are tempted to convict it of some gross error [...]. For example, some philosophers have attributed to jealousy the claim that the beloved is a thing capable of being possessed; or that any amount of the beloved's affection given to another diminishes the lover's share. Although these are contrary interpretive mistakes, we think that they betray a common impulse. The temptation to pretty up some emotions and to dumb down others are both symptoms of a psychological tendency which we will call moralism." (D'ARMS & JACOBSON, 2000, p. 75).

existe des cas dans lesquels le raisonnement moraliste pourrait être valide – contrairement aux exemples que nous avons présentés jusqu'à présent.

Il est évident qu'une manière très simple de résoudre l'invalidité *logique* de tels sophismes moralistes consiste à ajouter une prémisse intermédiaire, telle que, pour reprendre l'exemple de la différence de QI entre noirs et blancs, « s'il est moralement correct de traiter différemment les individus noirs et les individus blancs, alors les individus noirs et les individus blancs ne disposent pas des mêmes compétences intellectuelles ». Toutefois, cette solution ne résout rien de plus que la difficulté logique : en effet, comme le fait remarquer justement Mark Nelson, simplement ajouter une prémisse pour rendre valide un raisonnement n'apporte pas de soutien supplémentaire à la conclusion (NELSON, 2003). Ce n'est que lorsque des *raisons* d'adopter la prémisse intermédiaire auront été données que la conclusion reposera sur un fondement épistémique solide. En résumé, celui qui soutient « s'il est moralement correct de traiter différemment les individus noirs et les individus blancs, alors les individus noirs et les individus blancs ne disposent pas des mêmes compétences intellectuelles » peut bien se vanter d'avoir supprimé l'erreur logique de son raisonnement. Cependant, il n'est parvenu qu'à repousser le problème à un niveau supérieur : la question est désormais de savoir s'il y a de bonnes raisons de croire que cette prémisse intermédiaire est vraie.

1.2.2. LE SOPHISME MORALISTE MOTIVATIONNEL

Le sophisme moraliste qui nous intéresse dans ce travail consiste plus spécifiquement à adopter une certaine conception de la nature humaine – et, surtout, de ses motivations – sur la base d'une doctrine morale spécifique ; autrement dit, il consiste à concevoir l'être humain et ses motivations précisément de la manière dont on a besoin pour soutenir une certaine position morale. Une telle faute de raisonnement se rencontre, nous pensons, chez plusieurs philosophes moraux. Par exemple, nous suivons Jean-Louis Gardies pour affirmer que les partisans modernes du droit naturel – notamment Rousseau, Hobbes et Locke – ont certainement choisi la conception de la « nature humaine » qui convenait le mieux à leur position morale :

[La] nature humaine, dira-t-on, est si complexe que chacun lui donne, en fonction de ses préférences morales, sociales ou politiques, la forme qu'il souhaite. Il est

en effet difficile de nier que les jusnaturalistes ont pour la plupart façonné le concept de *nature*, d'une manière sans doute souvent inconsciente, en tout cas toujours inavouée, en sorte qu'ils puissent en déduire *naturellement* les règles qui faisaient l'objet soit de leur revendication soit plus souvent de leur volonté de conservation (GARDIES, 1972, p. 14).³⁴³

Cet acte de considérer comme « naturels » les traits de la nature humaine les mieux adaptés à la théorie morale favorisée par l'auteur était déjà courant, selon Gardies, également dans la philosophie antique.

En face des Sophistes, qui par l'observation de la nature humaine cherchaient à établir l'égalité de tous les hommes, qu'ils fussent barbares ou hellènes, Aristote soulignait déjà le fait de leur inégalité, pour fonder sur la nature l'institution de l'esclavage. Ainsi la prétention de déduire le *devoir-être* de *l'être* suscite-t-elle, outre les objections proprement logiques, une suspicion morale assez justifiée par les faits (GARDIES, 1972, p. 14).³⁴⁴

Du fait que, chez de nombreux auteurs, la sélection du trait de la nature humaine devant servir à fonder une théorie morale est motivée, en fait, par une conception morale antécédente, cela a pour conséquence que l'appel à la nature humaine suscite de légitimes suspicions *morales* – outre les suspicions logiques.

En plus des philosophes modernes et antiques, nous retrouvons des exemples de raisonnement moraliste chez des auteurs contemporains : nous pensons notamment aux métaéthiciens réalistes tels que Thomas Scanlon (1998) ou encore Rainer Forst (2012). Il nous semble que, chez ces auteurs également, la conception adoptée de la motivation morale dérive de leur position normative. En effet, ces auteurs *commencent* par déterminer ce qu'est le contenu moral de leur doctrine, avant d'attribuer à l'être humain une motivation

³⁴³ De son côté, David Hume accusait ces mêmes jusnaturalistes de commettre un *sophisme* naturaliste, c'est-à-dire de dériver des propositions morales à partir d'une certaine conception de la nature humaine. Toutefois, il est possible que leur erreur ait été différente : au lieu de commettre, comme le suggérait David Hume, un « sophisme naturaliste », en dérivant leur théorie morale de l'observation de la nature humaine, les théoriciens du droit naturel ont plutôt dérivé leur conception de la nature humaine d'une doctrine morale préalable. Ils ont donc commis un « sophisme moraliste » plutôt qu'un « sophisme naturaliste ».

³⁴⁴ Nous pouvons ajouter l'exemple du Stoïcien Epictète, concernant le caractère immoral et *donc* non naturel de la négligence : « [La] négligence est un mal pour notre personne morale et par là contraire à la nature ». (EPICTÈTE, 1950, livre II, chap. V)

morale correspondante. Par exemple, Thomas Scanlon attribue aux êtres humains le « désir d’agir d’une manière qui peut être justifiée à autrui » :

L’idée de pouvoir justifier son comportement auprès des autres personnes exerce une force motivationnelle.³⁴⁵

L’idée de pouvoir justifier son comportement aux autres rend compte de manière substantielle du fondement de la motivation morale.³⁴⁶

Or, le philosophe ne parvient à cette conception de la *motivation morale* qu’après avoir, préalablement, défini le *contenu* des exigences morales. En d’autres termes, l’auteur infère l’existence d’un désir de ne pas faire des actions qui ne peuvent pas être justifiées auprès d’autres personnes sur la base de l’existence d’un *devoir moral* de ne pas faire des actions qui ne peuvent pas être justifiées auprès des autres personnes (SCANLON, 1998, p. 4).

Plus formellement, l’argument de Scanlon se présente de la manière suivante :

- (1) Est immorale/injuste toute action qui ne peut pas être justifiée auprès des autres personnes – en des termes que ces personnes peuvent accepter.
- (2) ∴ Donc il est possible d’être motivé à renoncer à faire une action par la seule croyance que cette action ne peut pas être justifiée auprès des autres personnes.³⁴⁷

L’argument de Thomas Scanlon constitue un exemple du type d’argument qui nous intéresse particulièrement. En effet, cet argument consiste à attribuer *un certain contenu* à la motivation morale – ainsi que d’autres caractéristiques plus formelles, sur lesquelles nous

³⁴⁵ “[T]he idea that justifiability to others has motivational force.” (SCANLON, 1998, p. 170).

³⁴⁶ “The idea of justifiability to others provides a substantive account of the basis of moral motivation.” (SCANLON, 1998, p. 338).

³⁴⁷ Par souci de complétude – mais sans que cela ne change la structure de notre argument – il nous faut préciser que le philosophe ajoute la prémisse suivante pour permettre la transition de (1) à (2):

- (3) Le fait qu’une action ne peut pas être justifiée auprès des autres personnes constitue une raison de ne pas la faire.

Comme le montre cet ajout d’une prémisse intermédiaire, c’est la notion normative de *raison d’agir* – dont les raisons morales ne forment qu’un sous-ensemble – qui permet, dans la théorie de Scanlon, de dériver une conclusion *descriptive* au sujet des motivations humaine sur la base d’un énoncé normatif. Son raisonnement « moraliste », se présente donc, plus précisément, sous la forme suivante :

- (4) S a une raison – indépendante de ses désirs antécédents – de faire X.
- (5) ∴ S peut être motivé à faire X – indépendamment de ses désirs antécédents.

D’un énoncé *normatif* au sujet des *raisons* d’agir, Scanlon dérive un énoncé *descriptif* au sujet des capacités motivationnelles humaines. Il adopte donc bien le raisonnement « moraliste » que, jusqu’à présent, nous avons caractérisé comme étant fallacieux.

reviendrons – sur la base d’une théorie normative définissant *ce qu’il est bon, juste ou vertueux* de faire. Or, malgré l’apparence de sophisme, il s’agit, comme nous allons le soutenir par la suite, de la seule manière appropriée de raisonner, lorsque nous discutons de la motivation morale.

1.2.3. LE CARACTÈRE SUSPECT DE LA MOTIVATION MORALE

L’une des raisons pour lesquelles la motivation morale³⁴⁸ suscite des suspicions dans la littérature est qu’il est difficile de comprendre comment des personnes peuvent rationnellement être motivées à agir par des idées aussi abstraites et dépourvues de tout contenu que celles de *bien, de rectitude ou (éventuellement) de vertu*. Le cas le plus évident est celui de l’idée de *bien (moral)* ; si le bien moral est une propriété simple, sans contenu spécifique, alors la motivation morale correspondante se réduirait à un désir de faire le bien, dont il est difficile d’expliquer le caractère rationnel. En effet, ce désir de faire le bien pourrait être expliqué de deux manières : la première est que la perception ou la saisie intellectuelle du bien exerce un pouvoir causal direct sur la motivation des personnes. Cependant, dans ce cas, nous n’avons pas d’explication à la question *pourquoi*³⁴⁹ la personne accorde de l’intérêt ou de l’importance au bien ; elle le fait, tout simplement. *Quiconque* perçoit la propriété mystérieuse et simple du *bien* se trouve *de fait* motivé à le réaliser.

La seconde explication serait qu’il existe déjà, chez certaines personnes du moins, un désir général de « faire le bien » (*de dicto*). La force motivationnelle de ce désir est *déclenchée* lorsque la personne identifie une action particulière qui exemplifie la propriété du bien. Toutefois, dans ce cas, lorsqu’une personne agit en vue de faire le bien, elle ne fait que répondre à une « préférence pour le bien », préférence qui n’est pas plus raisonnable que n’importe quelle autre préférence *donnée*, comme, par exemple, une préférence pour le café. Il s’agit d’une sorte de *goût* pour la morale, dont il n’est pas possible de discuter et d’expliquer le caractère raisonnable.³⁵⁰

³⁴⁸ Pour rappel : motivation à faire le bien, à faire ce qui est juste ou à être une personne vertueuse.

³⁴⁹ Ce « pourquoi » n’indique pas que nous cherchons une raison non morale d’agir moralement. Il indique que nous cherchons à en savoir *plus* au sujet de la morale elle-même, au sujet de sa nature *substantielle*.

³⁵⁰ Scanlon critique Prichard en ce sens. Selon lui, si Prichard montre bien que la motivation morale ne peut pas être réduite à une motivation *non morale*, il est incapable d’expliquer quelle est la caractéristique du bien moral qui explique que nous lui accordions de l’intérêt. Or, s’il n’est pas possible de dire, au sujet du bien, autre chose que le fait qu’il s’agit du bien, nous n’avons pas expliqué ce qui justifie que nous lui accordions de

De plus, comme le fait remarquer Rawls, si la motivation morale devait se réduire à une simple « préférence », alors nous devrions conclure qu'il est vain de chercher à bâtir une société sur la base d'une motivation aussi instable, peu raisonnable et mal répartie qu'une « préférence pour le café ».

Ross soutient que le sens du juste est un désir pour un objet distinct (non analysable) puisqu'une propriété spécifique (non analysable) caractérise les actions qui constituent notre devoir. [...] Mais, dans cette interprétation, le sens du juste n'a aucune raison apparente, il ressemble à une préférence pour le thé plutôt que pour le café. Bien qu'une telle préférence puisse exister, en faire une règle de la structure de base de la société est profondément arbitraire (RAWLS, 2009 , pp. 517-518).

Pour finir, à la suite de Michael Smith, il est possible de reprocher à ce désir de « faire le bien » de générer une forme de « fétichisme pour la morale », c'est-à-dire une attention exagérée pour l'aspect moral des situations (SMITH, 1994, pp. 74-75).

Par contre, lorsque la motivation morale est définie par un contenu spécifique, c'est-à-dire *substantiellement*, elle devient beaucoup moins mystérieuse ; il devient possible de comprendre pourquoi les personnes accordent de l'importance au bien, au juste ou à la vertu. Bien entendu, pour connaître la nature du *contenu* de la motivation morale, il est nécessaire d'avoir, au préalable, défini le contenu de la morale. Il est nécessaire, autrement dit, de posséder une théorie morale normative.³⁵¹ Le contenu de la motivation morale ne peut pas être défini substantiellement avant d'avoir déterminé quelle doctrine morale, entre

l'importance (SCANLON, 1998, p. 150). Parce qu'il ne peut pas rendre compte de l'intérêt que nous portons à la morale par le biais du *contenu* de la morale, Prichard (1912) doit expliquer la motivation morale par le biais d'un désir contingent *de dicto* de « faire son devoir », désir, préférence ou *goût* que certaines personnes possèdent et d'autres ne possèdent pas. Ce désir est *donné* et *non motivé*, c'est-à-dire qu'il ne repose pas sur des *raisons*.

³⁵¹ Nous rejoignons ainsi Thomas Nagel lorsqu'il affirme que le seul moyen de répondre au sceptique, lorsque celui-ci demande "pourquoi devrais-je agir moralement?" est de développer une éthique *substantielle* plutôt que de faire de la métaéthique. Nous le rejoignons ainsi également pour affirmer que la philosophie morale *normative* est prioritaire par rapport à la psychologie et à la métaéthique. « Au bout du compte, il s'agit d'une compétition entre la crédibilité de l'éthique substantielle et la crédibilité d'une réduction psychologique externe de cette activité [...]. C'est pourquoi on ne peut défendre la raison morale qu'en abandonnant la métaéthique en faveur de l'éthique substantielle. Seul le poids intrinsèque de la pensée morale de premier ordre peut contrer les doutes subjectivistes ». ["So in the end, the contest is between the credibility of substantive ethics and the credibility of an external psychological reduction of that activity. [...] That is why we can defend moral reason only by abandoning metatheory for substantive ethics. Only the intrinsic weight of first-order moral thinking can counter the doubts of subjectivism." (NAGEL, 1997, p. 115;125).]

l'utilitarisme, le contractarianisme, le contractualisme, le déontologisme ou l'intuitionnisme, est la bonne théorie. Autrement dit, il ne peut exister de discussion crédible au sujet de la motivation morale *avant* de posséder une théorie morale normative.

Prenons par exemple l'utilitarisme : si l'utilitarisme est vrai, alors la motivation morale consiste à vouloir protéger et promouvoir le bien-être de tous les êtres susceptibles de souffrir. La motivation morale utilitariste consiste donc en une bienveillance généralisée, en un souci général du bien-être de tous les êtres sensibles. Or, il semble que nous pouvons facilement comprendre qu'une personne accorde de l'importance *au bien-être* des autres, alors que nous ne pouvons pas facilement comprendre pourquoi une personne accorde de l'importance à une propriété indéfinissable telle que le « bien ». Réciproquement, il est difficile de comprendre – même si cela n'est pas inconcevable – qu'une personne n'accorde jamais d'importance au bien-être d'autres personnes, alors qu'il n'est pas difficile de comprendre qu'une personne soit complètement indifférente à une notion abstraite et sans substance comme celle du bien.

Notons que, dans l'utilitarisme, la motivation morale n'est pas la bienveillance *partiale* que nous portons naturellement à l'égard de nos proches, mais d'une bienveillance *générale* qui correspond à l'exigence utilitariste d'accorder le même poids au bien-être de tous les êtres sensibles.

[P]our l'utilitarisme la motivation morale caractéristique est la bienveillance. Le terme est vague, et il peut susciter la confusion, particulièrement lorsqu'il suggère des sentiments chaleureux d'attachement personnel ou [...] tout type de sentiment que quiconque ressent avec plus de force pour certaines personnes que pour d'autres. La bienveillance utilitariste n'implique aucun attachement personnel et elle est immunisée contre la loi du carré inversé.³⁵²

Comme on le voit, dans l'utilitarisme, le contenu de la motivation morale est déterminé par le contenu de la théorie morale utilitariste ; la motivation morale ne consiste pas à vouloir *faire le bien* – où le bien est un terme simple, indéfinissable – mais à vouloir protéger et promouvoir le bien-être des autres personnes. Dans l'utilitarisme, une personne

³⁵² "[F]or utilitarianism the characteristic moral motive was benevolence. That term is vague, and it can also be misleading, particularly if it suggests warm feelings of personal attachment or, again, any kind of sentiment one naturally feels in greater degree for some people than others. Utilitarian benevolence involves no particular attachments, and it is immune to the inverse square law." (WILLIAMS, 2006, p. 81).

motivée à agir par sa croyance que l'action *est juste* est simplement une personne motivée à agir par la croyance que l'action permet de promouvoir le bien-être des êtres sensibles.

Dans une autre théorie morale, la nature de la motivation morale variera de manière correspondante. Comme nous l'avons déjà vu, la théorie contractualiste de Thomas Scanlon, par exemple, ne fournit pas seulement une conception de ce qui est *moralement correct ou incorrect (right and wrong)*, mais aussi une certaine conception de la motivation morale. Et cette conception de la motivation morale *dérive* de la conception morale contractualiste préalable – comme le remarque Bernard Williams :

L'idée centrale du contractualisme a été formulée par Thomas Scanlon, en relation avec sa conception de l'inconvenance morale (*moral wrongness*) : « Un acte est incorrect (*wrong*) si son accomplissement dans ces circonstances serait désapprouvé par n'importe quel système de règles pour la régulation générale du comportement que personne ne pourrait raisonnablement rejeter comme base pour un accord général informé et non contraint. [...] La théorie fournit également une conception de la motivation morale. La motivation morale basique est « un désir de pouvoir justifier son action aux autres sur une base que personne ne peut raisonnablement rejeter ». ³⁵³

Considérons encore d'autres théories normatives ; par exemple, si nous adoptons la doctrine morale de Thomas Nagel (1970) ou Charles Larmore (2008), alors la motivation morale consiste à désirer faire le bien des autres personnes autant que le sien propre. Si la théorie morale privilégiée est plutôt celle du *contractarianisme*³⁵⁴, alors il n'existe pas de

³⁵³ "The central idea of contractualism has been formulated by Thomas Scanlon, in relation to its account of moral wrongness: "An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced, general agreement. [...] The theory also gives an account of moral motivation. The basic moral motive is "a desire to be able to justify one's action to others on ground they could not reasonably reject." (WILLIAMS, 2006, p. 75)

³⁵⁴ « Le contractarianisme, qui provient de la conception du contrat social héritée de Hobbes, soutient que les personnes sont principalement motivées à agir par intérêt personnel, et qu'une évaluation rationnelle de leur intérêt les amènera à agir moralement (lorsque les normes morales sont déterminées par la maximisation de l'intérêt commun) et à consentir à l'autorité gouvernementale. Les contractariens arguent que chacun de nous est motivé à accepter la morale « d'abord parce que nous sommes vulnérables à la déprédation des autres, et ensuite parce que nous pouvons tous tirer avantage de la coopération avec autrui. » ["Contractarianism, which stems from the Hobbesian line of social contract thought, holds that persons are primarily self-interested, and that a rational assessment of the best strategy for attaining the maximization of their self-interest will lead them to act morally (where the moral norms are determined by the maximization of joint interest) and to consent to governmental authority. Contractarianism argues that we each are motivated

motivation spécifiquement morale, puisque le motif sollicité dans cette théorie est celui de l'intérêt personnel. Pour finir, si la doctrine morale retenue est celle de Rawls, alors la motivation morale consiste en un désir d'agir d'après des principes qui seraient acceptés par des personnes rationnelles derrière un voile d'ignorance. Plus précisément :

[L]e sentiment de la justice n'est pas un désir qui diffère de celui d'agir d'après des principes que des individus rationnels accepteraient dans une situation initiale donnant à chacun une représentation égale en tant que personne morale [...]
(RAWLS, 2009 , p. 518).

Le contenu de la motivation morale variera donc en fonction de la théorie normative *substantielle* adoptée – ce qui peut sembler trivial, mais qui, dans la littérature contemporaine, n'est pas du tout évident, puisqu'il existe une sorte de crainte générale de parler de la *substance* de la morale et de la motivation correspondante, et une tendance à vouloir uniquement se prononcer sur la *forme* de la morale et de la motivation correspondante.³⁵⁵

Il reste que l'attribution aux êtres humains d'une motivation morale *dont le contenu correspond précisément* à la doctrine morale défendue constitue toujours un cas de *sophisme moraliste*. Les individus sont supposés être motivés à agir précisément par l'idée du bien ou du juste telle qu'elle est décrite dans la théorie normative favorisée, ce qui paraît relever d'un tour de magie. Comment se fait-il que nous puissions, du simple fait que les individus ont le devoir moral de maximiser le bien-être du plus grand nombre (dans l'utilitarisme) ou de ne pas faire d'actions qui ne peuvent pas être justifiées auprès des

to accept morality "first because we are vulnerable to the depredations of others, and second because we can all benefit from cooperation with others" (Narveson 1988, 148)." (CUDD & EFTEKHARI, 2018).]

³⁵⁵ Bernard Williams déplore cette tendance, dans la philosophie contemporaine, à refuser de définir de façon *substantielle* le contenu de la morale : « Toute délimitation de la morale doit impliquer une référence au contenu des jugements, politiques, principes, ou quoique ce soit d'autre, qui sont décrits comme "moraux". Il peut sembler surprenant pour des personnes qui ne sont pas impliquées dans la philosophie morale que quelqu'un ait supposé qu'il puisse en être autrement ; mais il a été défendu, et cela fréquemment, que les vues morales (par opposition aux vues non morales) peuvent être identifiées sans référence à leur contenu, par des considérations tel que le fait qu'il s'agit de maximes pratiques qui sont reconnues comme dominant les autres maximes pratiques. » ["[A]ny significant delimitation of the moral must involve reference to the content of the judgments, policies, principles, or whatever, that are being described as 'moral'. It might perhaps surprise people innocent of moral philosophy that anyone has ever supposed this not to be so; but it has in fact been maintained, and frequently, that moral views (as opposed to non-moral ones) can be identified without reference to their content, by some considerations such as their being practical maxims which are acknowledged as overriding other practical maxims." (WILLIAMS, 1972, pp. 73-74)]. Selon nous, cette tendance s'explique par le fait que les métaéthiciens eux-mêmes présupposent que seuls les énoncés *descriptifs* (métaéthiques) sont susceptibles de composer une véritable science.

autres personnes (dans le contractualisme de Scanlon), conclure que les êtres humains disposent d'un désir de promouvoir le bien-être d'autrui ou d'un désir de pouvoir justifier leurs actions à autrui ? Il semble, à nouveau, que, par un tel argument, nous ne fassions rien de plus commettre une faute grossière de raisonnement.

1.2.4. RÉHABILITER LE RAISONNEMENT MORALISTE

Fort heureusement, il existe une manière, pour les philosophes, de défendre cette manière de raisonner – de l'existence d'un *devoir* à l'existence d'une motivation correspondante. Nous étudierons les arguments en faveur de cette manière de raisonner dans la prochaine section. Pour l'instant, voyons sous quelle forme se présente cette manière de raisonner:

(1) S a le devoir moral de faire X même si S ne possède aucun désir antérieur au jugement moral que l'action X permet de satisfaire.	Prémisse normative
(2) ∴ il est possible pour S d'être motivé à faire X même si S ne possède aucun désir antérieur au jugement moral que l'action X permet de satisfaire.	Conclusion descriptive

La seule explication disponible de (2) est que :

(3) Il est possible pour S d'être motivé à faire X par la seule croyance qu'il a le devoir moral de faire X	Conclusion descriptive

Dans cet argument, l'affirmation de l'existence d'une motivation fondée uniquement sur reconnaissance de la valeur morale de X dérive de l'affirmation de l'existence d'un *devoir moral* dont la validité ne dépend pas de la présence d'un désir antérieur au jugement moral. Autrement dit, de l'existence d'un devoir moral *catégorique* – i.e. dont la validité ne dépend pas des désirs de S antérieurs au jugement moral – l'argument dérive l'affirmation de l'existence d'une *motivation* morale catégorique – i.e. dont l'existence ne dépend pas des désirs de S antérieurs au jugement moral. La conception de la motivation morale est *dérivée* d'une conception du contenu substantiel des devoirs moraux.

Si nous appliquons la structure de cet argument aux exemples du contractualisme de Scanlon et de l'utilitarisme, cela donne les arguments suivants:

Pour le contractualisme :

(1) S a le devoir moral, indépendamment de ses désirs antérieurs, ³⁵⁶ de ne pas accomplir une action qui ne pourrait pas être justifiée auprès des autres personnes.	Prémisse normative
(2) ∴ il est possible pour S d'être motivé, indépendamment de ses désirs antérieurs, à ne pas accomplir une action qui ne pourrait pas être justifiée auprès des autres personnes	Conclusion descriptive

La seule explication disponible de (2) est que :

(3) il est possible pour S d'être motivé à ne pas accomplir une action par la seule croyance que cette action ne pourrait pas être justifiée auprès d'autres personnes.	Conclusion descriptive

Pour l'utilitarisme :

(1) S a le devoir moral, indépendamment de ses désirs antérieurs, de faire l'action qui maximise le bonheur du plus grand nombre.	Prémisse normative
(2) ∴ il est possible pour S d'être motivé, indépendamment de ses désirs antérieurs, à faire l'action qui maximise le bonheur du plus grand nombre.	Conclusion descriptive

La seule explication disponible de (2) est que :

(3) il est possible pour S d'être motivé à faire une action par la seule croyance que cette action maximise le bonheur du plus grand nombre.	Conclusion descriptive

³⁵⁶ Pour alléger la présentation, nous résumons l'idée de « désirs antérieurs au jugement moral » par « désirs antérieurs ».

Dans chacun des exemples, la conception de la motivation morale est *dérivée* de la conception des devoirs moraux. Il s'agit donc à chaque fois d'un cas de *sophisme moraliste*, dans lequel une thèse descriptive est dérivée d'une thèse normative.

Comme il est évident dans ces exemples, la conclusion à laquelle nous parvenons est plus faible qu'attendu : au lieu de conclure que S est motivé à agir par l'idée que l'action ne pourrait pas être justifiée auprès des autres personnes, ou que S est motivé à agir par l'idée que l'action maximise le plus grand bonheur du plus grand nombre, nous parvenons uniquement à la conclusion qu'il est *possible* pour S d'être motivé à agir par la seule idée qu'une action ne pourrait pas être justifiée auprès des autres personnes et qu'il est *possible* pour S d'être motivé à agir par la seule idée que l'action maximise le plus grand bonheur du plus grand nombre. Autrement dit, à partir d'une proposition portant sur le devoir, le présent argument ne nous permet pas de tirer une conclusion sur les motivations *actuelles* de S, mais uniquement sur les motivations *possibles* de S.

Malgré le fait qu'un tel argument permet uniquement d'attribuer des *possibilités ou capacités motivationnelles* aux individus plutôt que des motivations *effectives*, il reste néanmoins que sa conclusion est déjà très forte ; sa force et son importance proviennent de la clause « *par une croyance seule* ». En effet, comme nous y reviendrons plus loin, pour de nombreux philosophes, il est *conceptuellement impossible* d'être motivé à agir *par une croyance seule*, sans la présence d'un désir *indépendant* de cette croyance. Selon eux, le désir est *par définition* ce qui fournit toute motivation à agir, alors que la croyance est *par définition* dépourvue de force motivationnelle. De ce fait, selon eux, « être motivé à agir par une croyance seule » est une impossibilité conceptuelle.

Ainsi, même si la conclusion des arguments moralistes étudiées ne concerne que des *possibilités motivationnelles*, elle formule tout de même une thèse controversée ; à savoir la thèse selon laquelle il est possible d'être motivé à faire son devoir moral (ou à faire le bien ou à devenir une personne vertueuse), même si faire son devoir ne permet de satisfaire aucun des désirs antérieurs à la reconnaissance de la valeur morale de l'action.

La possibilité d'une motivation par la croyance morale seule, à laquelle, par exemple, Scanlon et Nagel adhèrent, représente déjà une thèse controversée. Mais certains philosophes, en l'occurrence Emmanuel Kant et John Rawls, sont allés encore plus loin : non seulement ils considèrent qu'il est possible pour un être humain d'être motivé à faire

l'action morale par la seule croyance que l'action est bonne, juste ou vertueuse, mais ils considèrent *en plus* que cette motivation morale est capable de *surmonter* ou de *dominer* l'ensemble des désirs non moraux qui s'opposeraient à la réalisation de l'action morale. Selon eux, non seulement un individu peut être motivé à faire le bien, à faire ce qui est juste, ou à être une personne vertueuse par la seule croyance qu'il est bien, juste ou vertueux de faire l'action, mais l'individu est, de surcroît, capable de *s'opposer* et de *dominer* les désirs non moraux – même très forts – pour accomplir une action bonne, juste ou vertueuse.

Nous retrouvons l'expression explicite de cette idée chez Rawls dans les passages suivants :

[P]uisque [l]es principes [de justice] sont dominants, le désir d'agir selon eux n'est satisfait que dans la mesure où il est également dominant vis-à-vis des autres désirs. C'est en agissant selon cette priorité que nous exprimons notre liberté à l'égard des contingences et des circonstances fortuites (RAWLS, 2009 , p. 615).

C'est un désir d'agir soi-même d'une certaine façon par-dessus tout, un effort qui contient en lui-même sa propre priorité (RAWLS, 2009 , p. 615).

Sans grand étonnement, l'argument qu'ils développent pour soutenir leur thèse est l'argument *moraliste* que nous sommes en train d'étudier. Dans sa forme générale, leur argument prend la forme suivante :

(1) Il existe pour les êtres humains des devoirs moraux catégoriques et dominants.	Prémisse normative
(2) ∴ il existe chez les êtres humains une motivation morale catégorique et dominante – i.e. les êtres humains peuvent être motivés à faire le bien, le juste et à devenir vertueux par la seule reconnaissance de la valeur morale d'une action, et ils sont capables de dominer leurs désirs non moraux pour y parvenir.	Conclusion descriptive

La notion de catégoricité, que nous avons déjà vue, résume l'idée d'indépendance à l'égard des désirs antérieurs à la reconnaissance de la valeur morale de l'action. La notion de *dominance*, de son côté, fait référence au fait, pour une considération, de *surpasser* toutes

les autres considérations. Par exemple, si le devoir moral de maximiser le bonheur du plus grand nombre est non seulement catégorique – i.e. valable indépendamment des désirs antérieurs de la personne – mais qu’il est en plus *dominant*, cela signifie que ce devoir moral ne peut jamais être annulé par le fait que, en le respectant, d’autres valeurs seraient mises en danger. Par exemple, si le devoir moral de subvenir aux besoins de sa famille est dominant par rapport à la réalisation, dans le monde, de valeurs esthétiques, alors il sera *toujours* moralement inacceptable d’abandonner sa famille pour réaliser un projet artistique – même si ce projet artistique possède des qualités esthétiques exceptionnelles et joue un rôle inestimable dans l’histoire de l’art, comme ce fut le cas pour le projet artistique de Gauguin.³⁵⁷ De la même manière, si le devoir de ne pas mentir est *dominant* relativement aux devoirs de *prudence*, alors il sera *toujours* interdit de mentir dans le but de s’épargner des conséquences personnelles désagréables, même lorsque dire la vérité est extrêmement coûteux pour soi.

Selon Emmanuel Kant et, à sa suite, John Rawls, les considérations de justice *dominent* toutes les autres considérations – qu’il s’agisse de considérations esthétiques, de bien-être personnel, d’efficacité économique, etc.³⁵⁸ Il n’est ainsi *jamais le cas* qu’une personne cesserait d’être soumise à un devoir de justice du simple fait que respecter ce devoir porterait atteinte à des valeurs non morales, telles que des valeurs esthétiques ou prudentielles, peu importe la gravité des atteintes. Or, selon eux, du fait qu’il existe de tels devoirs moraux catégoriques et dominants, il est possible de conclure que l’être humain dispose d’une motivation morale catégorique est dominante – i.e. capable de dominer

³⁵⁷ Le cas de Gauguin est pris en exemple par Bernard Williams pour caractériser la « chance morale » (WILLIAMS, 1981, pp. 20-39). En effet, Gauguin a abandonné sa famille pour se consacrer à la peinture. Il a néanmoins eu de la « chance », car son projet a finalement rencontré le succès artistique et l’influence historique que nous lui connaissons. La question est de savoir si ce succès artistique compense ou non le tort moral causé à la famille.

³⁵⁸ Le fait que les considérations morales *dominent* normativement les autres considérations n’implique pas que les considérations morales doivent constamment occuper l’esprit de l’individu, qu’elles doivent *dominer* les autres pensées – dans un sens psychologique. Comme le montre Scheffler, il suffit que la clause contrefactuelle suivante soit satisfaite : « Bien qu’elle réponde à des considérations qu’elle ne se représente pas en des termes spécifiquement moraux, une personne peut néanmoins satisfaire les deux conditions contrefactuelles suivantes : en premier lieu, elle n’aurait pas agi comme elle l’a fait si elle avait pensé que ce qu’elle faisait était mal ; en second lieu, si une considération lui était venue à l’esprit, qui semblait militer contre le caractère moralement permisible de son action, elle n’aurait pas agi ainsi. » [“[A]lthough acting in response to some feature of a situation that he does not represent to himself in overtly moral terms, a person may nevertheless satisfy two related counterfactual conditions: first, that he would not have acted as he did if he had believed that doing so was wrong; second, that if a consideration had come to his attention that seemed to militate against the moral permissibility of his act, he would not have performed the act.” (SCHEFFLER, 1992, p. 32).]

toutes les motivations contraires. Autrement dit, l'être humain est capable de surmonter tous ses désirs dans le but d'accomplir l'action morale ; sa motivation morale est potentiellement toute puissante. L'être humain est capable de se *libérer* de ses désirs non moraux pour accomplir son devoir moral.

1.3. LA VERSION MORALISTE DE « DEVOIR-IMPLIQUE-POUVOIR »

Bien entendu, l'argument proposé jusqu'à présent est toujours *logiquement invalide*. En effet, la transition entre la prémisse normative (1) et les conclusions descriptives (2) et (3) ne respecte pas les règles élémentaires de la logique. En effet, la logique ne permet toujours pas d'opérer une transition entre deux énoncés du type suivant :

(1) *a* est *M*

(2) $\therefore a$ est *N*

Pour justifier la transition de (1) à (2), il est nécessaire donc de faire appel à la *sémantique* de la notion de devoir. Cette sémantique devrait nous permettre d'établir une transition conceptuelle entre, d'un côté, les énoncés *normatifs* et, de l'autre, les énoncés *descriptifs* au sujet de la motivation humaine. Or, le principe sémantique qui permet d'établir cette transition est, à nouveau, comme dans le raisonnement naturaliste, le principe DIP-motivation.

Précisons l'élément suivant : il est vrai que le principe DIP-motivation est employé aussi bien dans le raisonnement naturaliste que dans le raisonnement moraliste. Toutefois, il est employé de deux manières différentes dans les deux types de raisonnement. En effet, dans le raisonnement naturaliste, il est employé pour construire un *modus tollens* sur la base de la *contraposée* de DIP. A l'inverse, dans le raisonnement moraliste, DIP est employé sous sa forme *non-contraposée* afin de construire un *modus ponens*. Ainsi, accepter DIP ne revient pas à accepter son usage *dans les deux directions*. Par exemple, pour les moralistes dans la lignée d'Emmanuel Kant, le seul usage légitime de DIP est l'usage *matériel* selon lequel « si dois, tu peux ! ». Dans notre travail, nous considérons également que, lorsque nous discutons des motivations, seul l'usage moraliste de DIP est légitime : « si tu dois, tu peux être motivé ! ».

Pour rappel, dans son usage par les empiristes dans le raisonnement naturaliste, DIP-motivation permettait de construire le raisonnement (sémantiquement valide) suivante :

(1) Les êtres humains ne peuvent pas être motivés à faire X.	Prémisse descriptive
(2) Si les êtres humains ne peuvent pas être motivés à faire X, alors les êtres humains n'ont pas le devoir moral de faire X	Principe analytique DIP-motivation.
(3) ∴ Les êtres humains n'ont pas le devoir moral de faire X.	Conclusion normative

Les naturalistes partent d'une prémisse descriptive *négative* pour obtenir, par le biais de la contraposée du principe DIP-motivation, une conclusion normative elle aussi négative.

Dans l'usage *moraliste* qui nous intéresse, le principe DIP-motivation permet de construire la déduction (sémantiquement valide) suivante :

(3) Les êtres humains ont le devoir moral catégorique (et dominant) de faire X.	Prémisse normative
(4) Si les êtres humains ont le devoir moral de faire X, alors les êtres humains peuvent être motivés à faire X.	Principe analytique DIP-motivation
(5) ∴ Les êtres humains sont capables de motivation morale catégorique (et dominante) pour faire X.	Conclusion descriptive

Dans cet exemple, nous partons d'une prémisse normative *positive* et obtenons, par le biais de DIP-motivation, une conclusion descriptive elle aussi positive. Cet argument est sémantiquement valide.

Dans les deux types de raisonnements, naturaliste et moraliste, la validité de l'argument dépend du fait que nous acceptons la thèse suivante : dans le discours moral ordinaire, nous employons une sémantique du *devoir* qui est telle qu'elle implique *la possibilité* – et plus particulièrement la possibilité motivationnelle. Or, nous avons vu qu'il existe différentes objections contre cette défense *sémantique* de DIP, notamment le fait qu'elle rend *contradictoire* l'idée de conflits moraux insolubles et qu'elle rend *impossible* la faiblesse de la volonté. En outre, l'interprétation sémantique de DIP ne correspond pas à la

forme historique du débat. Pour soutenir l'usage de DIP-motivation dans l'argument *moraliste*, il nous faut donc répondre à ces différentes inquiétudes.

1.3.1. LE PROBLÈME DES DILEMMES MORAUX

Selon la première objection, si le principe DIP est vrai, alors cela implique que les dilemmes moraux insolubles sont *impossibles* ; il est conceptuellement impossible qu'existe des situations de dilemmes moraux insolubles. Or, nous l'avons vu, de telles situations *existent* bel et bien dans la réalité – ainsi que dans certaines théories morales. La preuve de l'existence des dilemmes moraux insolubles est notamment l'expérience que nous avons de *sentiments* moraux résiduels, tels que le remords, la culpabilité, ou le regret, que nous ressentons lorsque nous échouons à respecter l'une de nos obligations morales (BÉLANGER, 2011, p. 15).

Première réponse

La première réponse à l'objection consiste à montrer que, en soi, DIP est compatible avec l'existence de dilemmes moraux du type : S a le devoir de faire A et S a le devoir de faire non-A.³⁵⁹

(1) DIP est compatible avec : $O(A) \wedge O(\neg A)$

En effet, DIP n'implique aucune contradiction dans les situations de dilemme du type « S a le devoir de faire A » et « S a le devoir de faire non-A » ; en vertu de DIP, « S a le devoir de faire A » implique « S peut faire A » et « S a le devoir de faire non-A » implique « S peut faire non-A ». Or, « S peut faire A » et « S peut faire non-A » ne sont pas contradictoires.³⁶⁰ Ce n'est que si nous commettons l'erreur – avec Ruwen Ogien (2001, p. 19) – de considérer que « S a le devoir de faire non-A » implique « S ne peut pas faire A » que DIP implique une contradiction.

³⁵⁹ Notons d'abord que le devoir de faire A et le devoir de faire non-A ne sont aucunement contradictoires en eux-mêmes. « S doit faire A » n'est pas contradictoire avec « S doit ne pas faire A ». La seule proposition avec laquelle « S doit faire A » est contradictoire est la proposition « il est faux que S doit faire A ». Plus formellement :

$O(A)$ et $\neg O(A)$ sont contradictoires
 $O(A)$ et $O(\neg A)$ ne sont pas contradictoires

³⁶⁰ Autrement dit, plus formellement :

$\Diamond(A) \wedge \neg\Diamond(A)$ sont contradictoires
 $\Diamond(A) \wedge \Diamond(\neg A)$ ne sont pas contradictoires

De ce fait, il n'est pas vrai que DIP ne permet pas de rendre compte de l'existence de dilemmes moraux. Aussi longtemps que la situation conflictuelle est exprimée par *deux* obligations distinctes, DIP n'implique pas que le conflit moral soit contradictoire. DIP est tout à fait compatible avec l'existence de dilemmes moraux ayant cette structure logique.

Par contre, il est vrai que DIP est incompatible avec le devoir de « faire A et non-A ».

(2) DIP est incompatible avec : $O(A \wedge \neg A)$

En effet, le devoir de faire « A et non-A » implique, en vertu de DIP, la possibilité de faire A et non-A. Si DIP est accepté et qu'il existe un devoir de faire « A et non-A », alors il faudrait admettre qu'il existe un monde possible dans lequel A et non-A sont toutes les deux réalisées, ce qui est contradictoire.

Cependant, nous ne sommes aucunement contraints de considérer qu'il existe des situations de dilemme moral ayant cette structure. En effet, la conjonction de deux devoirs (faire A et faire non-A) ne donne lieu à une obligation unique de faire une paire d'actions (faire A et non-A) que si nous acceptons le « principe de distribution déontique » (SAYRE-MCCORD, 1986, p. 181), ou, au minimum, le « principe déontique d'agglomération » (BÉLANGER, 2011, p. 22). Ces principes sont les suivants :

Principe de distribution déontique (PDD) : $[O(A) \wedge O(B)] \equiv O(A \wedge B)$

Principe déontique d'agglomération (PDA) : $[O(A) \wedge O(B)] \rightarrow O(A \wedge B)$

Etant donné que le principe conditionnel PDA est suffisant pour générer les contradictions qui nous intéressent, nous laissons de côté le principe biconditionnel PDD. Ainsi, nous retenons que le principe DIP n'implique une contradiction, en cas de dilemme moral, que s'il est combiné au principe PDA (au minimum) :

(3) « $O(A) \wedge O(\neg A)$ » implique « $O(A \wedge \neg A)$ » uniquement si nous acceptons (au minimum) le principe PDA

Or,

(4) Il existe des raisons de rejeter le principe PDA

La principale raison de rejeter PDA est que, sous certaines conditions, PDA implique que toute action est obligatoire. A supposer, en effet, que les conséquences *logiques* des actions obligatoires sont elles-mêmes obligatoires, alors les conséquences logiques de la proposition « A et non-A » doivent être considérées comme obligatoires. Or, de l'énoncé « A et non-A », qui est une contradiction, il est possible de tirer n'importe quelle conséquence logique (SAYRE-MCCORD, 1986, p. 188). Si nous acceptons PDA, n'importe quelle action (désignée par B) devient, dès lors, obligatoire.³⁶¹ Par exemple, s'il est obligatoire de sauver l'enfant qui se noie et de ne pas le sauver, alors il est obligatoire d'aller manger une glace.³⁶² Cette conséquence est absurde.

Deuxième réponse

Une autre manière de sauver DIP consiste à accepter le principe déontique d'agglomération (PDA) tout en refusant d'admettre l'existence de dilemmes moraux insolubles. Cela signifie que, bien que DIP et PDA mènent à une contradiction en cas de dilemme moral insoluble, cela n'est pas problématique, car de tels dilemmes n'existent pas.

- (1) La combinaison de DIP et PDA génère une contradiction lors de dilemmes moraux insolubles.
- (2) Il n'existe pas de dilemmes moraux insolubles
- (3) $\therefore$ La combinaison de DIP et PDA ne génère pas de contradiction.

La tactique consiste ainsi, cette fois, à concéder à l'objecteur que DIP (avec PDA) est incompatible avec l'existence de dilemmes moraux insolubles, mais à affirmer que cela n'est pas un problème, car les dilemmes moraux insolubles n'existent pas.

En premier lieu, il s'agit de montrer que les cas de dilemmes moraux du type $O(A) \wedge O(\neg A)$ n'existent pas : soit (1) parce que, dans les dilemmes moraux de ce type, les notions d'obligation sont *indexées*, soit (2) parce que les cas de dilemmes de ce type se réduisent à des cas de dilemmes moraux du type $O(A) \wedge O(B)$, lorsque A et B désignent deux actions qui

³⁶¹ Etant donné que « A et non-A » implique « B », alors le devoir de faire « A et non-A » implique le devoir de faire B. Plus formellement :

$[(A \wedge \neg A) \rightarrow B] \rightarrow [O(A \wedge \neg A) \rightarrow O(B)]$

³⁶² Marco Bélanger (2011, p. 138) propose un autre exemple : s'il est obligatoire de respecter un étudiant et de ne pas le respecter, alors il s'ensuit qu'il est obligatoire de le tuer.

ne peuvent pas être accomplies simultanément – et il faudra ensuite montrer que ces cas de dilemmes moraux ne posent pas de difficulté pour DIP.

1) *L'indexation*

La première option pour nier l'existence de dilemmes moraux du type $O(A) \wedge O(\neg A)$ consiste à considérer que la signification de la notion d'obligation est différente lorsqu'elle porte sur (A) et lorsqu'elle porte sur ($\neg A$). Autrement dit, la notion d'obligation est toujours *indexée*.

Pour illustrer ce point, considérons l'exemple classique suivant : un enfant est en train de se noyer. Sur la plage, un bon nageur s'apprête à aller le sauver. Toutefois, il s'avère que l'enfant en question est un futur dictateur – il est le fils du dictateur actuel et tout laisse croire qu'il sera aussi cruel que son père. Il s'avère même que cet enfant, dans le futur, commettra un atroce génocide. Il semble donc que, d'un côté, le touriste a le devoir moral de sauver l'enfant – en tant qu'il s'agit d'un enfant qui se noie. De l'autre, il semble que le touriste a aussi le devoir moral de ne pas le sauver – en tant qu'il s'agit d'un futur génocidaire qui se noie. De prime abord, il semble donc que le touriste fait face à un dilemme moral insoluble.

Pourtant, il semble évident que les deux devoirs contradictoires ne prennent pas leur source dans le même principe ; le devoir de sauver l'enfant découle du devoir d'aider une personne en danger. De son côté, le devoir de ne pas sauver le futur génocidaire dérive du devoir de faire l'action qui minimise la quantité totale de souffrance dans le monde. Notre touriste a donc l'obligation de sauver l'enfant en vertu d'un principe d'entraide et de ne pas le sauver en vertu du principe de minimisation de la souffrance dans le monde. Les principes dont dérivent ces deux devoirs (faire A et faire non-A) sont différents.

Cela signifie que, formellement, la manière adéquate de décrire la situation de dilemme moral n'est pas :

$O(A) \wedge O(\neg A)$ mais est plutôt

$O_{\text{-Entraide}}(A) \wedge O_{\text{-Minisouffrance}}(\neg A)$

Or, si nous acceptons cette réécriture, le « principe déontique d'agglomération » (PDA) ne s'applique plus. En effet,

$$O_E(A) \wedge O_M(\neg A) \text{ n'implique pas } O(A \wedge \neg A)$$

En bref, étant donné que, dans les situations de dilemmes du type « $O(A) \wedge O(\neg A)$ », la notion d'obligation est toujours *indexée*, alors PDA ne s'applique jamais dans ces circonstances et DIP ne mène plus une contradiction.

2) *L'impossibilité de faire A et B simultanément*

La seconde option consiste à traiter les cas de dilemmes moraux « $O(A) \wedge O(\neg A)$ » comme étant simplement des cas de dilemmes moraux du type « $O(A) \wedge O(B)$ », dans lesquels il n'est pas possible de faire à la fois A et B. Pour reprendre le cas du touriste de l'exemple précédent, nous pouvons décrire la situation de dilemme à laquelle il est confronté de la manière suivante : il a le devoir moral « de venir en aide à une personne en danger » (A) et le devoir moral « de minimiser la souffrance dans le monde » (B), mais il ne peut pas satisfaire ces deux actions simultanément. Le dilemme « $O(A) \wedge O(\neg A)$ » est ainsi un cas particulier d'un dilemme moral du type « $O(A) \wedge O(B)$ », quand il n'est pas possible de faire à la fois A et B.

Avec cette seconde solution, il semble toutefois, de prime abord, que nous n'avons pas complètement résolu le problème de la compatibilité logique de DIP avec l'existence de dilemmes moraux. En effet, l'existence de dilemmes moraux insolubles du type « $O(A)$ et $O(B)$ », n'est pas non plus compatible avec DIP, comme le montre le raisonnement suivant (BÉLANGER, 2011):

- 1) S doit faire A
- 2) S doit faire B
- 3) S ne peut pas faire A et B

Des prémisses (1) et (2) et du principe déontologique d'agglomération, il s'ensuit que :

- 4) S doit faire A et B.

De la prémisse (4) et du principe « devoir implique pouvoir », il s'ensuit que

- 5) S peut faire A et B

De la prémisse (3) et de la prémisse (5), il s'ensuit que

- 6) S ne peut pas faire A et B et S peut faire A et B

La conclusion de ce raisonnement est contradictoire. Cela montre que DIP est incompatible avec l'existence de dilemmes moraux portant sur deux actions qui ne peuvent pas être accomplies toutes les deux.

Afin d'éviter cette conclusion, il faut pouvoir montrer que ces cas de dilemmes moraux sont toujours uniquement *apparents*, et ne sont donc jamais *insolubles*.

La solution la plus simple dans ce sens consiste à montrer que, dans de nombreux cas, l'une des obligations est plus forte ou importante que l'autre. Si tel est le cas, si l'obligation *pro tanto* pour l'individu S de faire A est plus forte que l'obligation *pro tanto* pour S de faire B, alors S a l'obligation *toutes choses considérées* uniquement de faire A, sans devoir faire B. Ainsi, dans l'exemple précédent, si l'obligation morale de minimiser la souffrance dans le monde est plus importante que l'obligation morale de venir en aide à une personne en danger, alors notre touriste a l'obligation morale *toutes choses considérées* laisser se noyer le futur génocidaire.

La difficulté provient du fait qu'il n'est pas toujours le cas que l'une des obligations est plus importante que l'autre. En effet, cela n'est pas le cas dans l'exemple des jumeaux qui se noient à équidistance du rivage. Dans cet exemple, l'obligation de sauver l'enfant de gauche (A) et l'obligation de sauver l'enfant de droite (B) dérivent du même principe et sont d'égale

importance. Il n'est donc ni possible d'*indexer* la notion d'obligation, ni possible de privilégier une obligation plutôt que l'autre.

La solution est alors la suivante : lorsque deux obligations conflictuelles sont de même origine et de même importance (et uniquement dans ces cas précis), l'obligation résultante n'est *pas* une conjonction d'obligations du type « $O(A \wedge B)$ », mais une obligation *disjonctive* du type « $O(A \vee B)$ ». Plus formellement :

$[O(A) \wedge O(B)] \rightarrow O(A \vee B)$ lorsque $O(A)$ et $O(B)$ sont d'égale importance et proviennent de la même source et que A et B sont incompatibles.

Dans l'exemple qui nous intéresse, cela implique que, si l'obligation de sauver l'enfant de gauche (A) est de même origine et de même importance que l'obligation de sauver l'enfant de droite (B), alors le touriste a l'obligation morale de sauver l'un des deux enfants. Il a l'obligation morale de sauver soit l'enfant de gauche, soit l'enfant de droite. Mais il n'a pas l'obligation morale de sauver l'enfant de gauche *et* l'enfant de droite. En sauvant l'un ou l'autre, il satisfait à son obligation morale. Il ne commet donc aucune faute morale en ne sauvant pas le second enfant. Si le touriste éprouve néanmoins, par la suite, un sentiment de culpabilité, celui-ci n'est pas correct ou adéquat (*fitting*), i.e. il ne représente pas adéquatement son objet.

Pour intégrer cette exception au principe PDA dans la formulation du principe lui-même, nous pouvons reformuler le principe de la manière suivante (BÉLANGER, 2011, p. 139):

Principe PDA révisé $\Diamond(A \wedge B) \rightarrow [(O(A) \wedge O(B)) \rightarrow O(A \wedge B)]$

Cette solution peut sembler parfaitement *ad hoc*. En effet, nous affirmons que le principe déontique d'agglomération est valide en toute circonstance, *sauf* précisément lorsque les deux obligations conflictuelles sont d'égale importance. Cela revient à dire que PDA est valide, sauf *en cas de dilemme moral insoluble*. Or, il semble parfaitement gratuit de faire une telle affirmation dans un argument dont le but est *précisément* de contester l'existence de dilemmes moraux insolubles, dans le but de préserver PDA et DIP.

Il est cependant possible de justifier l'introduction d'une telle exception dans la formulation de PDA. La raison est la suivante : les situations *d'exception* sont clairement

identifiées. Il s'agit, en effet, des situations dans lesquelles les deux actions sur lesquelles portent les obligations d'égale importance, sont incompatibles. De ce fait, il ne s'agit pas véritablement d'une exception *ad hoc*, puisqu'elle porte sur *un ensemble* clairement identifié de situations.

L'élément important ici est de différencier les violations du principe qui forment des *contre-exemples* de celles qui forment des *exceptions*. Si nous identifions un *contre-exemple*, alors cela a pour conséquence d'invalider le principe. Par contre, si nous identifions une exception, cela n'invalidé pas le principe, mais permet plutôt de spécifier son champ d'application. Or, lorsqu'un ensemble de situations contredisant un principe partagent toutes une même caractéristique – dans notre cas, celle de porter sur des actions incompatibles – nous pouvons considérer ces situations comme des *exceptions* plutôt que comme étant des contre-exemples. Cela s'explique du fait qu'il est possible de formuler une *règle* concernant les cas où le principe ne s'applique pas. En conclusion, étant donné que les situations dans lesquelles PDA ne s'appliquent pas sont *clairement identifiées*, elles représentent des exceptions plutôt que des contre-exemples, et elles ne peuvent pas servir à invalider le principe.

En outre, il existe *d'autres domaines* dans lesquels le principe d'agglomération est *en général* valable, à l'exception des situations dans lesquelles les contenus sont incompatibles. C'est le cas, par exemple, pour les exigences de prudence (BÉLANGER, 2011, p. 129) :

S'il est prudent de faire A et s'il est prudent de faire B, il est en général prudent de faire A et B, *sauf* lorsque A et B sont des actions incompatibles.

Par exemple, s'il est dans mon intérêt de réviser mes examens et s'il est dans mon intérêt de ne pas fumer, alors il est dans mon intérêt à la fois de réviser mes examens et de ne pas fumer. Par contre, il est dans mon intérêt de réviser mes examens et il est dans mon intérêt d'aller me promener en montagne (pour ma santé), mais il n'est pas dans mon intérêt *de réviser et de marcher*. Pourquoi ? Parce qu'il n'est pas possible – ou du moins, pas prudent – de réviser et de marcher simultanément. L'agglomération des obligations de prudence est donc valide, *sauf* lorsqu'il n'est pas possible de faire les deux actions simultanément.

Un autre domaine où nous utilisons, dans le langage, le principe d'agglomération *avec des exceptions bien identifiées*, est celui des croyances.

Si S croit que p et S croit que q , alors S croit que p et q , sauf lorsque p et q sont contradictoires.

En effet, lorsque nous croyons que p et nous croyons que q , nous disons en général que nous croyons que p et q . Nous appliquons le principe d'agglomération. Toutefois, nous n'appliquons plus ce principe lorsque les croyances sont contradictoires : même si nous croyons que p et nous croyons que q , nous ne croyons pas, dans ces cas précis, que p et q . Admettons, par exemple, que pour différentes raisons, Jean croit que Pauline est coupable du meurtre de Françoise. Pour d'autres raisons, Jean croit également que Pauline n'est pas coupable du meurtre de Françoise. Néanmoins, Jean ne croit pas que Pauline est coupable et qu'elle n'est pas coupable.³⁶³ Au contraire, il sait qu'elle est *soit* coupable, *soit* non coupable.

Pour montrer la différence entre les situations *d'exception* et celles de *contre-exemples*, il convient d'identifier également des domaines dans lesquels il n'est pas possible de considérer les violations de PDA comme étant des exceptions, du fait que les violations ne partagent pas une caractéristique commune. C'est le cas notamment dans le domaine des désirs. En effet, il n'est pas le cas que le désir de faire A et le désir de faire B donnent lieu à un désir de faire A et B, *sauf* lorsque A et B sont incompatibles. En effet, il existe d'autres circonstances (que l'incompatibilité de A et B) dans lesquelles le principe d'agglomération des désirs ne s'applique pas. Ainsi, il est possible que je désire passer mon pyjama et que je désire sortir danser, mais il n'est pas le cas que je désire passer mon pyjama *et* sortir danser. Bien que les deux actions ne soient pas incompatibles, le principe d'agglomération des désirs ne s'applique pas. Cela montre que l'exemple présenté constitue bien un *contre-exemple* plutôt qu'une *exception* au principe d'agglomération.

³⁶³ Marco Bélanger (2011, p. 133) emploie la notion épistémique de *connaissance* pour montrer l'existence d'exceptions au principe d'agglomération dans des contextes épistémiques. Cela nous semble inapproprié, étant donné la nature factive de la connaissance : si je *sais* que p , cela implique que p est le cas. Et si je sais que q , cela implique que q est le cas. Or, il n'est pas possible que q soit le cas et que p soit le cas lorsque ces deux propositions se contredisent – ce qui est précisément censé être le cas dans l'exemple. Nous préférons donc utiliser le cas des croyances.

En résumé, il est possible de distinguer les circonstances dans lesquelles le principe PDA fait face à des *contre-exemples* des circonstances dans lesquelles il rencontre des *exceptions*. Or, le principe PDA, lorsqu'il est appliqué aux obligations, fait face à des *exceptions*, mais pas à des *contre-exemples*. C'est pourquoi nous pouvons légitimement accepter que PDA est un principe valide, *sauf* lorsque les actions A et B sont incompatibles.

En conclusion, nous avons avancé deux arguments pour ne pas rejeter l'analyse conceptuelle de DIP. Le premier consiste à montrer que DIP est compatible avec l'existence de dilemmes moraux lorsque nous n'admettons pas le « principe déontique d'agglomération ». Le second consiste à montrer que l'incompatibilité de DIP avec l'existence de dilemmes moraux insolubles (en admettant PDA) n'est pas problématique, car de tels dilemmes n'existent pas en fait : soit les notions d'obligation sont indexées, soit l'une des obligations est plus forte que l'autre, soit les deux obligations donnent lieu à une *disjonction* d'obligations plutôt qu'à une *conjonction* d'obligation, lorsque les obligations conflictuelles sont de même origine et d'égale importance.

La structure générale de nos deux arguments était donc la suivante :

Premier argument

- (1) DIP est compatible le fait d'avoir l'obligation de faire A et l'obligation de faire non-A.
- (2) DIP n'est pas compatible avec le fait d'avoir l'obligation de faire A et de faire non-A.
- (3) L'obligation de faire A et l'obligation de faire non-A ne donnent lieu à une obligation de faire A et non-A que si nous admettons le « principe déontique d'agglomération ».

MAIS :

- (4) Il existe des raisons de rejeter le « principe déontique d'agglomération », notamment le fait que, si nous l'acceptons, n'importe quelle action devient obligatoire en cas de dilemme moral insoluble.
- (5) ∴ DIP est compatible avec l'existence de dilemmes moraux insolubles.

Deuxième argument

(1) Les principes PDA et DIP mènent à une contradiction en cas de dilemme moral insoluble.

MAIS :

(2) Il n'existe pas de dilemme moral *insoluble*.

- a. Soit la notion d'obligation est indexée.
- b. Soit l'une des obligations est plus forte que l'autre.
- c. Soit la conjonction de deux obligations de même valeur portant sur des actions distinctes donne lieu à une obligation disjonctive.

(3) ∴ DIP et PDA peuvent être acceptés simultanément.

1.3.2. LE PROBLÈME DE LA FAIBLESSE DE LA VOLONTÉ

Selon le second reproche, DIP-motivation rend impossible la faiblesse de la volonté. En effet, si DIP est vrai, alors il est impossible *à la fois* d'avoir le devoir de faire X *et* de ne pas pouvoir être motivé à faire X. Autrement dit, les propositions [...]

(A) « S a le devoir moral de faire X »

et

(B) « S ne peut pas être motivé à faire X »

[...] ne peuvent pas être vraies simultanément. Or, comme le suppose l'objection, il arrive qu'elles *soient vraies* simultanément. En effet, il existe de nombreuses situations dans lesquelles nous avons un devoir de faire X mais nous ne pouvons pas être motivés à faire X – dans les termes d'Ogden, nous souffrons d'une « faiblesse insurmontable de la volonté ». Dans ces situations, nous ne considérons pas, en général, que notre faiblesse de volonté fasse disparaître notre devoir. Par exemple, le fait que nous ne pouvons pas nous motiver à rendre visite à notre grand-mère, comme nous l'avions promis, n'implique pas que nous n'avons pas l'obligation de respecter notre promesse de lui rendre visite. Cependant, si DIP est accepté, il faudrait considérer que les propositions (A) et (B) sont incompatibles, qu'elles ne peuvent pas être toutes les deux vraies. Il faudrait considérer que la faiblesse de la volonté *fait disparaître* le devoir moral. Selon l'objection, cela prouve que, dans le langage moral ordinaire, nous n'acceptons pas DIP dans son interprétation conceptuelle.

Notre réponse est simple : si DIP est conceptuellement vrai – ce que nous pensons être le cas – il est vrai que (A) et (B) ne peuvent pas être vraies ensemble. Mais au lieu de considérer que, si (B) est vraie, alors (A) est fausse, il faut plutôt considérer que, si (A) est vraie, alors (B) fausse. Au lieu de considérer que, si S ne peut pas être motivé à faire X, alors S n’a pas le devoir de faire X, il faut considérer que, si S a le devoir de faire X, alors il n’est pas le cas que S ne peut pas être motivé à faire X. Autrement dit, l’existence d’un devoir *exclut* qu’une personne ne puisse pas être motivée à faire X.

Nos objecteurs soutiendront sûrement qu’il est *totalelement absurde* de soutenir que nous ne pouvons pas souffrir de faiblesse de volonté dans l’accomplissement de nos devoirs.³⁶⁴ En effet, le fait de n’être pas motivé à faire ce que nous jugeons comme étant pourtant la chose à faire – toutes choses considérées – est un phénomène extrêmement courant.

A cela nous répondrons que ces situations existent effectivement ; il existe de nombreuses situations dans lesquelles

(A) « S a le devoir moral de faire X »

et

(C) « S n’est pas motivé à faire X »

Cependant, la compatibilité de (A) et (C) est bien différente de la compatibilité de (A) et (B). En effet, (B) est la thèse selon laquelle « S ne peut pas être motivé à faire X » et non pas la thèse (C) selon laquelle « S *n’est pas* motivé à faire B ». Autrement dit, (C) exprime une version *faible* de faiblesse de volonté – i.e. le simple fait de *ne pas être motivé*, alors que (B) exprime une version *forte* de la faiblesse de volonté – i.e. le fait de ne pas pouvoir être motivé. Or, la littérature contemporaine sur ce thème nous semble admettre uniquement la thèse faible (C) (MELE, 2009; STROUD & SVIRSKY, 2019). Le principe DIP est donc tout à fait

³⁶⁴ Contre cette affirmation, notons déjà que plusieurs auteurs admettent tout simplement que la faiblesse de la volonté est impossible (voir les références dans STROUD, et al., 2019). Si ces auteurs ont raison, alors le fait que DIP implique que la faiblesse de la volonté est impossible ne pose pas de difficulté pour notre théorie.

compatible avec la faiblesse de la volonté – comme un simple *défaut* de motivation, possiblement surmontable, plutôt que comme une *incapacité* à être motivé.³⁶⁵

Certes, il existe également de *véritables incapacités motivationnelles*. Par exemple, les personnes souffrant d'addictions ou de comportement compulsif ou de dépression, sont souvent bien incapables de se motiver à agir comme elles jugent qu'elles devraient agir. Toutefois, de tels troubles mentaux ne constituent pas des exemples de *faiblesse de volonté*, mais des exemples de *maladie*, de dysfonctionnement mental dont la correction ne dépend pas de la seule volonté du sujet. C'est d'ailleurs pour cette raison que, dans la littérature contemporaine, les cas de déficiences maladiques de motivation – l'addiction, la dépression, etc. – ne sont pas considérées comme des cas de faiblesse de la volonté, mais comme des cas de *contrainte*.

Une condition nécessaire pour qu'une action soit acratique [i.e. accomplie par faiblesse de volonté], dans la conception standard, est qu'elle ne doit pas être contrainte. Un héroïnomanie qui juge qu'il serait meilleur pour lui de ne pas utiliser la drogue maintenant, mais qui l'utilise toutefois maintenant parce qu'il a un désir irrésistible de le faire ne compte pas comme faisant un usage acratique de la drogue.³⁶⁶

Or, l'objection que nous avons considérée portait uniquement sur la faiblesse de la volonté et non pas sur les dysfonctionnements maladiques de la motivation. Il suffit donc pour notre argument que nous ayons montré que, même si DIP n'est pas compatible avec une version *forte* de la faiblesse de la volonté (B), il reste néanmoins compatible avec la version *faible* de la faiblesse de la volonté (C), version qui correspond à la caractérisation de la faiblesse de la volonté dans la littérature contemporaine. Ce n'est que parce que Ruwen

³⁶⁵ Nous pouvons même admettre une *incapacité* ponctuelle d'être motivé à agir. Il suffit – comme l'admettent au moins certains théoriciens de la faiblesse de la volonté – que cette incapacité ponctuelle puisse être surmontée par l'exercice, par la pratique, pour que nous considérions cette incapacité comme *surmontable*. Si l'incapacité ponctuelle peut être surmontée, et que ce qui empêche l'accomplissement du devoir, c'est précisément cette incapacité ponctuelle, alors le devoir ne disparaît pas ; au contraire, la personne souffrant d'incapacité ponctuelle de volonté a dès lors le devoir de pratiquer le contrôle de soi pour parvenir à raffermir sa volonté. Le caractère ponctuel ou durable de la faiblesse de la volonté est donc pertinent – contrairement à ce que semble soutenir Ogien (2001, p. 20). Nous pourrions construire, sur cette base, une « éthique pour agents faibles », pour reprendre les termes de Laurent Jaffro (2016), consistant à travailler pour raffermir sa volonté.

³⁶⁶ "One necessary condition for an action's being akratic, on a standard conception, is that it is uncompelled. A heroin addict who judges it best not to use the drug now but uses it now because he has an irresistible desire to do so does not count as akratically using the drug." (MELE, 2009, p. 403)

Ogien utilise une conception étonnamment forte de l'idée de *faiblesse de volonté* qu'il peut développer une objection contre DIP.

1.3.3. QUE FAIRE DE LUTHER ?

Un troisième problème que soulevait l'interprétation conceptuelle de DIP, et que ne soulevait pas l'interprétation *morale* du principe, était l'incompatibilité du principe avec la doctrine morale de Martin Luther. En effet, nous avons vu que le réformateur développe une position morale – ou théologique – qui contredit le principe DIP.³⁶⁷ Nous avons vu, également, que ses opposants soulevaient des arguments *moraux* contre la position de Luther ; si Luther a raison, Dieu apparaît comme un monstre cruel et capricieux. En effet, étant donné qu'il est *cruel* d'exiger l'impossible, et que Dieu exige l'impossible, alors Dieu est cruel – dans la doctrine luthérienne, selon ses objecteurs. Dans notre argument, le fait que l'opposition théologique portait sur le caractère *moral* ou *immoral* de la doctrine luthérienne devait servir à favoriser l'interprétation *morale* de DIP – par opposition à l'interprétation *conceptuelle* que nous favorisons.

Nous avons dit que le débat semblait admettre une interprétation *morale* de DIP, car les opposants de Luther employaient des arguments moraux. Cependant, si nous regardons de l'autre côté du débat, il n'est plus évident que le débat concernait exclusivement le caractère immoral des exigences impossibles à accomplir. En effet, dans les écrits de Luther, rapportés par Bernard Quilliet, l'argument du réformateur consistait à pointer le caractère fallacieux ou logiquement invalide de ses adversaires, à savoir les théologiens « humanistes » qui considéraient que la vie morale doit être réalisable pour les êtres humains. Selon Luther, ses opposants dérivent un énoncé au sujet de ce qui est *possible* à partir d'un énoncé au sujet de ce qui *doit être*, ce que la logique interdit.

Les maîtres d'école élémentaire et leurs élèves [...] le savent : par les verbes au mode impératif, on ne peut signifier rien de plus que ce qui doit être fait.

Comment donc se fait-il que vous autres, théologiens, vous dites des inepties [...],

³⁶⁷ Luther affirme avec force son refus de DIP contre Erasme dans plusieurs passages, dont celui-ci, particulièrement éloquent : « Tes comparaisons et tes maximes sont jolies et distinguées, mais quand tu plaides sur des sujets sacrés, tu les emploies de façon puérile et tout de travers : tu rampes sur le sol et tu ne penses rien qui soi au-dessus de la capacité humaine. Ce ne sont pas, en effet, des choses puériles, ni civiles, ni humaines, que Dieu, mais des choses divines, et elles excèdent la capacité humaine ». [Citation de Luther par Quilliet (2007, p. 260)]

si bien qu'à peine [après ?] avoir saisi un verbe à l'impératif, vous en déduisez l'indicatif, comme si d'emblée, dès qu'une chose est commandée, elle serait, aussi, nécessairement [... ?] possible à faire ?³⁶⁸

Or, comme nous l'avons montré dans ce travail, le raisonnement partant d'un énoncé normatif « S doit » pour conclure un énoncé descriptif « S peut » se fonde sur le principe *sémantique* « devoir-implique-pouvoir ». Aucun partisan de ce raisonnement ne prétend qu'il s'agit d'un raisonnement valide *en vertu uniquement* des composantes logiques du raisonnement. Ainsi, pour contredire Luther, il suffit de montrer que le raisonnement est valide, *non pas en vertu* des composantes logiques uniquement, mais en vertu de la sémantique de la notion de « devoir ». ³⁶⁹

En conclusion, le débat entre Luther et ses opposants, notamment Erasme, ne favorise aucune interprétation particulière de DIP. D'un côté, Luther fait appel à des considérations purement logiques concernant le passage fallacieux du devoir à la possibilité de l'accomplir. La seule manière de répondre à un tel argument serait de montrer que DIP est un principe *analytiquement vrai*. De l'autre côté, ses opposants font appel à des considérations *morales*. Le débat porte donc à la fois sur des considérations logiques et sur des considérations morales. Ce n'est que lorsque nous ignorons les arguments de l'une des parties du débat que nous parvenons à la conclusion que le débat théologique entre Luther et Erasme concernait uniquement le caractère moral/immoral des exigences divines impossibles à accomplir. Autrement dit, ce n'est qu'en ignorant une partie des arguments que l'on peut conclure que

³⁶⁸ Citation de Luther dans le livre de Bernard Quilliet (2007, p. 266)

³⁶⁹ Notons également que même Luther admet *au moins en partie* le principe DIP ; en effet, il admet que les exigences morales auxquelles sont soumis les êtres humains sont des exigences adaptées aux êtres humains, dans le sens qu'elles s'appliquent spécifiquement à des êtres humains, et non pas à d'autres créatures. Comme le fait remarquer Roger White (2005, p. 15) : « Il est important de noter, avant de continuer plus loin, quelque chose que Luther souligne contre Erasme : même s'il souhaite insister sur l'incapacité dans laquelle se trouve l'homme d'obéir à la Loi de Dieu, il ne faut pas en conclure que Dieu commande aux Hommes des choses dont il est absurde d'imaginer qu'un homme puisse l'accomplir ; la Loi de Dieu est adressée aux hommes, pas à des pierres ou à des oies, et ce qui est commandé est dans cette mesure quelque chose de possible pour l'être humain ; l'idée que quelqu'un obéit à la Loi Divine est une idée parfaitement intelligible. Ce n'est pas comme s'il était commandé à l'Homme de voler ou de porter un poids de dix tonnes ». [“It is worth stressing before going any further, something which Luther stresses against Erasmus: even if he wishes to insist on a human inability to obey the Law of God, this must not be taken as though God were commanding things which it would be absurd to imagine a man actually doing: God's Law is addressed to men, and not to stones or geese, and what is commanded is to that extent a genuine human possibility: the idea of someone actually obeying God's Law is a perfectly intelligible idea. It is not as though Man is being commanded to fly or carry a ten-ton weight.”]

le débat des théologiens favorise une interprétation *morale* de DIP plutôt qu'une interprétation *conceptuelle*.³⁷⁰

2. ET SI LES DEVOIRS MORAUX CATÉGORIQUES N'EXISTAIENT PAS ?

Nous avons vu comment il est possible de répondre aux objections avancées contre l'analyse conceptuelle de DIP. Il nous faut maintenant répondre aux difficultés que soulève le *raisonnement moraliste* qui nous intéresse spécifiquement, à savoir le raisonnement visant à soutenir l'existence d'une motivation morale catégorique (et éventuellement dominante). En effet, le raisonnement moraliste proposé peut être contesté de deux manières : (1) soit en contestant l'existence de devoirs moraux catégoriques, soit (2) en contestant la possibilité d'une motivation par la seule croyance morale. Nous allons étudier ces deux problèmes à tour de rôle.

Comme nous l'avons vu, l'argument proposé par les partisans de la thèse controversée de la possibilité d'une motivation morale indépendante des désirs antérieurs au jugement moral consiste à dériver une conclusion descriptive sur la base d'une prémisse normative : il existe des devoirs moraux *catégoriques*, c'est-à-dire des devoirs moraux dont la validité ne dépend pas des désirs antérieurs au jugement. Or, comme « devoir » implique conceptuellement « pouvoir-être-motivé », cela signifie qu'il existe une motivation morale catégorique – i.e. qu'il est possible d'être motivé à faire ce qui est juste indépendamment des antérieurs, par la seule considération de la rectitude de l'action.³⁷¹

³⁷⁰ Un autre argument pour favoriser l'interprétation conceptuelle par rapport à l'interprétation morale de DIP consiste à dire que, si deux personnes ne sont pas d'accord au sujet du principe « devoir-implique-pouvoir », leur désaccord porte sur la *compréhension* des notions morales. Il ne s'agit pas d'un désaccord normatif, comme lorsque deux personnes se disputent pour savoir si l'euthanasie est moralement permise ou non. Bloomfield développe un argument dans ce sens : « Si nous rencontrons quelqu'un qui blâme d'autres personnes parce qu'elles n'ont pas changé le passé, nous n'engagerions pas avec cette personne un débat mais nous considérerions que nous avons rencontré quelqu'un qui n'a pas compris ce qu'est la morale. Le dogme de Kant [i.e. « devoir-implique-pouvoir »] n'est pas un dogme envers lequel nous pouvons adopter librement une pro- ou anti-attitude, comme on le pourrait à l'égard de l'euthanasie ou de l'avortement ; rejeter le dogme de Kant implique de changer la nature même de la morale. » [“If we met someone who blamed others for not changing the past, we would not engage such a person in debate but would think that we had encountered someone who did not know what morality was. Kant's dogma is not something toward which we are equally free to take a pro or con attitude, as we can toward abortion or euthanasia; rejecting Kant's dogma entails changing the very nature of “morality” (BLOOMFIELD, 2006, p. 458).]

³⁷¹ Nous pourrions étendre l'argument aux *raisons catégoriques* d'agir, et ainsi montrer que notre raisonnement s'applique aussi aux valeurs et aux vertus morales.

Plus formellement, le raisonnement moraliste se présente de la manière suivante :

(1) Les êtres humains ont le devoir moral catégorique (et dominant) de faire X.	Prémisse normative
(2) Si les êtres humains ont le devoir moral de faire X, alors ils peuvent être motivés à faire X.	DIP-motivation
(3) ∴ Les êtres humains sont capables de motivation morale catégorique (et dominante) pour faire X.	Conclusion descriptive

Pour contester cet argument, il faut pouvoir refuser ou bien la prémisse (1) selon laquelle il existe des devoirs moraux catégoriques, ou bien l'interprétation conceptuelle de DIP – et donc de DIP-motivation (2). Nous avons vu qu'il n'existe pas de bonnes raisons de rejeter la défense conceptuelle de DIP. Nous pouvons toutefois concevoir que quelqu'un conteste la première prémisse. Selon l'objecteur, il n'existe tout simplement pas de devoirs moraux catégoriques.

A cette objection, nous répondons la chose suivante : bien entendu, il est possible de contester qu'il existe des devoirs moraux catégoriques, valables pour tout individu, indépendamment de sa constitution motivationnelle subjective. Pour défendre une telle position, il faut montrer *soit* qu'il n'existe pas du tout de devoirs moraux – comme le font les partisans de la théorie de l'erreur – *soit* que les devoirs moraux sont des devoirs hypothétiques, c'est-à-dire des devoirs dont la validité dépend des désirs antérieurs de la personne. Cependant, comme nous le montrerons, (A) il existe de bonnes raisons d'admettre qu'il existe des devoirs moraux catégoriques. En outre, (B) si une personne souhaite contester l'existence des devoirs moraux catégoriques, elle doit veiller à ne pas commettre l'une ou l'autre de deux erreurs qu'il est aisé de commettre dans ce contexte.

(A) Les bonnes raisons d'admettre qu'il existe des devoirs moraux catégoriques

Les bonnes raisons d'admettre l'existence de devoirs moraux catégoriques sont les suivantes :

i. La sémantique

En premier lieu, admettre l'existence de devoirs moraux catégoriques correspond à la sémantique du langage moral ; dans le langage moral ordinaire, nous acceptons que les

devoirs moraux sont valables indépendamment des désirs antérieurs au jugement. En général, la charge de la preuve revient donc plutôt aux métaéthiciens qui nient le caractère catégorique des devoirs moraux : ils doivent expliquer le fossé entre leur théorie métaéthique et le langage moral ordinaire. Le plus grand danger, pour notre position, provient donc de la théorie de l'erreur morale, qui admet la sémantique du langage ordinaire mais conteste que le langage moral (catégorique) puisse être vrai.³⁷²

ii. *L'unanimité des théories morales normatives*

Toutes les théories normatives – même l'éthique de la vertu, comme le concède Philippa Foot³⁷³ – admettent l'existence de devoirs moraux catégoriques. Dans l'utilitarisme, le devoir de faire l'action qui maximise le plus grand bonheur du plus grand nombre s'applique aux personnes indépendamment de leurs désirs antérieurs. Dans l'éthique déontologique kantienne, le devoir d'agir conformément à des principes qui pourraient être suivis par tout le monde s'applique aussi indépendamment des désirs antérieurs.³⁷⁴ Dans le contractualisme de Scanlon, il en va de même ; le devoir moral d'agir selon des principes que personne n'a de raison de rejeter s'applique à toutes les personnes, indépendamment de leurs désirs antérieurs (SCANLON, 1998). Dans l'éthique de la vertu, ce qu'un être humain *doit faire* est défini par des faits objectifs au sujet de la nature humaine – de ses besoins, de son épanouissement, etc. Ce que chaque personne *doit faire* est donc défini par des faits indépendants des désirs antérieurs de cette personne (FOOT, 2001). Nous voyons donc que les principales théories *normatives* disponibles conçoivent les principes moraux comme étant des principes *catégoriques*. Si le raisonnement moraliste développé dans ce chapitre est correct, ces théories devraient donc toutes admettre aussi l'existence d'une *motivation morale catégorique* correspondante – le désir de réaliser le plus bonheur du plus grand nombre, d'agir selon des principes que personne n'a de raison de rejeter, de devenir une bonne personne, un bon être humain, etc.

³⁷² Malgré tout, il semble que les théoriciens de l'erreur – à l'exclusion des abolitionnistes – admettent la nécessité de faire appel à une motivation spécifiquement morale. Cela est nécessaire même si, selon eux, cette motivation repose sur une erreur.

³⁷³ « Je me demande désormais, étant donné le caractère évidemment indigeste de l'idée que la morale est un système d'impératifs hypothétiques, comment j'ai pu l'accepter, même pour une courte période. » [“I now wonder why, given the obvious indigestibility of the idea that morality is indeed a system of hypothetical imperatives, I should have accepted it even for a short while.” (FOOT, 2001, p. 61)].

³⁷⁴ « L'impératif catégorique [...] s'énonce comme suit : *Agis uniquement d'après la maxime grâce à laquelle tu peux vouloir en même temps qu'elle devienne une loi universelle.* » (KANT, 2007 (1785), p. 54).

(B) Les deux erreurs à éviter

Une personne qui souhaite contester l'existence de devoirs moraux catégoriques doit faire attention à ne pas commettre l'une des deux erreurs suivantes :

- i. *Soutenir que les devoirs catégoriques ne sont pas possibles, car il n'existe pas de motivation catégorique.*

Le point particulièrement important pour notre propos est le suivant : la direction de raisonnement appropriée en philosophie morale va de la recherche *normative*, au sujet de la forme et du contenu des exigences morales, vers la recherche *descriptive* au sujet des possibilités motivationnelles humaines. La conception de la motivation morale dépend de la recherche *normative* préalable. Tout argument prenant la direction *inverse*, c'est-à-dire partant d'une description de la motivation morale pour développer une doctrine normative, est injustifié. En conséquence, il n'est notamment pas possible de nier l'existence de devoirs moraux catégoriques à l'aide d'un argument concernant l'impossibilité d'une motivation morale catégorique. Si quelqu'un souhaite nier l'existence d'une motivation morale catégorique, il doit montrer qu'il n'existe pas de devoirs moraux catégoriques. Mais il ne peut pas contester l'existence de devoirs catégoriques en commençant par nier l'existence d'une motivation morale catégorique. Comme l'affirme Christine Korsgaard (1986), le scepticisme au sujet de la motivation morale catégorique dérive du scepticisme au sujet de l'existence de devoirs moraux catégoriques, et non l'inverse – contrairement à ce qui a largement été soutenu dans la littérature.

- ii. *Confondre les devoirs moraux catégoriques et les devoirs moraux inconditionnels?*

Le rejet de la possibilité des devoirs moraux catégoriques repose souvent sur une erreur particulièrement pernicieuse – à laquelle même Kant a succombé. Cette erreur consiste à confondre les devoirs moraux *catégoriques* et les devoirs moraux *inconditionnels*. Les devoirs moraux *catégoriques* sont des devoirs qui s'appliquent aux sujets moraux indépendamment de leurs *motivations* antérieures ou *non morales*. Les devoirs moraux *inconditionnels* sont des devoirs qui s'appliquent aux sujets moraux *indépendamment de toute condition empirique* – que cette condition concerne la motivation ou autre chose. Or, considérer que les devoirs moraux sont des devoirs *catégoriques* ne contraint aucunement à

les considérer comme des devoirs *inconditionnels*. Il est tout à fait possible de soutenir à la fois (A) que les devoirs moraux sont valides uniquement *si certaines conditions empiriques* sont satisfaites, tout en soutenant, simultanément, (B) que les devoirs moraux sont valides indépendamment des désirs des sujets moraux antérieurs au jugement moral.³⁷⁵

Il n'existe en fait pas de devoirs moraux *inconditionnels*. En effet, tout devoir moral n'est valide que *si certaines conditions empiriques sont satisfaites*. Parfois, l'énonciation de ces conditions est laissée implicite. Par exemple, étant donné que « devoir » implique conceptuellement « pouvoir », il est toujours présupposé implicitement dans l'énonciation d'un devoir que les conditions empiriques sont réunies qui rendent possible le respect ou la violation de ce devoir.

La condition sous laquelle toute norme – individuelle ou générale – est valide est la somme des circonstances dans lesquelles elle peut être observée ou violée. Une norme qui décrète comme étant obligatoire quelque chose qui se produit dans le cours naturel des événements, telle que la norme « les êtres humains doivent mourir » (qui ne peut pas être violée), ou quelque chose qui ne peut pas se produire dans le cours naturel des événements, telle que la norme « les êtres humains doivent vivre pour l'éternité » (qui ne peut pas être observée), [...] n'est pas considérée comme valide.³⁷⁶

Ainsi, même des normes qui *semblent* inconditionnelles, telle que la norme « aime ton ennemi ! », ne sont que des normes *conditionnelles*, puisqu'elles ne sont valides *au minimum* que sous la condition que l'ennemi en question existe.

Même la norme générale, formulée de façon inconditionnelle « Aime ton ennemi ! » n'est valide que de façon conditionnelle, c'est-à-dire sous la condition

³⁷⁵ Pour rappel, nous soutenons que DIP permet d'exclure les normes qu'il est logiquement, physiquement, psychologiquement, biologiquement ou économiquement impossible de respecter. Nous soutenons également que *s'il existait des incapacités motivationnelles* empêchant de faire des actions, alors ces incapacités supprimeraient bien les devoirs moraux portant sur ces actions. Il n'existe cependant pas, selon nous, de telles incapacités motivationnelles.

³⁷⁶ "The condition under which any norm—individual or general—is valid is the sum of the circumstances in which it can be observed or violated. A norm which decrees as obligatory something which must occur in the natural course of events, such as the norm 'Human beings are to die' (which cannot be violated), or something which cannot occur in the natural course of events, such as the norm 'Human beings are to live forever' (which cannot be observed), is pointless and is therefore not considered to be valid." (KELSEN, 1991, p. 20).

que la personne a un ennemi. Sa formulation correcte est donc : « Si quelqu'un a un ennemi, alors il doit l'aimer ! ». ³⁷⁷

En conclusion, s'il peut effectivement sembler suspect de soutenir que les devoirs moraux seraient valides de façon *inconditionnelle*, il n'est pas suspect de considérer que les devoirs moraux sont valides de façon *catégorique*. Ce n'est que par une confusion des deux notions que le caractère *catégorique* des devoirs moraux devient problématique.

3. LE PROBLÈME FINAL : L'IMPOSSIBILITÉ DE MOTIVATION SANS MOTIVATION

3.1. L'OBJECTION

Nous avons suggéré que l'argument moraliste peut servir à inférer la possibilité d'une motivation morale *catégorique*, c'est-à-dire d'une motivation à agir par la seule considération de la valeur morale de l'action, indépendamment des désirs antérieurs, et même à l'encontre, des désirs antérieurs de la personne. La conclusion du raisonnement moraliste est ainsi qu'il est possible d'être motivé à agir moralement, même en l'absence d'un désir pour une fin indépendante de la valeur morale elle-même.

Toutefois, il existe une objection puissante contre cette conclusion. Selon cette objection, il est *conceptuellement impossible* d'être motivé à agir par une croyance (morale) seule, en l'absence d'un désir *indépendant* de cette croyance. En effet, dans cette perspective, les désirs sont *par définition* ce qui motive à agir. Il est donc nécessaire que soit présent un désir à chaque fois qu'il y a motivation à agir. L'idée d'une motivation à agir en l'absence d'un désir – i.e. par une croyance seule – est une impossibilité conceptuelle. Or, étant donné que les théories morales qui admettent l'existence de devoirs moraux catégoriques sont contraintes, en supposant DIP-motivation, d'admettre l'existence correspondante d'une motivation catégorique, alors de telles doctrines doivent être rejetées. L'idée d'une motivation catégorique – sans désir, par la seule croyance – est une contradiction dans les termes. Affirmer qu'une croyance peut motiver à agir en l'absence d'un désir indépendant, cela revient à dire qu'une croyance peut motiver à agir en l'absence

³⁷⁷ "Even the categorically formulated general norm 'Love your enemies' is valid only conditionally, that is, on the condition that a person has an enemy. It is not valid if a person has no enemy, and so its correct formulation is: 'If someone has an enemy, he is to love him.'" (KELSEN, 1991, p. 20)

de motivation. Mais « motiver à agir en l'absence de motivation » est une idée contradictoire, et donc impossible. Pour avoir une motivation, il faut avoir une motivation.

Cette objection conceptuelle contre la possibilité d'une motivation par la seule croyance morale fut déjà formulée à l'encontre de la thèse d'Emmanuel Kant par Christian Garve, dans les termes suivants :

[I]l faut que l'homme ait un *mobile* qui le mette en mouvement [...].³⁷⁸

Elle se retrouve aujourd'hui dans les propos du philosophe anti-kantien *par excellence*, Bernard Williams :

Seule la motivation motive. [...] On peut penser qu'il s'agit d'une tautologie et c'est, à peu de choses près, ce dont je pense qu'il s'agit.³⁷⁹

Cette objection conçoit comme *incohérente* la thèse selon laquelle une croyance *seule* peut motiver à agir. En effet, lorsque nous sommes motivés à agir, nous devons reconnaître qu'il y a bien « quelque chose » qui exerce la force motivationnelle. Ce « quelque chose » qui nous motive, nous le désignons par le terme générique de « désir » ou « pro-attitude ». Il nous faut donc bien reconnaître que, lorsque nous sommes motivés à agir, il y a nécessairement présence d'un désir ou d'une « pro-attitude ». L'idée d'une motivation par la seule croyance morale ne fait en conséquence aucun sens.

La réponse à cette objection prend la forme d'un dilemme : soit l'objection repose sur une conception *tautologique* du désir et, dans ce cas, elle est entièrement inoffensive ; soit elle repose sur une conception *substantielle* du désir et, dans ce cas, l'idée d'une motivation sans désir n'est plus *contradictoire*.

Pour rappel, la thèse de Williams est la suivante : « une motivation sans désir est impossible ». De ce fait, la thèse implique qu'une motivation par une croyance seule est impossible. Dans sa structure générale, le dilemme auquel est confrontée cette thèse est le suivant : soit le désir est défini de manière formelle comme tout « ce qui motive à agir », dans quel cas la thèse de Williams exprime bien une vérité tautologique mais est inoffensive. Soit le désir est défini substantiellement comme étant *une catégorie spécifique* de

³⁷⁸ (KANT, 1986 (1793))

³⁷⁹ "Only motivation motivates. [...] This might be thought as a tautology and that is, just about, what I take it to be" (WILLIAMS, 1995, p. 117).

motivation, dans quel cas l'idée d'une motivation sans désir n'est plus une contradiction dans les termes.

Au final, ce n'est que lorsque l'objection de Williams repose sur un usage indifférencié de deux significations du terme « désir » qu'elle acquiert l'apparence trompeuse d'une thèse *à la fois* substantielle et tautologique.³⁸⁰

3.2. LE DILEMME

Il nous faut maintenant présenter le dilemme plus en détail. Nous verrons ainsi que ni la conception formelle du désir, ni la conception substantielle du désir, ne permet de contester la possibilité d'une motivation par une croyance seule.

Selon la première définition de la notion de « désir », celle-ci est équivalente à la notion de « motivation » comme étant simplement « ce qui motive à agir » ou « ce qui fait agir ». Autrement dit, « désir » est considéré comme synonyme de « motivation », comme étant simplement « ce qui motive à agir ». Le problème avec cette conception tautologique du désir, comme le remarque Michael Huemer, est qu'elle est compatible avec toutes les conceptions de la motivation morale.

Le cœur de la thèse humienne que les motivations à agir requièrent des désirs ne consiste en rien de plus qu'en la tautologie suivante, que même les plus fervents rationalistes peuvent accepter : les motivations à agir requièrent des motivations à agir.³⁸¹

Cette conception formelle du « désir », comme étant « ce qui motive à agir », n'exclut pas la possibilité d'une motivation morale catégorique, puisqu'elle n'exclut pas la possibilité ou bien (i) que la reconnaissance de la valeur morale d'une action suscite une motivation

³⁸⁰ Cette critique de la conception purement *formelle* ou *logique* du désir est présente aussi chez Thomas Nagel. Le raisonnement concerne les *raisons d'agir* en général, plutôt que les raisons morales spécifiquement, mais la conclusion du raisonnement est la même : « Le fait qu'un désir soit une condition nécessaire (parce qu'il s'agit d'une conséquence logique) du fait qu'une raison motive à agir, cela n'implique pas que ce désir soit une condition nécessaire pour l'existence d'une raison d'agir ; si le désir est motivé par la raison d'agir, alors il ne peut pas faire partie des conditions d'existence de cette raison elle-même. » [“The fact that the presence of a desire is a logically necessary condition (because it is a logical consequence) of a reason's motivating, does not entail that it is a necessary condition of the presence of the reason; and if it is motivated by that reason it cannot be among the reason's conditions.” (NAGEL, 1970, p. 30).]

³⁸¹ “The core Humean thesis that motives for action require desires amounts to no more than the following tautology, with which even the staunchest rationalist will agree: Motives for action require motives for action. (HUEMER, 2005, p. 165).”

nouvelle, c'est-à-dire un *désir nouveau*, dont l'existence *dépend* de la croyance morale, ou bien (ii) que la croyance morale elle-même motive à agir.

Dans le second sens de « désir », ce concept ne réfère pas simplement à « ce qui motive à agir », mais réfère à une conception plus spécifique du désir. Par exemple, un désir peut être considéré comme un état motivationnel *affectif* (NAGEL, 1970, p. 4) ou comme un état motivationnel ayant une phénoménologie spécifique ou encore comme étant un état motivationnel portant sur le plaisir attendu (LEMAIRE, 2007). Peu importe laquelle de ces conceptions *spécifiques* du désir est adoptée, il n'est alors plus *conceptuellement* impossible d'être motivé à agir en l'absence d'un désir.

En effet, le terme « désir » n'étant plus identifié à simplement « ce qui motive à agir », il devient tout à fait envisageable d'être motivé à agir en l'absence d'un désir. Admettons, par exemple, que le désir soit défini spécifiquement comme étant un état motivationnel *affectif*, c'est-à-dire un état motivationnel dans lequel on *se trouve*, mais qui n'est pas sujet à une évaluation rationnelle. Tout état mental motivationnel qui ne serait pas affectif, mais qui serait dérivé d'une croyance et susceptible d'évaluation rationnelle, ne serait ainsi pas un désir. Si le désir est défini de cette manière, alors il n'est plus tautologiquement vrai que toute motivation requiert la présence d'un désir ; il n'est pas *tautologiquement vrai* que toute motivation à agir dépend d'un état motivationnel affectif, non sujet à une évaluation rationnelle. Il est tout à fait *conceptuellement possible* que la motivation soit fournie par un état motivationnel *non affectif*, fondé sur une croyance.

Comme le décrit parfaitement Thomas Scanlon, aucune considération purement conceptuelle ne peut exclure qu'il existe des motivations qui soient *l'effet motivationnel* d'une croyance plutôt que *sa condition motivationnelle* :

[Si le concept de] désir est considéré comme incluant n'importe quelle "pro-attitude" qu'un agent peut avoir à l'égard d'une action ou d'un résultat, quel que soit le contenu ou le fondement de cette attitude, [alors] il est incontestable que les désirs dans ce sens large sont capables de nous pousser à l'action, et il est plausible d'affirmer qu'ils sont les seules choses capables de le faire, puisque tout ce qui nous motive (au moins dans les actions intentionnelles) sera vraisemblablement considéré comme un désir. [...] Mais de nombreux éléments de cette classe [de désirs] sont ce que Nagel appelle des « désirs fondés » ; c'est-

à-dire, ces désirs ne sont pas *la source* de la motivation, mais plutôt la conséquence motivationnelle de quelque chose d'autre, comme la reconnaissance que quelque chose est un devoir moral, ou que quelque chose est soutenu par une raison d'un autre type.³⁸²

En conclusion, il n'existe pas de raison *conceptuelle* pour rejeter la possibilité d'une motivation par la seule croyance morale. Tout argument devant servir à rejeter cette possibilité doit dépasser la seule analyse conceptuelle des notions de « désirs » et de « motivation ».

4. PEUT-ON COMPLÈTEMENT IGNORER LES DIFFICULTÉS MOTIVATIONNELLES DES ÊTRES HUMAINS ?

Avant de clore ce travail, nous devons encore clarifier notre position en spécifiant le rôle que nous accordons, finalement, aux considérations motivationnelles, dans une théorie politique.

Nous avons soutenu la thèse selon laquelle il ne faut pas tenir compte des défaillances motivationnelles de l'être humain *dans la formulation* des principes moraux eux-mêmes. Plus précisément, nous avons soutenu que les défaillances motivationnelles humaines ne constituent pas des raisons pour *limiter* ou *diminuer* les exigences morales. Est-ce que cela signifie que nous pouvons ignorer complètement les considérations au sujet des difficultés motivationnelles que les êtres humains pourraient rencontrer dans l'accomplissement de leur devoir ?

A cette question, nous répondons par la négative : le fait que la considération des défaillances motivationnelles ne permet pas de *limiter* les exigences morales n'implique pas qu'il ne faut pas en tenir compte. Bien au contraire, la prise en compte des défaillances

³⁸² "[If the concept of] a desire is taken to include any "pro-attitude" that an agent may have toward an action or outcome, whatever the content or basis of this attitude may be, [then it] is uncontroversial that desires in this broad sense are capable of moving us to act, and it is plausible to claim that they are the only thing capable of this, since anything that moves us (at least to intentional action) is likely to count as such a desire. [...] But many elements of this class are what Nagel calls "motivated desires"; that is to say, they do not seem to be sources of motivation but rather the motivational consequences of something else, such as the recognition of something as a duty, or as supported by a reason of another kind." (SCANLON, 1998, p. 37).

humaines ouvre la voie à la formulation de *nouveaux* devoirs moraux.³⁸³ En effet, sous hypothèse que chaque individu a le devoir d'aider à la réalisation d'un monde moralement bon/juste, il a le devoir d'assurer les conditions qui permettent sa réalisation et d'aider les personnes à accomplir leur devoir, particulièrement lorsque cela est difficile. Autrement dit, lorsque le devoir est difficile à réaliser, nous avons le devoir d'aider les personnes à l'accomplir. Et, *surtout*, nous avons le devoir moral de soutenir les institutions qui facilitent la réalisation des devoirs moraux.³⁸⁴

En suivant Dubbink et Van de Ven, la prise en compte des limitations motivationnelles humaines fonde *pour tous les membres de la société* l'ensemble suivant de devoirs moraux :

- i. Un devoir de soutenir les circonstances qui rendent possible la motivation morale : « Si les êtres humains souffrent souvent d'un manque de motivation morale, alors ils ont le devoir moral de veiller aux circonstances dans lesquelles ils doivent accomplir leur devoir. »³⁸⁵
- ii. Un devoir de se changer soi-même, si la motivation actuelle est insuffisante, afin de renforcer sa motivation morale.
- iii. Un devoir d'organiser les problèmes collectifs au niveau approprié, c'est-à-dire au niveau collectif.

Ainsi, même s'il est vrai que, dans une société « bien ordonnée », pour reprendre les termes de Rawls, il existe une forte convergence entre les exigences de prudence et les

³⁸³ Pour Thomas Nagel, la prise en compte des défaillances motivationnelles de l'être humain sert à départager, d'un côté, les actions moralement *obligatoires* et, de l'autre, les actions moralement *surrogatoires*. « La vertu surrogatoire consiste à adhérer aux principes de la morale impersonnelle avant que ceux-ci n'aient été modifiés pour tenir compte des limitations normales de la nature humaine ».[“[S]upererogatory virtue is adherence to the claims of impersonal morality prior to their modification to accommodate the normal limitations of human nature” (NAGEL, 1986, p. 204)]. Dans notre approche, la considération des défaillances ne joue pas le même rôle ; l'existence de défaillances motivationnelles crée de nouveaux devoirs moraux, notamment le devoir d'aider à accomplir son devoir, et de soutenir les formes de société dans lesquelles la vie morale est possible, mais ne supprime ou ne diminue la gravité d'aucun devoir moral.

³⁸⁴ Contrairement aux approches économiques, qui souhaitent faire que l'action morale devienne *intéressante* grâce à l'organisation d'incitations, l'approche que nous proposons ne rend pas l'action morale *intéressante* ; s'il est vrai que la difficulté à accomplir son devoir est modérée à l'aide de la mise en place d'institutions collectives, cela n'est possible que parce que *les membres de la société* acceptent de renoncer à une partie de leurs intérêts pour participer à la réalisation du bien moral. Si l'action morale devient moins difficile à accomplir pour S dans un certain modèle de société, cela n'est possible que parce que les autres membres de sa société ont accompli leur devoir moral d'aider à la réalisation du bien, de la justice et de la vertu, au détriment de la satisfaction de leurs propres intérêts.

³⁸⁵ “If human beings are often lacking in moral motivation, they have a moral duty to attend to the circumstances under which they are required to act.” (DUBBINK & VAN DE VEN, 2012, p. 234).

exigences morales, ce n'est pas cette convergence qui explique le respect des normes morales. Au contraire, c'est l'existence d'un devoir moral « de second ordre »³⁸⁶, qui exige de favoriser les conditions qui permettent l'accomplissement des devoirs moraux « de premier ordre », qui explique la motivation première des citoyens de vivre dans une société « bien-ordonnée ».³⁸⁷

Par exemple, admettons qu'il soit moralement obligatoire de cesser ou de fortement diminuer les activités émettrices de gaz à effet de serre, pour atteindre un seuil compatible avec la préservation de l'environnement. Pour certains acteurs économiques, respecter ce devoir implique de souffrir d'un important désavantage concurrentiel, menant potentiellement à la cessation de l'activité. Il est même possible que leurs activités économiques soient *essentielle*ment incompatibles avec la préservation de l'environnement. Les activités liées à l'extraction d'énergie fossile, à l'industrie de l'automobile thermique, à l'aviation de tourisme, etc. semblent tout simplement irréconciliables avec les objectifs de lutte contre le changement climatique. La question qui se pose est alors la suivante : les entreprises investies dans ces productions, ont-elles le devoir moral de préserver l'environnement, même si cela implique, pour elles, de renoncer complètement à leur activité économique ?

Etant donné notre argumentation précédente, nous répondons sans surprise par la positive : les entreprises ont le devoir moral de cesser leur activité – indépendamment de ce que font les autres. Toutefois, étant donné l'importance du sacrifice personnel engagé, le reste de la population est soumis au devoir moral *de second ordre* d'offrir à ces entreprises des conditions facilitant la réalisation de devoir de premier ordre de préserver l'environnement. Les autres citoyens auraient, par exemple, le devoir moral de second ordre de soutenir des mesures publiques permettant une reconversion professionnelle des employés.³⁸⁸

³⁸⁶ Un devoir moral « de second ordre » est un devoir portant sur d'autres devoirs moraux, et qui ne porte pas directement sur une action spécifique, contrairement aux devoirs moraux « de premier ordre ».

³⁸⁷ Bien entendu, cette motivation morale première peut être soutenue par des motivations non morales complémentaires ; notamment l'intérêt personnel, mais aussi le désir d'apparaître comme une personne vertueuse, le plaisir quasi-esthétique de vivre dans une vie « bien ordonnée », etc.

³⁸⁸ Cependant, il est important de garder à l'esprit que l'entreprise a le devoir moral de préserver l'environnement, *même si* elle n'est pas soutenue dans ses démarches par le reste de la population, car aucune considération de difficulté ou de sacrifice ne supprime le devoir.

Le devoir moral « de second ordre » de favoriser les circonstances qui facilitent l'accomplissement des devoirs moraux « de premier ordre » implique également que chacun a le devoir d'aider au développement et au maintien des dispositions morales, chez soi-même et chez les autres. C'est pourquoi un système social tel que préconisé dans l'approche « économique »³⁸⁹, dans lequel l'État est responsable de susciter l'action morale chez les acteurs à l'aide d'incitations économiques, est non seulement *irréaliste*, mais *immoral* : il met en danger les capacités de jugement moral et de motivation morale des acteurs économiques.

L'agentivité morale [i.e. la capacité à formuler des jugements moraux et agir conformément à ces jugements] peut être endommagée, bloquée et même peut-être complètement éradiquée [...]. Elle a besoin d'entraînement, de développement et d'expérience. Elle a besoin d'un environnement qui la renforce. En conséquence, l'environnement offert par les institutionnalistes³⁹⁰ moraux aux agents moraux est trop hostile.³⁹¹

Ainsi, l'existence de défaillances motivationnelles répandue chez les êtres humains nous donne le devoir de travailler à renforcer nos propres capacités morales ainsi que celles de nos concitoyens. Plus encore, nous avons le devoir moral de soutenir les institutions qui facilitent la réalisation de nos devoirs moraux. Cela implique, notamment, d'accepter la prise en charge collective des problèmes collectifs, en ne s'opposant pas aux lois censées permettre de les régler. En outre, ayant connaissance de nos défaillances motivationnelles, des cas de faiblesse de volonté et d'irrationalité, nous devons accepter les mesures institutionnelles dont le but est de *nous aider* à surmonter ces défaillances.³⁹²

³⁸⁹ Pour rappel, l'approche « économique » vise à « faire l'économie de la vertu », c'est-à-dire à obtenir un comportement extérieurement vertueux en s'appuyant exclusivement sur l'intérêt personnel des personnes.

³⁹⁰ Les « institutionnalistes » moraux auxquels il est fait référence sont les théoriciens qui préconisent que les représentants politiques sont responsables de mettre en place les incitations permettant aux acteurs économiques de ne se préoccuper que de leur intérêt personnel. Ils correspondent aux « économistes » de notre travail.

³⁹¹ «Moral agency can be damaged, impaired and perhaps even completely eradicated [...]. It needs training, development and experience. It needs a reinforcing environment. Hence, the environment that the moral institutionalists offer the moral agent is too hostile.» (DUBBINK & VAN DE VEN, 2012, p. 236).

³⁹² Notre proposition semble compatible avec la thèse de John M. Doris (2002), selon laquelle l'échec moral est le plus souvent attribuable aux circonstances de choix qu'à une déficience de la personnalité. Néanmoins, notre conclusion ne s'accorde pas avec la thèse de l'auteur selon laquelle il faut, en conséquence, modifier nos exigences morales ou les transférer à l'État.

Ainsi, l'une des tâches ou des devoirs de l'organisation politique consiste à *rendre accessible* la vie morale aux citoyens, afin de faciliter l'accomplissement du devoir.

Une tâche importante, peut-être même la plus importante, de la pensée et de l'action politique est d'arranger le monde pour que chacun puisse vivre une bonne vie sans faire le mal, sans porter atteinte à autrui, sans bénéficier de manière inéquitable de l'infortune des autres, etc. L'harmonie morale et pas seulement la paix civile est le but correct de la politique, et il serait désirable de l'obtenir sans imposer à tout le monde le type de conversion personnelle qui serait nécessaire pour supprimer complètement les tensions entre la morale et le bien personnel.³⁹³

Il est important de souligner qu'une telle approche n'implique pas de *renoncer* aux motivations morales – contrairement à l'approche économique. Elle n'implique pas non plus d'attribuer l'entière responsabilité des fautes morales – lorsqu'elles concernent des problèmes collectifs – à un manquement de l'État dans la mise en place d'incitatifs. Elle implique uniquement le *devoir* – pour *chaque individu* – de favoriser les conditions sociales et politiques dans lesquelles la vie morale est accessible sans trop de difficultés et sans trop de sacrifices. Et puisque l'existence d'institutions sociales et politiques appropriées facilite l'accomplissement de la vie morale, chacun a le devoir moral de les soutenir et de les promouvoir.

La conception de l'usage approprié des lois, des sanctions et des incitations économiques qui émerge de notre approche est la suivante : en premier lieu, les lois et les sanctions servent de *garantie* ou *d'assurance*. En effet, il peut être irrationnel de poursuivre seul un objectif qui ne peut être atteint que collectivement. Pour qu'il soit rationnel de *faire sa part*, il est donc nécessaire de savoir que les autres personnes participent également à l'effort commun. La sanction par la loi permet alors d'offrir cette forme d'assurance – en fournissant une norme, publiquement reconnue, d'évaluation des comportements.

³⁹³“An important, perhaps the most important task of political thought and action is to arrange the world so that everyone can live a good life without doing wrong, injuring others, benefiting unfairly from their misfortune, and so forth. Moral harmony and not only civil peace is the right aim of politics, and it would be desirable to achieve it without putting everyone through the type of deep personal Conversion needed to make a clash between morality and the good life impossible.” (NAGEL, 1986, p. 206).

Les sanctions légales sont [parfois] nécessaires pour résoudre le « problème de l'assurance ». Il est parfois contraire à la raison, pour une personne particulière, d'adhérer à un principe moral particulier, si les autres personnes n'adhèrent pas, elles aussi, à ce principe. Dans de telles situations, le soutien coercitif au principe dont il est question peut être nécessaire pour fournir l'assurance requise de la participation des autres personnes. [De telles sanctions] fournissent seulement à chacun une *raison de croire* que les autres vont agir de manière adéquate.³⁹⁴

Les lois et sanctions n'ont ainsi pas pour but de *susciter* un comportement moral en l'absence de motivation morale, mais de soutenir la motivation morale en offrant une forme de garantie que les efforts déployés pour atteindre des objectifs moraux ne sont pas vains.³⁹⁵

Pour finir, dans les problèmes d'action collective, un obstacle important provient de la difficulté à *identifier* précisément son devoir moral. En effet, lorsque nous considérons le cas du changement climatique, par exemple, il est extrêmement difficile de savoir quel est exactement le style de vie moralement approprié. Dans ce cas, l'ensemble des personnes disposant des informations nécessaires à la prise de décision ont le devoir moral de fournir cette information – en tant qu'elles ont le devoir moral d'offrir les conditions qui permettent l'accomplissement de l'action morale. Elles ont, réciproquement, le devoir moral *de ne pas abuser* de la motivation morale des personnes en leur fournissant une information mensongère – par exemple, en créant des labels écologiques trompeurs.

³⁹⁴ “[L]egal sanctions are [sometimes] required to solve the “assurance problem”. Sometimes adherence to a particular moral principle by a given agent is contrary to reason unless most others adhere to it as well. And in such cases, coercive support for the principle in question may be necessary to provide the required assurance of performance by others. [Such sanctions] merely provide each with *reason for believing* that the others will behave in the required way.” (MCMAHON, 1981, p. 274).]

³⁹⁵ John Rawls offre une interprétation similaire des lois et sanctions dans les sections 42 et 52 de la *Théorie de la justice* (RAWLS, 2009).

5. CONCLUSION

Dans ce chapitre, nous avons soutenu la thèse selon laquelle il n'est possible de déterminer les caractéristiques (substantielles et formelles) de la motivation morale que lorsque nous bénéficions, au préalable, d'une théorie morale normative. Nous avons ainsi affirmé qu'aucune considération motivationnelle ne peut servir à imposer des limites ou restrictions aux devoirs moraux des êtres humains. Au contraire, la prise en compte des défaillances et limitations motivationnelles donnent lieu à de *nouveaux devoirs*, de second ordre, dérivant du devoir moral d'aider à la réalisation des devoirs moraux de premier ordre. Mais ces seconds devoirs d'assistance ne font absolument pas disparaître les devoirs moraux de premier niveau.

CONCLUSION

Les réponses aux objections anti-moralistes (économique et métaéthique) que nous avons proposées dans les premiers chapitres, combinées à l'ébauche d'une défense *positive* du raisonnement moraliste que nous avons développée dans le dernier chapitre, nous amènent aux conclusions générales suivantes :

(1) Devoir-implique-pouvoir

Notre étude du rapport entre la recherche empirique au sujet des motivations humaines et la recherche éthique au sujet des devoirs moraux nous a amenés à étudier le principe « devoir-implique-pouvoir ». Dans ce contexte, nous avons cherché à déterminer deux choses : en premier lieu, nous nous sommes demandés quelle était la meilleure interprétation de ce principe. Nous avons alors soutenu que l'interprétation *conceptuelle* devait être favorisée. De ce fait, selon nous, un énoncé du type « les êtres humains ont le devoir moral de se téléporter, bien qu'ils soient incapables de se téléporter » est un énoncé contradictoire, car il implique à la fois que les êtres humains *peuvent* se téléporter et que les êtres humains *ne peuvent pas* se téléporter.

En second lieu, nous avons tenté d'étendre le principe DIP aux limites *motivationnelles* de la nature humaine. Ce mouvement d'extension se justifie par le fait que, dans certaines circonstances, ce qui explique qu'une personne soit *incapable de faire X* est le fait que cette personne ne peut simplement pas *être motivée* à faire X. Dans ce contexte, nous avons néanmoins constaté que les limites motivationnelles des êtres humains possèdent une caractéristique que ne possèdent pas les autres limites (physiques, biologiques, psychologiques, etc.) : les limites motivationnelles n'existent pas ou, du moins, elles ne peuvent pas être connues. De ce fait, il n'est pas possible de rejeter une théorie morale par une référence à de prétendues impossibilités motivationnelles humaines. Au contraire, le principe DIP-motivation permet d'étendre notre conception des possibilités motivationnelles humaines, puisqu'il peut être utilisé de manière non contraposée dans un argument moraliste, selon lequel « si tu dois, tu peux être motivé ! ».

(2) La direction d'argumentation

Notre étude permet également de conclure qu'il ne fait pas sens de discuter de la motivation morale *avant* d'avoir déterminé quelle théorie morale est vraie. Une théorie normative ne peut pas être rejetée en faisant appel à une certaine conception de la motivation humaine avec laquelle elle serait incompatible. La direction adéquate d'argumentation ne débute donc pas par une recherche empirique au sujet des motivations, à partir de laquelle une théorie morale particulière serait ensuite déduite ou rejetée. Le soutien motivationnel à une doctrine morale provient du contenu de cette doctrine elle-même, et ne doit pas être recherché dans des motivations indépendantes de la doctrine morale en question. Le soutien motivationnel provient d'une élucidation et clarification des concepts moraux, ainsi que d'une clarification des devoirs et droits de chaque personne, mais il ne provient pas de motivations non morales indépendantes. En effet, le rôle de la théorie morale est précisément de mettre en lumière une nouvelle source de motivation, proprement morale, et ne consiste pas à organiser des ressources motivationnelles non morales.

Nous rejoignons ainsi Thomas Nagel pour affirmer qu'un système moral ne doit pas être considéré comme l'étude des conséquences d'une motivation humaine fondamentale – par exemple du désir de préservation de soi de Hobbes – sur les relations entre les hommes et comme la déduction consécutive d'un modèle de société. L'éthique n'est pas une simple branche de la psychologie empirique. L'éthique ne dépend pas de l'identification d'une motivation antérieure à la construction du système moral. Au contraire, l'éthique permet de découvrir une motivation nouvelle, dépendante des principes moraux. Dans cette perspective, ce n'est pas une motivation non morale qui rend possible l'éthique, mais c'est au contraire l'éthique qui explique et rend possible la motivation morale.

[L]éthique, ne consiste pas à s'approprier une fondation motivationnelle antécédente et indépendante, sur laquelle construire ensuite les principes éthiques, mais consiste plutôt à dégager une structure motivationnelle qui est

spécifiquement éthique et qui est expliquée précisément par ce qui explique les principes éthiques eux-mêmes.³⁹⁶

En conséquence, il n'est pas possible de discuter de motivation morale avant de disposer d'une doctrine morale. A la limite, le mieux que nous puissions faire, c'est *comparer* les motivations morales associées aux différentes théories morales : est-ce qu'une théorie utilitariste, basée sur le désir de promouvoir le bien-être de tous les êtres sensibles, a besoin de recourir plus souvent à des incitations et sanctions qu'une théorie contractualiste, basée sur le désir de faire des actions qu'autrui n'a pas de raison de rejeter ou sur le désir de « ne pas faire une exception pour soi » ? Dans tous les cas, il n'est pas possible de discuter de la motivation morale *indépendamment* des théories normatives. L'appel à la « motivation morale » ne peut donc jamais servir à favoriser une position métaéthique plutôt qu'une autre.

(3) Des prédispositions externalistes

Notre approche favorise l'externalisme, dans le débat métaéthique sur la nature du *jugement moral* : il est possible qu'une personne pleinement rationnelle ne soit pas motivée à faire ce qui est bien, juste, vertueux – étant donné que ce qui est bien, juste, vertueux, est toujours défini de manière substantielle. Ce qui n'implique pas, à l'inverse, que la motivation morale n'est pas rationnelle : au contraire, en lui attribuant un contenu spécifique, il devient nettement plus clair que la faute commise par une personne qui fait le mal est d'ignorer une chose qui a de l'importance, d'ignorer une raison d'agir substantielle, mais qu'il ne s'agit pas d'une faute de raisonnement, d'une faute logique ou d'une incohérence. Cette approche rend mieux compte de la spécificité de la *faute morale* que les approches qui essaient de réduire la faute morale à une faute logique.

En effet, les théories internalistes reposent sur l'idée que le seul moyen de montrer que la morale est importante consiste à réduire la faute morale à une faute grave de logique.

³⁹⁶ "[E]thics, rather than appropriating an antecedently comprehensible motivational foundation on which to build its requirements, actually uncovers a motivational structure which is specifically ethical and which is explained by precisely what explains those requirements." (NAGEL, 1970, p. 12).

Cependant, ce faisant, elle dénature la particularité de la faute morale. Comme l'affirme Thomas Scanlon³⁹⁷ :

Les conceptions formelles [...] ne fournissent pas une description adéquate de ce qui ne va pas lorsqu'une personne ne [se soucie pas de la morale]. La force particulière des exigences morales semble bien différente de celle, disons, des principes de logique, même si toutes les deux sont, en quelque sorte, « inéluctables ». Et la faute impliquée dans l'échec à être motivé par des exigences morales ne semble pas être une forme d'incohérence.³⁹⁸

A l'inverse, les théories externalistes, qui offrent une conception *substantielle* du jugement moral, ne réduisent pas la faute morale à une faute de logique. De ce fait, elles ne peuvent pas affirmer que toute personne qui rejette la morale se contredit elle-même ou qu'elle commet une faute grave de raisonnement. Cette incapacité ne constitue toutefois pas un véritable désavantage, car les théories externalistes rendent mieux compte de la nature spécifique de la motivation et de la faute morale. En outre, comme le fait remarquer Scanlon, dans les faits, les théories substantielles de la morale sont celles qui ont rencontré le plus de succès.

[D]ans les faits, les conceptions de la morale qui ont obtenu le plus grand soutien ont en général été les conceptions substantielles. Les idées de la volonté divine ou d'une fin humaine, par exemple, ont semblé fournir à de nombreuses personnes une conception adéquate parce qu'elles semblaient avoir la priorité et l'importance nécessaires.³⁹⁹

³⁹⁷ Thomas Pink (2004) émet le même type de critique à l'encontre des théories qui tentent de réduire la faute morale à une faute logique. De la même manière, Julia Markovits (2014), pourtant partisane de l'internalisme, reconnaît cette difficulté pour sa théorie. Pour y répondre, elle définit la faute morale comme une faute logique ayant pour spécificité d'engendrer des conséquences négatives importantes sur le bien-être d'autrui. Ces conséquences sur le bien-être d'autrui expliquent, selon elle, que la faute morale soit plus grave que d'autres types de fautes de raisonnement. Toutefois, en répondant de cette manière, la philosophe glisse vers une conception substantielle de la morale que pourtant elle rejette.

³⁹⁸ "Formal accounts [do] not give a very satisfactory description of what is wrong with a person who fails to [care about morality]. The special force of moral requirements seems quite different from that of, say, principles of logic, even if both are, in some sense, "inescapable." And the fault involved in failing to be moved by moral requirements does not seem to be a form of incoherence. For these reasons, looking for a substantive account seems to me a more promising strategy." (SCANLON, 1998, p. 151).

³⁹⁹ "[In] fact the accounts of morality that have drawn the widest support have generally been substantive ones. The ideas of God's will and the human telos, for example, seemed to many to provide successful accounts of morality because they seemed to have the necessary priority and importance." (SCANLON, 1998, p. 151).

Ainsi, contrairement à ce qu'il peut sembler de prime abord, pour expliquer l'importance et la priorité de la morale, il vaut mieux montrer quel est son contenu substantiel plutôt que de tenter de réduire la faute morale à une faute de logique. Dans ce sens, une personne qui rejette la morale ne se contredit pas elle-même, mais elle ignore délibérément quelque chose qui pourtant mérite son attention. Et cette seconde faute est plus grave que la première.

(4) Un internalisme moral modéré

Notre approche permet de défendre une position internaliste modérée dans le débat métaéthique sur l'internalisme *moral* : si S a le devoir moral de faire X, alors il est possible pour S d'être motivé à faire X – ce qui ne signifie pas que S *est motivé* à faire X. Contrairement, toutefois, à l'internalisme moral *empiriste*, où la présence d'un désir fait partie des *conditions* de vérité des énoncés moraux, c'est l'inverse qui est le cas dans l'internalisme moral *moraliste* : l'énoncé « S peut être motivé à faire X indépendamment de ses désirs antécédents » est *rendu vrai* par le fait que S a le devoir moral de faire X indépendamment de ses désirs antécédents.

(5) Une utopie réaliste

Les arguments que nous avons soutenus devraient permettre de développer une philosophie politique *utopiste* de façon *réaliste*, selon les termes de John Rawls.

La philosophie politique est utopiste de façon réaliste lorsqu'elle fait reculer ce que la réflexion ordinaire conçoit comme les limites des possibilités politiques pratiques. (RAWLS, 2006, p. 18)

En effet, nous étions partis de l'idée que les motivations humaines – avec leurs défaillances – représentaient des limites au type de société réalisable. Nous avons découvert que, au final, cela n'est pas le cas. Au contraire, il s'est avéré que les considérations morales, lorsqu'elles ont un contenu *substantiel*, offrent l'occasion d'une motivation *entièrement nouvelle*. De ce fait, une société fondée sur des principes moraux – substantiels, biens définis et connus de tous – n'est pas irréalisable, puisque les considérations morales *étendent* les possibilités motivationnelles. En conséquence, nous avons fait reculer, par notre raisonnement, notre conception des limites des possibilités politiques.

ANNEXES

1. LA MÉTAÉTHIQUE

Le domaine de recherche en philosophie morale est partagé en trois secteurs : l'éthique appliquée, l'éthique normative et la métaéthique. L'éthique appliquée représente le domaine le moins abstrait de la philosophie morale : il vise à élucider des problèmes moraux spécifiques, à en cerner les enjeux et à proposer des solutions. Ce domaine de recherche regroupe ainsi l'éthique environnementale, l'éthique animale, la bioéthique, l'éthique des affaires, mais aussi les éthiques professionnelles, telles que l'éthique médicale, journalistique, juridique, etc.

L'éthique normative représente un domaine de recherche plus général que l'éthique appliquée : son but est de systématiser les règles morales particulières à l'aide de principes moraux généraux. Trois principaux systèmes moraux concurrents sont ainsi comparés relativement à leur capacité respective à systématiser nos évaluations ou intuitions morales particulières: l'utilitarisme soutient que la valeur morale des actions est déterminée entièrement par leurs conséquences sur le bien-être général. Le déontologisme soutient que la valeur morale des actions est déterminée par leur valeur intrinsèque ou par leur conformité à un principe moral fondamental, indépendamment des conséquences. Finalement, l'éthique de la vertu soutient que la valeur des actions provient des dispositions, des traits de caractère dont elles sont des manifestations.

La métaéthique constitue le domaine de recherche le plus abstrait de la philosophie morale. Pour résumer son contenu, nous pouvons dire que le problème ne consiste pas à déterminer *ce qui est bien ou mal, juste ou injuste, vertueux ou vicieux*, mais à déterminer *ce que nous faisons* lorsque nous affirmons que quelque chose est bien ou mal, juste ou injuste, vertueux ou vicieux. Il s'agit d'*expliquer* l'usage que nous faisons des concepts moraux et de déterminer si cet usage correspond ou non à une réalité indépendante de nos attitudes morales. En résumé, la métaéthique est le domaine de la philosophie qui étudie la signification du langage moral, analyse les concepts moraux et élucide leurs présupposés ontologiques et épistémologiques (MILLER, 2013, pp. 2-3).

1.1. LES DIFFÉRENTES POSITIONS DU DÉBAT MÉTAÉTHIQUE CONTEMPORAIN

Dans cette section, nous allons brièvement présenter les principales positions métaéthiques occupant le champ actuel de la discussion. Pour ce faire, nous allons procéder à l'aide d'oppositions, en partant des plus générales pour terminer avec les oppositions les plus subtiles. Ainsi, nous commencerons par exposer (1) l'opposition entre le cognitivisme et le non-cognitivisme, puis (2) entre le réalisme et la théorie de l'erreur, (3) entre l'objectivisme et le non-objectivisme, (4) entre le non-naturalisme et le naturalisme, (5) entre le naturalisme analytique et le naturalisme synthétique et, pour finir, (6) entre le naturalisme synthétique réductionniste et non réductionniste. Nous terminerons l'exposition en proposant un tableau récapitulatif des principales oppositions.

1.1.1. COGNITIVISME VS NON-COGNITIVISME

La principale distinction métaéthique concerne l'opposition entre le cognitivisme et le non-cognitivisme. Selon le cognitivisme, les énoncés moraux possèdent une valeur de vérité. Ils sont vrais ou faux. Par exemple, l'énoncé moral « il est immoral de commettre un adultère » est, selon le cognitivisme, soit vrai, soit faux. Selon le non-cognitivisme, à l'inverse, les énoncés moraux ne possèdent pas de valeur de vérité. Ils ne sont donc ni vrais, ni faux. Il n'est ainsi ni vrai ni faux, pour le non-cognitivisme, que l'adultère est immoral.

De manière corrélative, le cognitivisme et le non-cognitivisme s'opposent en général concernant la nature des états mentaux exprimés dans les énoncés moraux. Dans la très grande majorité des cas, le cognitivisme s'accompagne de la conception des jugements moraux comme étant des *croyances*, c'est-à-dire des états mentaux visant à représenter le monde tel qu'il est. Ainsi, l'énoncé « il est immoral de commettre un adultère » exprime la croyance que l'adultère est immoral. Le cognitivisme peut donc être caractérisé comme suit :

Cognitivism Les énoncés moraux (i) possèdent des conditions de vérité et (ii) expriment des croyances.

Réciproquement, le non-cognitivisme s'accompagne d'une conception des énoncés moraux comme étant l'expression d'états mentaux non-cognitifs. Dire que l'adultère est immoral, ce n'est pas attribuer la propriété « être immoral » à l'adultère, mais c'est exprimer

son sentiment de désapprobation à l'égard de l'adultère, ou sa préférence pour que l'on ne pratique pas l'adultère.

Non-cognitivism Les énoncés moraux (i) sont dépourvus de conditions de vérité et (ii) expriment des *pro-attitudes*.

Dans le non-cognitivism, la signification des énoncés moraux ne provient pas de leurs conditions vérité, contrairement à ce qui est le cas pour les énoncés descriptifs en général. La difficulté pour les non-cognitivistes consiste alors à *expliquer* le sens des expressions morales autrement que par les conditions de vérité – pour éviter de conclure que les expressions morales sont *dépourvues de sens*. En d'autres termes, étant donné que les concepts moraux ne jouent *aucun rôle* descriptif, il faut pouvoir expliquer autrement que par les conditions de vérité le sens des énoncés moraux (ROOJEN van, 2018).

Une première approche du non-cognitivism explique la signification des énoncés moraux par les *actes de langage* que ces énoncés permettent d'accomplir. Il suffit de savoir quel est l'acte de langage que les énoncés moraux permettent d'accomplir pour savoir quelle est la signification de ces énoncés. Pour les émotivistes, par exemple, les énoncés moraux permettent d'exprimer nos émotions ou de susciter des émotions chez autrui. Ainsi, selon eux, l'énoncé « l'avortement est immoral » signifie « Bouh ! l'avortement » et « la charité est moralement bonne » signifie « Hourrah ! la charité » (AYER, 1936; STEVENSON, 1944).

Un autre exemple de cette interprétation du langage moral en termes d'actes de langage est le prescriptivism développé par Hare (1952). Pour les prescriptivistes, les phrases morales ont pour fonction de *prescrire* un comportement, de donner des ordres. Ainsi, selon eux, « l'avortement est immoral » signifie « ne pratiquez pas l'avortement ! ». Le rôle des énoncés moraux est de *donner des ordres* afin de modifier le comportement des autres personnes. Et la *signification* des énoncés moraux est donnée par leur fonction.

Emotivism Le langage moral sert à exprimer des émotions et à susciter des émotions chez autrui.

Prescriptivism Le langage moral sert à diriger les actions et à prescrire un comportement à autrui. Il sert à formuler des *ordres*.

Une seconde approche, plus récente, du non-cognitivism, ne rattache pas la signification des termes moraux aux actes de langage que ces termes permettent d'effectuer. Cette seconde approche identifie la signification des termes moraux aux états mentaux que ces énoncés servent à exprimer. Cette thèse prend le nom « d'expressivisme ». Dans cette approche, pour déterminer la signification des énoncés moraux, il suffit de savoir quelle est l'attitude mentale que ces énoncés expriment, lorsqu'ils sont énoncés sincèrement (JAQUET & NAAR, 2019, p. 34). Par exemple, si « le mensonge est immoral » exprime une attitude de désapprobation du mensonge, alors la signification de l'énoncé « le mensonge est immoral » est la désapprobation du mensonge. Il existe plusieurs formes d'expressivisme en fonction de la nature de l'attitude exprimée par l'énoncé moral. Par exemple, selon Gibbard, un énoncé moral exprime une attitude d'acceptation ou de rejet d'une norme ou d'un ensemble de normes (GIBBARD, 1990).

Expressivisme des normes Un énoncé moral exprime l'acceptation d'une norme ou d'un ensemble de normes.

Les premières versions du non-cognitivism ont rencontré une importante difficulté. Cette dernière consiste dans le fait que le non-cognitivism peine à rendre compte de certaines caractéristiques syntaxiques et logiques du discours moral. Celui-ci possède en effet l'apparence d'un langage *descriptif*; d'abord, les énoncés moraux ressemblent syntaxiquement à des descriptions du monde. En outre, il est possible de construire des raisonnements parfaitement valides à l'aide de jugements moraux. Or, le non cognitivism ne parvient pas à rendre compte de la validité de tels raisonnements. Dans la littérature, ce problème prend le nom du « problème de Frege-Geach », du nom du philosophe qui l'a formulé (GEACH, 1965) et qui lui-même s'appuie sur la théorie de la signification développée par Frege.

Admettons par exemple que nous jugions qu'il est mal de faire du commerce d'armement dans des zones en guerre. Nous savons en outre que Gilles fait du commerce d'armement en Syrie. Etant donné que la Syrie est une zone de conflit, nous concluons que Gilles commet une faute morale. Un tel raisonnement paraît tout à fait valide :

1. Il est mal de faire du commerce d'armement dans des zones de guerre.
2. S'il est mal de faire du commerce d'armement dans des zones de guerre, alors il est mal que Gilles fasse du commerce d'armement dans la zone de guerre syrienne.
3. Il est mal que Gilles fasse du commerce d'armement dans la zone de guerre syrienne.

Le problème pour le non-cognitivism est alors le suivant : si la prémisse « il est mal de faire du commerce d'armement dans des zones de guerre » n'est pas susceptible d'être vraie ou fausse, alors elle ne peut pas être utilisée dans un raisonnement valide, de la même manière que l'expression « waouh » ou « beurk » ne peut pas servir de prémisse pour un raisonnement valide. En outre, le non-cognitivism semble contraint d'attribuer une signification différente aux termes moraux lorsque ceux-ci apparaissent dans des contextes assertifs (prémisse 1) ou dans des contextes non assertifs (prémisse 2). En effet, alors que l'énoncé « il est mal de faire du commerce d'armement » signifie « Bouh ! le commerce d'armement » dans la première prémisse, il n'exprime par contre pas la désapprobation dans le contexte conditionnel de la seconde prémisse. Or, il est admis qu'un énoncé ne change pas de signification dans un contexte assertif ou non assertif, selon la théorie de la signification de Frege.

Simon Blackburn a développé le programme bien connu du *quasi-réalisme* pour tenter de répondre au problème de Frege-Geach. Le but de ce programme consiste à concilier le *projectivisme* moral, une forme de non-cognitivism, avec une explication de l'usage (descriptif) que nous faisons des termes moraux dans le discours ordinaire (BLACKBURN, 1984, p. 180).

Afin de rendre compte notamment de la possibilité des raisonnements intégrant les propriétés évaluatives, Simon Blackburn introduit les opérateurs évaluatifs « Hourrah ! » et « Bouh ! » (H !, B !), afin de montrer que quelqu'un qui admettrait les prémisses du raisonnement précédent mais rejetterait la conclusion se trouverait dans une situation de « clash » entre ses propres attitudes d'approbation et de désapprobation.

Le raisonnement complet pourrait donc être rendu en termes d'attitudes d'approbation et de désapprobation de la manière suivante :

1. B ! (commerce d'armement)
2. H! [B ! (commerce d'armement) ; B ! (commerce d'armement par Gilles)].
3. B ! (commerce d'armement par Gilles)

La seconde prémisse exprime l'approbation (H !) de la combinaison d'une attitude de rejet (B!) à l'égard du commerce d'armement dans les zones de guerre en général et d'une attitude de rejet (B !) à l'égard du commerce d'armement par Gilles en Syrie. Dans ce contexte, quelqu'un qui exprimerait la combinaison des attitudes B ! (commerce d'armement) et H ! (commerce d'armement par Gilles) aurait une conjonction d'attitude que lui-même désapprouve en (2). Il ferait face à une situation de « clash » entre ses diverses attitudes. C'est pourquoi, pour éviter ce « clash », il doit adopter l'attitude exprimée dans la conclusion, à savoir « B ! (commerce d'armement par Gilles) ». Le problème avec cette solution vient du fait que si quelqu'un possède une attitude que lui-même désapprouve, cela ne suffit pas pour qu'il commette une erreur *logique*. En désapprouvant ses propres attitudes, il les condamne *moralement*, mais non pas d'un point de vue *logique* (WRIGHT, 1985).

En conclusion, nous pouvons résumer la dialectique entre le cognitivisme et le non-cognitivisme par la réponse que ces courants apportent à la question suivante : « Les énoncés moraux possèdent-ils des conditions de vérité et expriment-ils des états mentaux cognitifs ? ». Les cognitivistes répondent par l'affirmative à cette question, alors que les non-cognitivistes répondent par la négative. Schématiquement, l'opposition est la suivante :

Les énoncés moraux possèdent-ils des conditions de vérité et expriment-ils des états mentaux cognitifs ?			
OUI Cognitivism	NON Non-cognitivism		
	La signification des énoncés moraux provient de l'acte de langage que ces énoncés permettent d'accomplir		La signification des énoncés moraux provient de l'état mental que ces énoncés expriment
	Emotivism	Prescriptivism	Expressivism des normes Quasi-réalisme

1.1.2. RÉALISME VS THÉORIE DE L'ERREUR

Le non-cognitivismisme affirme que les énoncés moraux expriment des attitudes conatives ou affectives et sont dépourvus de valeur de vérité. Le cognitivismisme, de son côté, admet que les énoncés moraux décrivent le monde et expriment des croyances. Au sein de ce *cognitivismisme*, nous pouvons distinguer deux grandes familles de théories métaéthiques. Selon la première, bien que les énoncés moraux ont pour prétention de décrire le monde, qu'ils possèdent des conditions de vérité et qu'ils expriment des croyances, ces énoncés sont toujours *faux*. On appelle ce type de doctrine la « théorie de l'erreur ». Elle a été défendue notamment par John Mackie (1977) et Richard Joyce (2001).

Théorie de l'erreur Tous les énoncés et jugements moraux sont faux.

Les partisans de la théorie de l'erreur admettent l'ensemble des affirmations suivantes :

« Il est faux que la peine de mort est injuste. »

« Il est faux que le viol est immoral. »

« Il est faux que le courage est une vertu et l'orgueil un vice. »

« Il est faux qu'il faut respecter ses parents. »

« Il est faux que nous devons nous abstenir de manger des chats. »

De la fausseté de ces énoncés, il ne faut pas conclure que les théoriciens de l'erreur admettent les thèses contraires, par exemple que la peine de mort est juste ou que le viol est moralement acceptable. Pour eux, les énoncés « la peine de mort est juste » et « le viol est moralement acceptable » sont tout aussi faux que les énoncés « la peine de mort est injuste » et « le viol est immoral ». N'importe quel énoncé moral est un énoncé faux.

Si ces énoncés sont faux, ce n'est pas parce que leur contraire est vrai, mais parce que ces énoncés attribuent aux choses des propriétés qui n'existent pas dans la réalité. De la même manière que des jugements qui admettent l'existence d'objets qui n'existent pas (« Le Père-Noël est un goinfre », « Ma voisine est une sorcière », « L'actuel roi de France est chauve », « Mon docteur peut soigner mes douleurs par la pensée ») sont tous faux, les énoncés moraux qui attribuent aux actions des propriétés qui n'existent pas, en l'occurrence des propriétés morales *objectives* et *non naturelles*, sont aussi faux. Dire qu'il est mal de

commettre un vol, c'est attribuer une propriété objective et non naturelle à l'action de voler, propriété morale qui n'existe pas en fait.

L'argument en faveur de la théorie de l'erreur est ainsi le suivant:

- a. Les énoncés moraux présupposent l'existence de faits objectifs non naturels.
(Prémisse sémantique)
- b. Il n'existe pas de faits objectifs non naturels – i.e. seuls existent les faits naturels.
(Prémisse ontologique)
- c. $\therefore$ les énoncés moraux sont tous faux.

Le principal défi qui se pose, pour ces doctrines, consiste à déterminer si, étant dans l'erreur, il faut se débarrasser de tout langage moral ou si, bien que faux, le langage moral remplit certaines fonctions, notamment sociales, qui exigent que nous continuions à faire « comme si » les faits moraux existaient. Dans ce contexte, différentes variantes de la théorie de l'erreur se confrontent. Les principales sont les théories de l'erreur *abolitionnistes*, *révisionnistes* et *conservatives*. Les premières soutiennent que nous devons renoncer au langage moral, les secondes que nous devons *réviser* la manière dont nous interprétons le langage moral et les dernières considèrent que nous devons continuer à faire comme si le langage moral décrivait bien des faits qui existent.

A l'inverse de la théorie de l'erreur, d'autres théories métaéthiques cognitivistes considèrent qu'*au moins* certains énoncés moraux et jugements moraux sont vrais. Il arrive donc que les énoncés moraux représentent le monde de façon adéquate. On parle dans ce cadre de *réalisme moral*.

Réalisme moral Certains énoncés moraux sont vrais.

Schématiquement, le débat entre les théoriciens de l'erreur morale et les théoriciens du réalisme moral peut être représenté par le tableau suivant :

Les croyances morales sont-elles parfois vraies ?			
OUI Le réalisme moral	NON La théorie de l'erreur		
	Abolitionniste	Révisionniste	Conservative

1.1.3. OBJECTIVISME VS NON-OBJECTIVISME

Nous avons vu que la théorie de l'erreur nie l'existence de *vérités* morales – tous les énoncés moraux étant *faux* – alors que le réalisme moral affirme l'existence de vérités morales. Il existe néanmoins différentes variantes du réalisme moral.

D'un côté, les théories dites *faiblement réalistes* soutiennent qu'il existe bien des vérités morales, mais que ces vérités ne sont pas *indépendantes* des croyances et attitudes des agents. Au contraire, ce sont précisément les attitudes et croyances des agents (réels ou hypothétiques) qui rendent vraies les croyances morales ou énoncés moraux. Ces théories peuvent être qualifiées de théories réalistes *non objectivistes* puisque les faits qui rendent vrais les énoncés moraux ne sont pas des faits *objectifs*, c'est-à-dire qu'ils ne sont pas des faits qui existent indépendamment des attitudes de certains agents à leur égard.

Réalisme moral non objectiviste Les faits qui rendent vrais les énoncés moraux ne sont pas des faits indépendants des attitudes de certains agents.

Cette perspective non objectiviste englobe les conceptions métaéthiques (i) relativistes, (ii) constructivistes, (iii) subjectivistes et (iv) volontaristes.

Selon (i) le relativisme moral, la vérité et la fausseté des jugements moraux d'une personne sont déterminées par l'acceptation ou le rejet de ces jugements au sein du groupe auquel appartient la personne qui émet le jugement.⁴⁰⁰

Relativisme « X est moralement bon » affirmé par le locuteur S est vrai si et seulement si X est approuvé par le groupe auquel S appartient.

Selon (ii) le *subjectivisme* moral, la vérité ou la fausseté des jugements moraux exprimés par une personne est déterminée par ce que cette personne (seule) juge comme étant acceptable ou inacceptable, ce que cette personne approuve ou désapprouve.

Subjectivisme « X est moralement bon » affirmé par le locuteur S est vrai si et seulement si S approuve X.

Selon (iii) le *constructivisme* moral, les vérités morales sont déterminées par les attitudes *conditionnelles* du locuteur, c'est-à-dire par ce que le locuteur *désirerait faire* s'il se

⁴⁰⁰ Le plus célèbre des relativistes contemporains est sans doute Gilbert Harman (2000 ; 1975)

trouvait dans certaines conditions hypothétiques ou idéales. Le constructivisme soutient donc que les valeurs morales décrivent du point de vue de l'agent (de ses désirs, valeurs, ...) lorsque ce point de vue est soumis à certaines contraintes de rationalité, notamment aux contraintes de cohérence entre les désirs et de vérité des croyances, d'absence de biais, et d'absence de faiblesse de volonté et de dépression, etc. (KORSGAARD, 1996; SMITH, 1994; RAWLS, 2009).

Constructivisme L'énoncé « S a le devoir moral de faire X » est vrai si et seulement si S approuverait X dans des circonstances idéales ou hypothétiques C.

Selon les versions *universalistes* du constructivisme, tous les êtres humains choisiraient d'agir de la même manière s'ils étaient tous pleinement rationnels. Selon les versions *individualistes* du constructivisme, il est possible que plusieurs personnes pleinement rationnelles choisissent d'agir de manière différente, même en conditions idéales.

La dernière variante du non-objectivisme est appelée (iv) le *volontarisme* (juridique, divin ou du spectateur impartial). Elle consiste à affirmer que la valeur de vérité des énoncés moraux est déterminée par la volonté d'un agent indépendant du locuteur. Cet agent peut notamment être l'autorité étatique dans le volontarisme juridique (KELSEN, 1991) ou Dieu dans le volontarisme divin (ADAMS, 1987; MURPHY, 2002; QUINN, 1978) ou encore un spectateur impartial et bienveillant dans la théorie du spectateur impartial (FIRTH, 1952). Dans tous les cas, ce qui est moralement bon ou mauvais est déterminé par ce qu'un agent indépendant du locuteur approuve ou désapprouve. Par exemple, l'énoncé « il est immoral de commettre un adultère » est vrai si et seulement si (a) la loi ou l'autorité légale condamne l'adultère (volontarisme juridique), (b) Dieu désapprouve l'adultère (volontarisme divin) ou (c) le spectateur impartial désapprouve l'adultère (théorie du spectateur impartial).

Volontarisme L'énoncé « il est moral/immoral de faire X » affirmé par S est vrai si et seulement si un agent A (différent de S) approuve/désapprouve X.

Pour les doctrines relativistes, subjectivistes, constructivistes ou volontaristes, la vérité des énoncés moraux n'est pas indépendante des attitudes des agents. Bien plus, la vérité des énoncés moraux est *déterminée* par les attitudes de certains agents. A l'inverse, d'autres théories réalistes soutiennent que la vérité des énoncés moraux est *indépendante* des attitudes de l'agent. On parle cette fois de *réalisme moral fort* ou *objectiviste*.

Réalisme moral objectiviste

Les énoncés moraux sont vrais ou faux indépendamment des attitudes des agents. Les faits moraux ne sont pas des faits au sujet des attitudes des agents.

L'opposition entre le réalisme moral objectiviste et non objectiviste réside donc dans la réponse apportée à la question suivante : « Les faits moraux sont-ils des faits indépendants des attitudes d'un ou plusieurs agents ? ». Le réalisme moral objectif répond par l'affirmative, alors que le réalisme moral non objectif répond par la négative. Schématiquement, à nouveau, le débat peut être décrit de la manière suivante :

Les faits moraux sont-ils des faits indépendants des attitudes des attitudes d'un ou plusieurs agents ?				
OUI Le réalisme moral objectiviste	NON Le réalisme moral non objectiviste			
	Subjectivisme	Relativisme	Constructivisme	Volontarisme

1.1.4. RÉALISME MORAL NATURALISTE VS NON-NATURALISME

Il nous faut distinguer maintenant, parmi les théories morales réalistes objectivistes, les théories *naturalistes* et les théories *non naturalistes*. Selon les premières, les faits moraux qui rendent vrais les énoncés moraux sont des faits naturels, alors que, pour les secondes, les faits moraux qui rendent vrais les énoncés moraux sont des faits non naturels.

Les partisans du *non-naturalisme moral* s'accordent avec les partisans du naturalisme moral sur le fait qu'il existe des faits moraux indépendants des attitudes des agents. Cependant, les premiers nient que ces faits soient naturels, par exemple, que nous puissions les découvrir à l'aide des méthodes des sciences empiriques, alors que les seconds affirment que les faits moraux *sont* naturels.

Selon les non-naturalistes, les faits moraux possèdent des caractéristiques qui empêchent de les considérer comme des faits naturels. Par exemple, selon certains non naturalistes, les faits moraux ne sont pas « localisables » dans le monde spatio-temporel de la même manière que le sont les faits naturels, ils sont éternels et immuables, ils sont dépourvus de pouvoir causal, et ils offrent des raisons *catégoriques* d'agir, c'est-à-dire des

raisons d'agir qui ne dépendent pas des désirs individuels. A l'inverse, les faits naturels sont situés dans le temps et l'espace, exercent un pouvoir causal et offrent uniquement des raisons *hypothétiques* d'agir, c'est-à-dire des raisons d'agir qui dépendent des désirs individuels. Puisque les faits moraux et les faits naturels sont si différents les uns des autres, les non-naturalistes concluent que les faits moraux *ne sont pas* des faits naturels (ENOCH, 2011, p. 80; PARFIT, 2011, pp. 264-377; DANCY, 2006).

La position non naturaliste peut donc se résumer de la manière suivante :

Réalisme moral non naturaliste Les faits moraux objectifs ne sont pas des faits naturels.

A titre de précision, un non naturaliste n'est pas forcé d'admettre que les propriétés morales n'entretiennent *aucun* lien avec les propriétés naturelles ; il affirme uniquement que les propriétés naturelles ne *suffisent pas* pour expliquer les propriétés morales d'un objet.

La conséquence du non-naturalisme moral est le rejet d'une thèse métaphysique que *l'ensemble* des autres positions métaéthiques admet, à savoir la thèse du *naturalisme* métaphysique, selon laquelle seuls existent des faits naturels, c'est-à-dire des faits que nous pouvons étudier par les méthodes des sciences empiriques.

En métaéthique, cette thèse naturaliste se traduit par l'exigence de rendre compte des pratiques morales, du discours moral et éventuellement des faits moraux dans des termes exclusivement naturels. Il faut pouvoir expliquer les pratiques morales à l'aide de faits susceptibles d'être étudiés par les sciences empiriques.

[Pour rendre compte des pratiques morales], il faut que nous restions dans le cadre d'une hypothèse naturaliste [...]. Nous devons faire la supposition que les effets qu'a le monde sur les opinions que nous avons à son sujet, ainsi que les effets que nous avons sur le monde, sont explicables sensiblement de la même manière que n'importe laquelle des autres interactions ayant lieu dans le monde en général, autrement dit, au moyen de la science empirique. Il faut adopter cette hypothèse parce qu'il n'est pas crédible de supposer que les agents ou les objets du monde soient dotés de pouvoirs ou de propriétés relevant du spiritisme, pouvoirs et propriétés qui tiendraient en échec les explications scientifiques (SMITH, 2004, p. 492).

Le non-naturalisme est la seule position métaéthique qui ne respecte pas cette contrainte naturaliste, puisqu'il admet dans son ontologie des entités non naturelles – i.e. les faits moraux. Toutes les autres positions – non-cognitivism, théorie de l'erreur, non-objectivisme et les naturalismes moraux – respectent cette contrainte naturaliste. Le non-respect de la contrainte naturaliste constitue même le principal défaut du réalisme moral non naturaliste.

Nous avons vu que le réalisme moral non naturaliste consiste à soutenir qu'il existe des faits moraux indépendants des attitudes des agents et que ces faits ne sont pas naturels. De leur côté, les partisans du réalisme moral objectif *naturaliste* soutiennent également que les faits moraux sont indépendants des attitudes des agents, mais ils soutiennent, en outre, que ces faits sont des faits naturels.

Un fait naturel est à *minima* traité comme un fait que l'on peut décrire sans recourir à des concepts moraux/normatifs: par exemple, le fait que la guerre cause beaucoup de souffrance est un fait naturel, car il ne contient aucun terme normatif.⁴⁰¹ Cette spécification n'est toutefois pas suffisante, puisqu'il existe des faits que l'on peut décrire sans recours à des termes normatifs mais qui ne sont néanmoins pas des faits naturels. Par exemple, le fait que Dieu est tout puissant peut être décrit sans recourir à des termes moraux ou normatifs, mais il ne s'agit pas d'un fait naturel. Les autres critères proposés sont souvent les suivants : un fait est naturel s'il exerce un rôle causal dans le monde spatio-temporel et/ou un fait est naturel s'il peut être étudié à l'aide des méthodes de la science empirique – i.e. s'il est observable ou s'il explique des phénomènes observables.

Plutôt que de tenter de définir plus précisément les caractéristiques des faits naturels – cette définition est sujette à controverse – nous allons nous contenter de lister des exemples non problématiques de faits naturels possiblement pertinents dans le domaine moral:

⁴⁰¹ Nous rappelons que l'important consiste dans le fait que l'énoncé ne contient pas de concepts normatifs *dans des phrases non subordonnées*. En effet, il est tout à fait possible qu'un terme normatif apparaisse dans une phrase subordonnée sans pourtant que l'énoncé soit considéré comme normatif : par exemple, « les conservateurs américains considèrent que l'avortement est immoral » est un énoncé qui contient un concept normatif. Cependant, le concept normatif apparaît dans une phrase subordonnée, et ne suffit donc pas pour que l'énoncé dans son ensemble soit un énoncé normatif. Bien au contraire, il paraît évident que l'énoncé « les conservateurs américains considèrent que l'avortement est immoral » est un énoncé descriptif naturel étudiable par les sciences empiriques traditionnelles ; il consiste à affirmer que les conservateurs désapprouvent l'avortement, ce que nous pouvons vérifier par l'étude de leur comportement ou par l'étude de leur programme politique.

- (1) Les êtres humains sont motivés à agir par intérêt personnel.
- (2) Le traité de non-prolifération des armes nucléaires est un traité qui favorise la paix mondiale.
- (3) L'organisation de l'Etat Islamique est une organisation qui sème la terreur dans les zones qu'elle occupe.
- (4) Faire de la danse rend Christian heureux.
- (5) Tous les êtres humains désirent le bonheur.
- (6) Manger une glace au chocolat procure du plaisir.

Tous ces énoncés – possiblement pertinents pour l'évaluation morale – ne contenant aucun terme normatif, peuvent être étudiés à l'aide des méthodes des sciences empiriques et jouent un rôle causal explicatif. Ils décrivent donc de manière non controversée des faits naturels.

Etant donné ce qui précède, nous pouvons définir le réalisme moral *naturaliste* selon la thèse suivante :

Réalisme moral naturaliste (i) Tous les faits (y compris les faits moraux) sont des faits naturels et (ii) nous acquérons la connaissance morale de la même manière que nous acquérons la connaissance dans les autres sciences empiriques.

A la question « les faits indépendants des attitudes des agents qui rendent vrais les énoncés moraux, sont-ils des faits naturels ? » les partisans du non-naturalisme moral répondent par la négative alors que les partisans du naturalisme moral répondent par l'affirmative. L'opposition peut donc schématiquement être représentée par le tableau suivant :

Les faits moraux indépendants des attitudes des agents, sont-ils des faits naturels ?	
OUI Naturalisme moral	NON Non-naturalisme

1.1.5. NATURALISME MORAL ANALYTIQUE VS SYNTHÉTIQUE

Parmi les partisans du réalisme moral *naturaliste*, nous pouvons encore distinguer différents courants. Pour certains naturalistes, les concepts moraux peuvent être *définis* en termes exclusivement naturels. Les termes moraux sont *synonymes* de termes naturels. Cette position métaéthique est qualifiée réalisme moral naturaliste *analytique*. Par exemple, un naturaliste analytique pourrait soutenir que l'énoncé «la torture cause de grandes souffrances » a la même signification que «la torture est immorale ».

Le réalisme moral naturaliste *analytique* Les concepts moraux sont des synonymes de concepts naturels.

De cette synonymie entre un terme moral (« être immoral ») et un terme naturel (« causer de la souffrance »), les naturalistes concluent que la propriété d'être immoral est *identique* à la propriété naturelle de causer de la souffrance.

Le réalisme moral naturaliste *analytique* ne constitue pas l'approche métaéthique naturaliste la plus répandue.⁴⁰² La plupart des naturalistes considèrent plutôt que les termes moraux *ne sont pas* synonymes de termes naturels. Toutefois, selon ces derniers, malgré l'absence de synonymie des termes, il reste le cas que les termes moraux font référence à des propriétés naturelles. Autrement dit, les propriétés morales sont identiques à des propriétés naturelles, même si la signification des termes moraux et celle des termes naturels diffèrent. De la même manière que la pensée « ceci est de l'eau » et « ceci est du H₂O » n'est pas la même pensée, quand bien même la propriété d'être de l'eau est identique à celle d'être du H₂O, la pensée « ceci est immoral » et « ceci cause de la souffrance » n'est pas la même pensée, quand bien même la propriété d'être immoral et celle de causer de la souffrance sont identiques.

Cette position est qualifiée de réalisme moral naturaliste *synthétique*, puisque la vérité des énoncés d'identité entre les concepts naturels et les concepts normatifs ne provient pas de la seule *signification* des concepts.

⁴⁰² A l'exception de la théorie du fonctionnalisme moral, défendue notamment par Franck Jackson (1998).

Naturalisme moral synthétique

Les concepts moraux ne sont pas *synonymes* de concepts naturels, mais ces concepts font néanmoins référence à une propriété naturelle.

A la question « les concepts moraux sont-ils synonymes de concepts naturels ? » les partisans du naturalisme moral *analytique* répondent « oui » alors que les partisans du naturalisme moral *synthétique* répondent « non ». Schématiquement, nous obtenons ainsi le tableau suivant :

Les concepts moraux sont-ils synonymes de concepts naturels ?	
OUI Naturalisme moral analytique	NON Naturalisme moral synthétique

1.1.6. NATURALISME MORAL SYNTHÉTIQUE RÉDUCTIONNISTE VS NON RÉDUCTIONNISTE

Il existe encore une ligne de partage *au sein du naturalisme moral synthétique*. D'un côté, les approches *réductionnistes* considèrent que les propriétés morales ou faits moraux sont *réductibles* à des propriétés naturelles *non morales* ou à des faits naturels *non moraux* – physiques, chimiques, biologiques, psychologiques, économiques, etc. Par exemple, si nous disons que « la guerre est immorale », le concept *immoral* fait référence à la propriété de *générer beaucoup de souffrances*, qui est une propriété naturelle *non morale*. Cette approche est réductionniste car elle considère que les *faits moraux* ne sont rien de plus que des faits naturels *non moraux* – ou que les propriétés morales ne sont rien de plus que des propriétés naturelles *non morales*. En ce sens, il n'existe pas à proprement parler de faits *moraux*, mais seulement un *langage moral* décrivant des faits naturels d'une manière spécifique. Cette thèse naturaliste réductionniste est défendue notamment par Peter Railton (1986), Richard Brandt (1979) ou Mark Schroeder (2005).⁴⁰³

Le naturalisme synthétique réductionniste

Les faits moraux sont réductibles à des faits naturels *non moraux*

La seconde approche du réalisme moral naturaliste synthétique est *non réductionniste*. Ses partisans considèrent, contrairement aux partisans du réductionnisme, que les faits

⁴⁰³ Il nous semble, sans certitude, que l'éthique de la vertu devrait également être classée parmi les positions naturalistes réductionnistes (FOOT, 2001; HURSTHOUSE, 1999; GEACH, 1956; THOMPSON, 1995).

moraux ne sont pas identiques à des faits naturels *non moraux* et que les propriétés morales ne sont pas identiques à des propriétés naturelles *non morales*. Selon eux, au contraire, les propriétés morales sont des propriétés naturelles *irréductiblement morales*. Il s'agit de propriétés *morales* naturelles. Les propriétés morales *surviennent* sur des propriétés non morales (psychologiques, biologiques, physiques, ...), mais elles ne sont pas *identiques* à de telles propriétés non morales.

L'idée est que les propriétés morales partagent suffisamment de caractéristiques communes avec ce que nous classifions comme étant des *propriétés naturelles*, notamment par le fait qu'elles sont indispensables pour expliquer certains faits observables. Ces propriétés s'ajoutent néanmoins, dans l'ontologie du monde, aux propriétés naturelles non morales – physiques, chimiques, économiques, etc. Autrement dit, nos meilleures théories explicatives de certains phénomènes observables doivent faire appel à des *propriétés naturelles irréductiblement morales*.⁴⁰⁴

Les principaux défenseurs de cette dernière approche sont appelés les « réalistes de Cornell », dont font partie notamment Richard Boyd (1996), David Brink (1989) et Nicholas Sturgeon (1988).

Le naturalisme non réductionniste

Les faits moraux sont des faits naturels
irréductiblement moraux

A la question « les faits moraux sont-ils réductibles à des faits naturels *non moraux* ? », les partisans du naturalisme moral réductionniste répondent par l'affirmative, alors que les partisans du naturalisme moral non réductionniste répondent par la négative. Par un tableau, nous pouvons représenter le débat de la manière suivante :

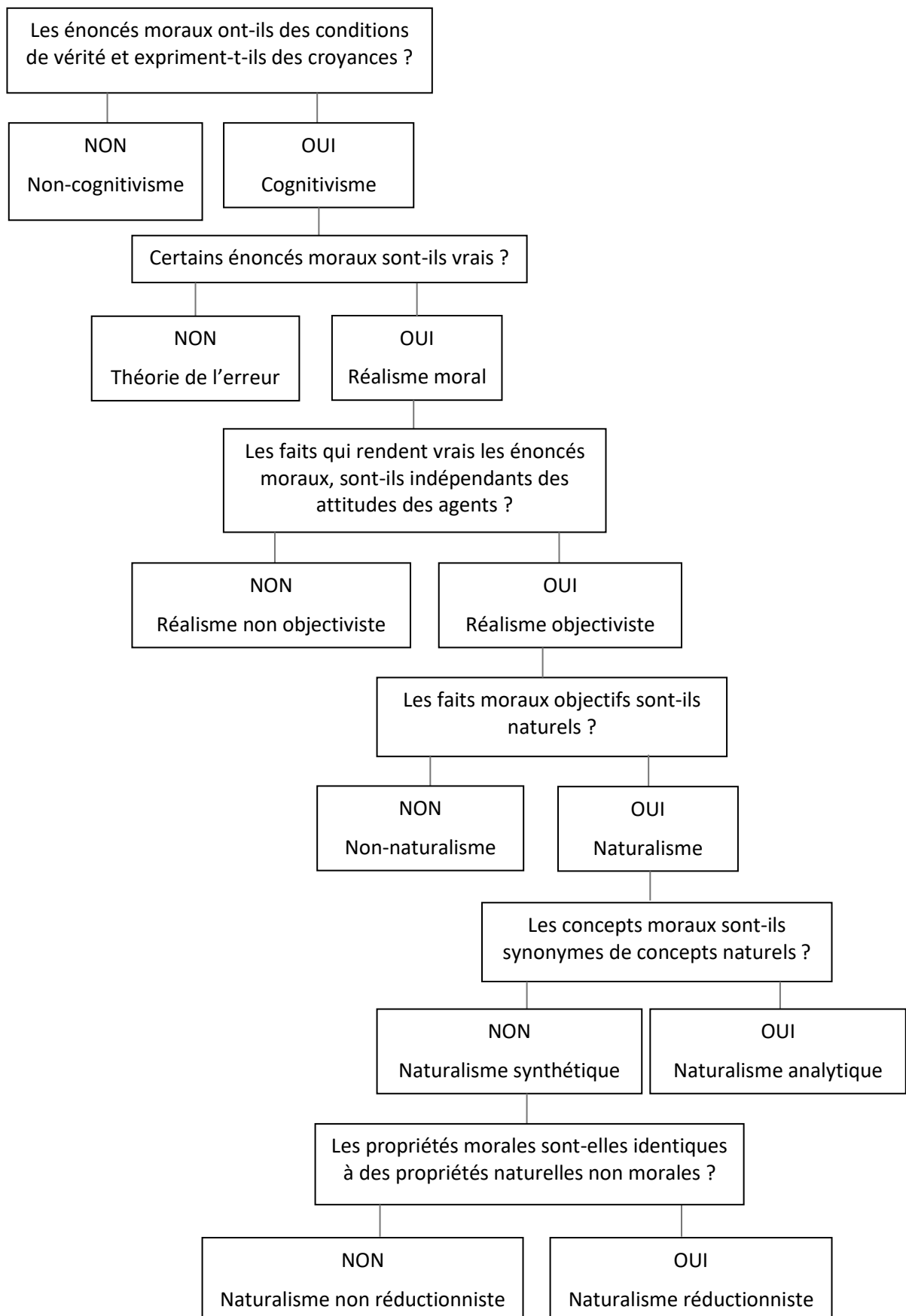
⁴⁰⁴ Il n'est pas évident de distinguer la doctrine du réalisme naturaliste *non réductionniste* de la doctrine du réalisme moral *non naturaliste*. La principale différence provient, selon nous, du rôle causal que jouent les propriétés morales ou les faits moraux dans les deux doctrines. Les premiers attribuent des pouvoirs causaux aux *faits moraux* alors que les seconds attribuent des pouvoirs causaux aux *états mentaux*, à savoir aux croyances morales, qui représentent ces faits moraux. La doctrine de David Enoch est à ce titre représentative de la conception non naturaliste de l'explication de l'action : « La manière dont les faits normatifs – tels que le fait que R est une raison pour A de ϕ (dans ces circonstances) – expliquent l'action n'est pas qu'ils *causent* l'action, mais plutôt qu'ils sont ce en vue de quoi l'agent agit – dans les termes de notre approche, ils sont le contenu d'une croyance qui joue le rôle causal approprié ». [“(T)he way in which normative facts – such as that R is a reason for A to ϕ (in the circumstances) – explain action is not by *causing* action, but rather by being that in light of which an agent acts – in terms of the account here, by being the content of a belief that plays the appropriate causal role.” (ENOCH, 2011, p. 233)].

Les faits moraux sont-ils identiques à des faits naturels non moraux ?	
OUI Naturalisme moral synthétique réductionniste	NON Naturalisme moral synthétique non réductionniste

1.1.7. TABLEAU DES DOCTRINES MÉTAÉTHIQUES

Afin de résumer l'exposition des différentes positions métaéthiques que nous avons proposée dans les sous-sections précédentes, nous proposons d'inscrire l'ensemble des positions dans un schéma commun, en fonction des réponses aux principales questions métaéthiques.

Tableau résumé des différentes positions métaéthiques



2. LES NÉGATIONS DE NORMES SONT-ELLES DES NORMES ?

Il est possible de contester que la conclusion d'un argument fondé sur DIP soit un véritable énoncé normatif. Selon cette objection, un énoncé du type « S n'a pas l'obligation de faire X » n'est pas un véritable énoncé normatif ; seuls les énoncés normatifs *positifs* sont de véritables énoncés normatifs. Autrement dit, selon l'objection, la négation d'une norme n'est pas elle-même une norme.

Du fait que la négation d'une norme n'est pas elle-même une norme, l'objection soutient que DIP ne permet pas d'inférer de véritables conclusions normatives. Ainsi, dans le raisonnement suivant, la conclusion n'est pas considérée comme normative :

- (1) Les êtres humains sont motivés à agir exclusivement par intérêt personnel
- (2) Aider un étranger dans le besoin ne permet pas de satisfaire l'intérêt personnel
- (3) Si les êtres humains ne peuvent pas être motivé à faire X, alors ils n'ont pas le devoir moral de faire X.
- (4) ∴ les êtres humains n'ont pas le devoir moral d'aider un étranger dans le besoin.

Etant donné que (4) est une négation de norme, alors (4) ne constitue pas un véritable énoncé normatif. DIP ne permettrait donc pas, selon l'objection, de tirer des conclusions normatives sur la base de prémisses exclusivement descriptives.

Nous rejetons cette objection. Selon nous, les négations de normes ou d'obligation sont tout autant des énoncés normatifs que des affirmations de normes ou d'obligation. Il nous faut cependant appuyer notre position en deux temps. D'abord, (A) nous donnerons des raisons de considérer les négations de normes comme étant elles-mêmes des normes. Nous verrons ensuite (B) une objection importante qui peut être avancée à l'encontre de nos arguments. Nous tenterons de répondre à cette objection en distinguant deux niveaux de discours.

2.1. L'INTER-DÉFINISSABILITÉ DES NOTIONS NORMATIVES ET LES NÉGATIONS DE NÉGATION

Deux considérations formelles doivent nous inciter à considérer les négations de normes comme étant elles-mêmes des normes. En premier lieu, les notions normatives peuvent être définies les unes par rapport aux autres, ce qui implique que, pour toute

négarion d'obligation, il existe un énoncé normatif positif correspondant. Ainsi, la négation d'une norme peut être définie en termes de permission. En effet, $\neg OA$ se définit par $P \rightarrow A$. De ce fait, la négation « les êtres humains n'ont pas le devoir moral d'aider un étranger dans le besoin » peut se traduire par « les êtres humains ont la permission de ne pas aider un étranger dans le besoin ». Or, le second énoncé est un énoncé normatif *positif*, attribuant à une action une propriété normative. Cela montre que les négations de normes sont bien des énoncés normatifs, puisqu'elles correspondent à des énoncés normatifs positifs. L'exemple de raisonnement précédent, fondé sur DIP, nous amènerait donc à une conclusion normative de la façon suivante :

- (1) Les êtres humains sont motivés à agir exclusivement par intérêt personnel
- (2) Aider un étranger dans le besoin ne permet pas de satisfaire l'intérêt personnel
- (3) Si les êtres humains ne peuvent pas être motivé à faire X, alors ils n'ont pas le devoir moral de faire X.
- (4) $\therefore$ les êtres humains ont la permission de ne pas aider un étranger dans le besoin.

Bien entendu, il est possible de contester que l'attribution d'une permission soit un véritable énoncé normatif puisque, possiblement, il n'indique pas de manière dont il faut se comporter. En effet, s'il est permis à S de faire X, cela n'implique pas que S doive faire X, ou que S fasse quelque chose de mal s'il ne fait pas X, ni même qu'il serait bon que S fasse X. Autrement dit, le fait qu'il soit permis à S de faire X ne contraint aucunement S à modifier le monde d'une manière ou d'une autre. Il ne s'agit donc pas d'un énoncé normatif proprement dit.

Toutefois, les énoncés de permission ont pour effet d'imposer des obligations à *des tierces personnes*. Ainsi, s'il est *permis* à un individu S1 de faire X, cela implique qu'un autre individu S2 a l'interdiction d'empêcher S1 de faire X. Par exemple, si S1 a la permission de vendre ses propres organes, alors S2 a l'interdiction d'empêcher S1 de vendre ses propres organes – i.e. S2 a l'obligation morale de ne pas empêcher S1 de vendre ses propres organes. Nous retrouvons dans ce contexte un énoncé d'obligation positif. Cela prouve que les énoncés de permission *permettent bien* de prescrire certains comportements, puisqu'ils impliquent des obligations pour des tierces personnes.

À notre réponse, il est possible néanmoins d'objecter ceci : s'il est permis de faire tout ce qu'il n'est pas obligatoire de faire, alors toute incapacité d'agir implique une permission.

En effet, puisqu'une incapacité de faire implique une négation d'obligation, et qu'une négation d'obligation implique à son tour une permission, alors toute incapacité de faire implique une permission. Comme le constate von Wright :

Si nous acceptons la vue selon laquelle une permission est une simple absence ou un manque d'interdiction, alors il est clair qu'il existe un grand nombre de choses qu'un individu ne peut pas faire mais qu'il a la permission de faire. En fait, toute chose qu'un agent ne peut pas faire, il a la permission de le faire.⁴⁰⁵

L'être humain étant incapable de sauter plus haut que la lune, il a la permission de ne pas sauter plus haut que la lune ; étant incapable de poursuivre autre chose que son propre intérêt, il a la permission de poursuivre son propre intérêt ; étant incapable de devenir une tortue, il a la permission de ne pas devenir une tortue, etc. Chaque incapacité donne lieu à une permission, suscitant ainsi une explosion de vérités normatives. Cette vue implique qu'il n'existe pas d'action *dépourvue* de propriété normative. L'ensemble du monde pratique est normativement chargé.

Pour certains auteurs, cette conséquence est inacceptable ; selon eux, il existe des actions *moralement neutres*, comme se laver les dents. De telles actions sont simplement *dépourvues* de propriété normative. Pour d'autres auteurs, cette conséquence n'est pas inacceptable : même se laver les dents est une action moralement chargée, puisqu'il semble bien immoral d'empêcher une personne de se laver les dents. Nous ne rentrons pas dans ce débat, mais indiquons seulement que notre position implique que la morale est partout : chaque action possède une propriété morale.⁴⁰⁶

Une deuxième considération formelle qui doit nous inciter à concevoir les négations de normes comme étant elles-mêmes des normes provient du traitement des *négations de négations* de normes. En effet, selon la logique classique, une double négation est équivalente à une affirmation positive. Cela implique qu'une double négation d'une norme est elle-même une norme. Par exemple, $\neg\neg$ (S a le devoir moral de faire X) implique que S a

⁴⁰⁵ "If we accept the view that permission is mere absence or lack of prohibition, then it is clear that there are any number of things which one is permitted to do but which one cannot do. In fact, anything which an agent cannot do, he would then be permitted to do" (von WRIGHT, 1963, p. 108).

⁴⁰⁶ Pour une présentation et discussion des différentes positions sur ce thème, se référer au livre de Samuel Scheffler, *Human Morality* (1992). Scheffler lui-même soutient, comme nous, que la morale est *omniprésente* (*pervasive*).

le devoir moral de faire X. Cependant, il est étrange de dire qu'une négation de norme n'est pas une norme, alors qu'une double négation de norme est bien une norme. De la même manière, il serait étrange de dire que, lorsque nous répondons à la question « S a-t-il le devoir moral de faire X ? », nous affirmons quelque chose de normatif si nous répondons par l'affirmative et quelque chose de non normatif si nous répondons par la négative (OGIEN, 2001, pp. 16-17).

2.2. LA COMPATIBILITÉ AVEC LE NIHILISME MORAL

Mark Nelson conteste toutefois notre conclusion, selon laquelle les énoncés exprimant des négations d'obligations sont eux-mêmes des énoncés d'obligation. Son argument est le suivant : selon la doctrine métaéthique de l'erreur morale, défendue notamment par Mackie (1977), tous les énoncés normatifs sont faux. Il est ainsi faux que Georges a le devoir moral de rendre visite à sa grand-mère, il est faux qu'il est moralement inacceptable d'abandonner ses propres enfants et il est faux que François a la permission morale de vendre ses propres organes. La théorie de l'erreur morale implique donc, selon Mark Nelson, que toute *négation* d'énoncé normatif est vraie. Or, cela est problématique pour notre propos : en effet, si la vérité de l'énoncé « il n'est pas le cas que les êtres humains ont l'obligation morale de venir en aide à un étranger dans le besoin » est compatible avec la théorie de l'erreur morale, cela signifie qu'il est possible que l'énoncé *soit vrai* alors qu'il n'existe aucune propriété morale. Or, selon Nelson, un énoncé qui peut être vrai même si les propriétés normatives n'existent pas n'est pas un énoncé proprement normatif.

Une proposition n'est pas une proposition normative (*Ought-proposition*) si elle ne contient aucun engagement moral, i.e. si elle n'affirme pas qu'une action est juste ou injuste, vertueuse ou vicieuse, etc. Une manière de préciser cette idée est de dire [...] qu'une proposition ne contient pas d'engagement moral si elle est compatible avec le nihilisme moral ou la théorie de l'erreur morale.⁴⁰⁷

Notre réponse sur ce point est moins aisée. En effet, nous sommes forcés d'admettre que la logique propositionnelle ne permet pas de distinguer les cas où la négation d'un

⁴⁰⁷ "[A] proposition is not an Ought-proposition if it carries no moral commitment, i.e., if it does not claim that some action is right or wrong, virtuous or vicious, etc. One way of spelling this out is to say [...] that a proposition carries no moral commitment if it is compatible with moral nihilism or an error theory of ethics." (NELSON, 1995, pp. 557-558).

énoncé normatif est vraie en vertu de la vérité de la théorie de l'erreur et les cas où la négation d'un énoncé normatif est vraie en vertu de l'existence d'une permission correspondante. En effet, nous ne pouvons pas distinguer la situation où $\neg OA$ est rendue vraie par le nihilisme moral et la situation où $\neg OA$ est rendue vraie par le fait que A est une action *permise*. La différence entre les deux situations ne peut pas être expliquée au niveau de la logique propositionnelle de premier ordre. Selon nous, il est nécessaire de faire appel à la logique des prédicats pour saisir la différence entre les deux situations. Plus encore, nous pensons que la thèse de la théorie de l'erreur ne peut être formulée adéquatement qu'à l'aide d'énoncés *de second ordre* au sujet de la vérité ou fausseté d'énoncés de premier ordre.⁴⁰⁸

Plus concrètement, il nous faut distinguer la vérité des énoncés normatifs négatifs et la vérité de la théorie de l'erreur morale de la manière suivante :

La théorie de l'erreur morale $\neg(\exists a)(Oa)$

Énoncé normatif négatif $(\exists a)(\neg Oa)$

L'avantage de différencier ces deux manières de formaliser l'affirmation que « A n'est pas obligatoire » est que leurs implications nous permettent de caractériser de manière appropriée à la fois le nihilisme moral et l'inter-définissabilité des termes moraux que nous trouvons en logique déontique – avec l'implication souhaitable qu'une négation d'obligation reste un énoncé normatif. En effet, selon cette approche :

$$\neg(\exists a)(Oa) \text{ implique } \neg(\exists a)(P-a)$$

Cette implication est cohérente avec le nihilisme moral, selon lequel *n'importe quel énoncé normatif* est faux, peu importe qu'il concerne une obligation ou une permission ou une interdiction. Et la négation de la négation de l'énoncé, à savoir $\neg\neg(\exists a)(Oa)$, implique simplement la fausseté de la théorie de l'erreur. Nous voyons donc que c'est bien à ce niveau que la théorie de l'erreur doit être formulée ; peu importe la nature de l'attribution de propriétés normatives au sein d'un énoncé normatif, cet énoncé sera toujours faux.

⁴⁰⁸ Le fait que la théorie de l'erreur morale ne puisse être formulée que comme une théorie de second ordre n'est pas étonnant : cela montre simplement que la doctrine de l'erreur morale est une doctrine *métaéthique* alors que les énoncés normatifs négatifs appartiennent à l'éthique normative.

Notre approche permet également de rendre compte du fait que la vérité d'une négation d'énoncé normatif n'est pas compatible avec la théorie de l'erreur morale, du fait que, par inter-définissabilité des termes déontiques, elle implique la vérité d'autres énoncés normatifs positifs. En effet, selon notre approche :

$$(\exists a)(\neg Oa) \text{ implique } (\exists a)(P-a)$$

Cette implication est compatible avec la logique déontique, selon laquelle l'absence d'obligation signifie quelque chose de véritablement normatif, à savoir la permission de ne pas faire A. Comme il est évident dans ce formalisme, les énoncés normatifs négatifs ne sont pas compatibles avec la négation de l'énoncé existentiel *dans son ensemble*. Leur vérité est donc incompatible avec le nihilisme moral.

En outre, dans ce formalisme, la double négation de l'énoncé normatif se caractérise comme suit : $(\exists a)(\neg\neg Oa)$. Et cette double négation implique l'obligation de faire A, qui est un énoncé proprement normatif.

2.3. CONCLUSION

Ainsi, en formalisant la théorie de l'erreur au niveau approprié, c'est à dire en tant que théorie *de second ordre* portant sur les énoncés normatifs de premier ordre, il devient évident que l'affirmation d'un énoncé normatif *négatif* n'est pas compatible avec la théorie de l'erreur. Cela prouve, encore une fois, qu'un énoncé moral *négatif* est un énoncé moral à part entière. En conclusion, le principe DIP permet bien d'obtenir une conclusion normative sur la base de prémisses purement descriptives.

3. LEXIQUE

Définition de l'altruisme	L'acteur A est altruiste seulement si, par son action, il vise à satisfaire les intérêts de l'acteur B, à réaliser le bien de l'acteur B, lorsque $A \neq B$.
Définition de l'égoïsme	L'acteur A est égoïste seulement si, par son action, il vise à satisfaire ses propres intérêts, à réaliser son propre bien.
Définition de la motivation morale	L'acteur A dispose d'une motivation proprement morale seulement si, par son action, il vise à réaliser le bien, à faire ce qui est juste ou à être une personne vertueuse.
DIP-Général	Si les êtres du type T ont le devoir moral de faire X, alors il est possible pour les êtres du type T de faire X.
DIP-motivation	Nécessairement, si S a le devoir moral de faire X, alors peut être motivé à faire X
DIP-motivation-général	Si les êtres du type T ont le devoir moral de faire X, alors il est possible pour les êtres du type T d'être motivés à faire X.
Externalisme du jugement moral	(a) Il est faux que: nécessairement, si S juge que X est une action moralement bonne, alors S est motivé à faire X et que si S juge que X est immorale, alors S est motivé à renoncer à faire X. (b) Pour que S soit motivé à faire X, S doit avoir un désir – indépendant du jugement moral – que l'action X permet de satisfaire.
Internalisme du bien « affectif » de S	Une entité X est « un bien affectif » pour S si et seulement il est approprié pour S de se soucier ou d'accorder de la valeur à X – étant donné que X présente un effet positif sur l'état affectif de S.
Internalisme du bien « pratique » de S	Une entité X est « un bien pratique » pour S si et seulement si X est l'objet approprié de la motivation intéressée de S – étant donné que X favorise l'intérêt personnel de S.

Internalisme du bien personnel	Quelque chose est bon pour une personne S seulement si l'objet présente une certaine relation R à la motivation ou à l'état affectif de S.
Internalisme moral	Nécessairement, si S a le devoir moral de faire X, alors il existe dans l'ensemble motivationnel de S un désir que l'action X permet de satisfaire.
L'accusation d' « erreur naturaliste »	Il n'existe pas de relation <i>conceptuelle</i> entre des concepts moraux et des concepts descriptifs. Les « principes sémantiques de transition » sont donc tous <i>faux</i> . Toute doctrine qui prétend qu'il existe une telle relation analytique est <i>fausse</i> et commet une « erreur naturaliste ».
L'argument de la possibilité de l'amoraliste	Il est possible qu'il existe un individu qui juge que X est une action moralement bonne/obligatoire, mais qui n'est pas motivé à faire X – même si l'individu est psychologiquement normal, rationnel, et que sa « vision » morale fonctionne correctement.
L'internalisme conditionnel du jugement moral	Nécessairement, si le sujet S juge que l'action X est moralement bonne, alors S est motivé à faire X, et si S juge que l'action X est immorale, alors S est motivé à renoncer à faire X, <i>toutes choses égales par ailleurs</i> – i.e. en l'absence de troubles psychologiques, de faiblesse de la volonté et de « troubles de la perception morale ».
L'internalisme du jugement moral	Nécessairement, si le sujet S juge que l'action X est moralement bonne, alors S est motivé à faire X et si S juge que l'action X est immorale, alors S est motivé à renoncer à faire X.
La loi de Hume au niveau logique	Il est logiquement interdit de dériver une conclusion normative (au sujet de ce qui devrait être le cas) à partir de prémisses exclusivement descriptives (au sujet de ce qui est le cas). Toute doctrine morale qui viole cette interdiction logique commet le <i>sophisme naturaliste</i> (descriptiviste): elle établit un passage de

	l'être au devoir-être qui est fallacieux (CLAVIEN, 2007, p. 223).
La contrainte humaniste	Une théorie politique normative est défectueuse et donc fausse si elle impose des standards ou exigences qui ignorent la nature humaine (ESTLUND, 2011, p. 208).
La version contraposée de DIP	Nécessairement, si S ne peut pas faire X, alors il n'est pas le cas que S a le devoir moral de faire X.
La version contraposée de DIP-motivation	Nécessairement, si S ne peut pas être motivé à faire X, alors il n'est pas le cas que S a le devoir moral de faire X.
La version contraposée de l'internalisme moral	Nécessairement, s'il n'existe pas dans l'ensemble motivationnel de S un désir que l'action X permet de satisfaire, alors il n'est pas le cas que S a le devoir moral de faire X.
Principe « devoir-implique-pouvoir » (DIP)	Nécessairement, si S a le devoir moral de faire X, alors S peut faire X
Principe sémantique de transition	Principe d'après lequel il existe une relation analytique entre un concept descriptif (naturel) N et un concept normatif (moral) M

BIBLIOGRAPHIE

- ADAMS, R. (1987). *The Virtue of Faith and Other Essays in Philosophical Theology*. Oxford: Oxford University Press.
- ALVAREZ, M. (2007). The Causalist/Anti-causalist Debate in the Theory of Action: What It Is and Why It Matters. Dans A. LEIST, *Action in Context* (pp. 103-123). Berlin: De Gruyter.
- ANDREONI, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, 100(401), pp. 464-477.
- ANSCOMBE, E. (1957). *Intention*. Oxford: Basil Blackwell.
- ANSCOMBE, E. (1958). Modern Moral Philosophy. *Philosophy*, 33(124), pp. 1-19.
- ANSCOMBE, E. (1976). Contraception and Chastity . Dans M. D. BAYLES, *Ethics and Population*. Cambridge, MS: Schenkman .
- APEL, K.-O. (1996-1998). *Discussion et responsabilité*. (C. Bouchindhomme, Trad.) Paris: Cerf.
- ARNSPERGER, C., & VAN PARIJS, P. (2000). *Ethique économique et sociale*. Paris: La Découverte.
- ARROW, K. (1950). A Difficulty in the Concept of Social Welfare. *Journal of Political Economy*, 58(4), pp. 328-346.
- ARROW, K. (1992). Moral Thinking and Economic Interaction. Dans *Social and Ethical Aspects of Economics: A Colloquium in the Vatican* (pp. 17-23). Vatican City: Pontifical Council for Justice and Peace.
- AUDI, R. N. (2001). A Kantian Intuitionism. *Mind*, 439, pp. 601-635.
- AUSTIN, J. L. (1956 - 1957). A Plea for Excuses: The Presidential Address. *Proceedings of the Aristotelian Society*, 57, pp. 1-30.
- AVI-YONAH, R. S. (2002). Review: Why Tax the Rich? Efficiency, Equity and Progressive Taxation. *The Yale Law Journal*, 111(6), pp. 1391-1416.

- AYER, A. J. (1936). *Language, Truth and Logic* . New York: Dover .
- BAYERTZ, K. (2004). *Warum Überhaupt Moralisch Sein?* München, Deutschland : C.H. Beck Verlag .
- BECK, L. W. (1966). Conscious and Unconscious Motives. *Mind*, 75(298), 155-179.
- BECK, L. W. (1974). "Was-Must Be" and "Is-Ought" in Hume. *Philosophical Studies* , 26(3/4).
- BEE, M. ((à paraître)). The Plesure of Exchange: Adam Smith's Third Kind of Self-Love. *Journal of the history of economic thought*.
- BÉLANGER, M. (2011). *Existe-t-il des dilemmes moraux insolubles?* . Paris: L'Harmattan .
- BENTHAM, J. (1996). *The Collected Works of Jeremy Bentham: An Introduction to the Principles of Moraly and Legislation*. Oxford : Clarendon Press.
- BERKER, S. (2014). Does Evolutionary Psychology Show that Morality is Mind-Dependent? Dans J. D'ARMS, & D. JACOBSON, *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics* (pp. 215-252). Oxford: Oxford University Press.
- BETTKE, K., & NYBORG, K. (2004). Moral hazard and moral motivation: Corporate social responsibility as labor market screening. *University of Oslo working paper*, 25.
- BETTKE, K., KVERNDORCK, S., & NYBORG , K. (2003). An economic model of moral motivation. *Journal of Public Economics*, 87(9-10).
- BLACKBURN, S. (1984). *Spreading the Word*. Oxford: Oxford University Press.
- BLOOMFIELD, P. (2006). Two Dogmas of Metaethics. *Philosophical Studies*, 132, pp. 439-466.
- BOATRIGHT, J. (1999). Does Business Ethics Rests on a Mistake ? *Business Ethics Quarterly* , 9.
- BOETTKE, P. (2012). *Living Economics: Yesterday, Today, Tomorrow*. Oakland: Independent Institute .
- BOUDON, R. (1977). The Logic of Relative Frustration . Dans R. BOUDON, *The Unintended Consequences of Social Action* . London: Palgrave Macmillan.
- BOWLES, S. (2008, June). Policy Designed for Self-Interested Citizens May Undermine "The Moral Sentiments". *Science*, 320 , 1605-1609.

- BOYD, R. (1996). How to be a Moral Realist". Dans S. a. DARWALL, *Moral Discourse and Practice: Some Philosophical Approaches* (pp. 105-137.). New York: Oxford University Press.
- BRANDT, R. (1979). *A Theory of the Good and the Right*. Westminster, MD: Prometheus Books.
- BRENNAN, G. (1993). Economics. Dans P. PETTIT, & R. E. GOODIN, *A Companion to Contemporary Political Philosophy* (pp. 123-157). Oxford, Cambridge Mass.: B. Blackwell.
- BREWER, S. (1993). Quelques raisonnements théoriques sur des raisonnements pratiques a propos du raisonnement théorique. Dans P. CHANGEUX, *Fondements Naturels de l'Ethique*. Paris: Odile Jacob.
- BRINK, D. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- BRINK, D. (1997). Moral Motivation. *Ethics*, 108(1), pp. 4-32.
- BROAD, C. D. (1949). Egoism as a Theory of Human Motives. *Hibbert Journal*, 48, pp. 105-114.
- BRUNI, L., & SUGDEN , R. (2013). Reclaiming Virtue Ethics for Economics. *Journal of Economic Perspectives*, 27(4), pp. 141-164.
- BUCHANAN, J., & TULLOCK, G. (1965). *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press.
- BÜRKI, M. (2012). L'éthique paie. *Swissquote: ePrivate banking Magazine*(3).
- BURTT, S. (1990). The Good Citizen's Psyche: On the Psychology of Civic Virtue. *Polity*, 23(1), pp. 23-38.
- CAILLÉ, A., LAZZERI, C., & SENEL, M. (2001). Introduction générale . Dans A. CAILLÉ, C. LAZZERI, & M. SENEL, *Histoire raisonnée de philosophie morale et politique* . Paris : La Découverte.
- CALORI, F. (2011). "Laut Denken": de la transparence chez Kant . *Raison Publique* .
- CHAUVIER, S. (2013). *Ethique sans visage: le problème des effets externes* . Paris : Vrin .

- CIMA, M., TONNAER, F., & HAUSER, M. (2010). Psychopaths Know Right from Wrong but don't Care. *Social cognitive and affective neuroscience*, 5(1), pp. 59-67.
- CLAVIEN, C. (2007). Comment les données scientifiques et les théories évolutionnistes transforment l'éthique normative. Dans C. CLAVIEN, & C. EL-BEZ, *Morale et évolution biologique ; entre déterminisme et liberté* (pp. 220-244). Lausanne: PPUR.
- CLAVIEN, C. (2012). Nouvelle Sociologie et Morale Naturalisée. *SociologieS*.
- CLAVIEN, C., & CHAPUISAT, M. (2016). The evolution of utility functions and psychological altruism. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 56, pp. 24-31.
- COHON, R. (2006). Hume on Promises and the Peculiar Act of the Mind . *Philosophy Faculty Scholarship* , 13 .
- CUDD, A., & EFTEKHARI, S. (2018). Contractarianism. (E. N. ZALTA, Éd.) *The Stanford Encyclopedia of Philosophy*.
- DAMASIO, A. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New-York: Avon Books.
- DANCY, J. (1993). *Moral Reasons*. Oxford: Basil Blackwell.
- DANCY, J. (2000). *Practical Reality*. Oxford : Clarendon Press .
- DANCY, J. (2006). Nonnaturalism. Dans D. COPP, *The Oxford Handbook of Ethical Theory* (pp. 122-145). Oxford: OxfordUniversity Press.
- D'ARMS, J., & JACOBSON, D. (2000). The Moralistic Fallacy: On the Appropriateness of Emotions. *Philosophy and Phenomenological Research*, 61(1), pp. 65-90.
- DASGUPTA, P. (2011). Modern Economics and the Idea of citizenships. Dans *Social and Ethical Aspects of Economics* (éd. 2nd, pp. 47-58). Vatican City: Pontifical Council for Justice and Peace.
- DAVIDSON, D. (1963). Actions, Reasons and Causes. *The Journal of Philosophy*, 60(23), pp. 685-700.
- DENNETT, D., & STEGLICH-PETERSEN, A. (2008). Blackburn. Dans D. DENNETT, & A. STEGLICH-PETERSEN, *The Philosophical Lexicon*.

- DESCARTES, R. (1897). Lettre à Mersenne du 25 décembre 1639. Dans R. DESCARTES, & A. e. Tannery (Éd.), *Oeuvres* (Vol. II). Paris: Leopold Cerf.
- DIDEROT, D. (1963). *Oeuvres politiques*. Paris : Garnier.
- DIETSCH, P., & CLAVEAU, F. (2008). Concurrence fiscale et responsabilité étatique. *Ethique Publique*, 10(1).
- DIETSCH, P., & RIXEN, T. (2014). Tax Competition and Global Background Justice. *The Journal of Political Philosophy*, 22(2), pp. 150-177.
- DORIS, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge : Cambridge University Press.
- DORIS, J., STICH, S., PHILLIPS, J., & WALMSLEY, L. (2015). *Moral Psychology: Empirical Approaches*. (E. N. ZALTA, Éd.) Consulté le 08 22, 2019, sur Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/archives/win2017/entries/moral-psych-emp/>
- DUBBINK, W. (2015). A Moral Grounding of the Duty to Further Justice in Commercial Life. *Ethical Theory and Moral Practice*, 18, pp. 27-45.
- DUBBINK, W., & VAN DE VEN, B. (2012). On the Duties of Commission in Commercial Life. A Kantian Criticism of Moral Institutionalism. *Ethical Theory and Moral Practice*, 15, pp. 221-238.
- DUBBINK, W., & van LIEDEKERKE, L. (2019). Rethinking the Purity of the Moral Motives in Business: Kant Against MORal Purism. *Journal of Business Ethics*.
- Economist, T. (1992, February 8). Let Them Eat Pollution. *The Economist*, p. 66.
- ECONOMIST, T. (1992, February 8). Let Them Eat Pollution. *The Economist*, p. 66.
- ELSTER, J. (1978). *Logic and Society: Contradictions and Possible Worlds*. Chichester; New-York; Brisbane; Toronto : John Wiley and Sons.
- ELSTER, J. (2009). *Le Désintéressement: traité critique de l'homme économique (vol. 1)*. Paris: Seuil.
- ELSTER, J. (2010). *L'Irrationalité: traité critique de l'homme économique (vol. 2)*. Paris: Seuil.
- ENOCH, D. (2011). *Taking Morality Seriously: A Defense of Robust Realism* (éd. Oxford:). Oxford : Oxford University Press.

- ESLTUND, D. (1990). Mutual Benevolence and the Theory of Happiness. *The Journal of Philosophy*, 87(4), pp. 187-204.
- ESTLUND, D. (2011). Human Nature and the Limits (If Any) of Political Philosophy. *Philosophy and Public Affairs*, 39(3), pp. 207-237.
- EXPRESS, L. (2019, 04 29). *Visqueux, baveux, repoussants... Ces "animaux moches" qu'on peine à protéger*. (L'Express) Consulté le 08 23, 2019, sur [lexpress.fr: https://www.lexpress.fr/actualite/societe/environnement/visqueux-baveux-repoussants-ces-animaux-moches-qu-on-peine-a-proteger_2075462.html](https://www.lexpress.fr/actualite/societe/environnement/visqueux-baveux-repoussants-ces-animaux-moches-qu-on-peine-a-proteger_2075462.html).
- FEHR, E., & FALK, A. (2002). Psychological Foundations of Incentives. *European Economic Review*, 46(4-5), pp. 687-724.
- FINLAY, S., & SCHROEDER, M. (2017, August 18). *Reasons for Action: Internal vs. External*. (E. N. ZALTA, Éd.) Consulté le December 07, 2018, sur The Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/archives/fall2017/entries/reasons-internal-external/>
- FIRTH, R. (1952). Ethical Absolutism and the Ideal Observer. *Philosophy and Phenomenological Research*, 12(3), pp. 317-345.
- FLANAGAN, O. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge : Harvard University Press.
- FLEURBAEY , M., & MONGIN, P. (1996). Choix social. Dans M. CANTO-SPERBER, *Dictionnaire d'éthique et de philosophie morale* (pp. 295-304). Paris: Presses Universitaires de France.
- FLEURBAEY, M. (1994). L'absence d'envie dans une problématique "post-welfariste". *Recherches économiques de Louvain*, 60(1), pp. 9-41.
- FLEURBAEY, M. (1996). *Théories économiques de la justice*. Paris: Economica.
- FOOT, P. (2001). *Natural Goodness*. Oxford : Clarendon Press.
- FORST, R. (2012). *The Right to Justification: Elements of a constructivist Theory of Justice* . (J. Flynn, Trad.) New-York: Columbia University Press.
- FRANKENA, W. K. (1939). The Naturalistic Fallacy. *Mind*, 48(192), 464-477.

- FRANKFURT, H. (1969). Alternate Possibilities and Moral Responsibility. *Journal of Philosophy*, 66(23), pp. 829-839.
- FRITZ, A. (2009). *Der Naturalistische Fehlschluss*. Freiburg (CH): Academic Press Freiburg .
- FYOT, J.-L. (1952). *Dimensions de l'homme et science économique*. Paris: Presses Universitaires de France .
- GARDIES, J.-L. (1972). *Essai sur les fondements a priori de la rationalité morale et juridique*. Paris: L. G. D. J. .
- GARDIES, J.-L. (1987). *L'Erreur de Hume*. Paris : Presses Universitaires de France.
- GARDIES, J.-L. (2006). Le raisonnement moral et juridique est-il vérifonctionnel ? *Philosophia Scientiae* , 10(1), pp. 9-21.
- GARDIES, J.-L., & KALINOWSKI, G. (1974). Un logicien déontique avant la lettre: Gottfried Wilhelm Leibniz. *Archives for Philosophy of Law and Social Philosophy*, 60(1), pp. 79-112.
- GAUDÉ, L. (2004). *Le Soleil des Scorta* . Arles : Actes Sud.
- GAUTHIER, D. (1986). *Morals by Agreement*. Oxford: Oxford University Press.
- GEACH, P. (1956). Good and Evil. *Analysis*, 17, pp. 32-42.
- GEACH, P. (1965). Assertion. *Philosophical Review*, 74, 449–465.
- GIBBARD, A. (1990). *Sagesse des choix, justesse des sentiments: une théorie du jugement normatif*. (S. Laugier, Trad.) Paris: Presses Universitaires de France.
- GNEEZY, U., & RUSTICHINI, A. (2000). A Fine Is a Price. *The Journal of Legal Studies*, 29(1), pp. 1-17.
- GNEEZY, U., & RUSTICHINI, A. (2000). Pay Enough or Don't Pay at All. *The Quaterly Journal of Economics*, 115(3), pp. 791-810.
- GREENE, J. D. (2008). The Secret Joke of Kant's Soul. Dans W. SINNOT-ARMSTRONG, *Moral Psychology* (Vol. 3, pp. 35-80). Cambridge MA: MIT Press.
- HABERMAS, J. (1999). *Morale et communication*. (C. Bouchindhomme, Trad.) Paris: Flammarion.

- HAMEL, C. (2015). Vertu civique, intérêts personnels et bien commun. Repenser la politique de la vertu. *Les ateliers de l'éthique*, 10(1), pp. 100-128.
- HARE, R. M. (1952). *The Language of Morals*. Clarendon: Oxford University Press.
- HARMAN, G. (2000). *Explaining Value and Other Essays in Moral Philosophy*. Oxford: Oxford University Press.
- HAUSMAN, D. (1992). *The Inexact and Separate Science of Economics*. Cambridge : Cambridge University Press.
- HAUSMAN, D. (2011). *Preference, Value, Choice and Welfare*. New-York: Cambridge University Press.
- HAUSMAN, D., & MCPHERSON, M. (2007). The Philosophical Foundations of Mainstream Normative Economics. Dans D. HAUSMAN, *The Philosophy of Economics: an Anthology* (éd. Third, pp. 226-251). Cambridge: Cambridge University Press.
- HAUSMANN, D. a. (2006). *Economic Analysis, Moral Philosophy and Public Policy* (éd. Second Edition). New-York: Cambridge University Press.
- HAYEK, F. (2006 (1960)). *The Constitution of Liberty*. London: Routledge.
- HEATH, J. (2004). A Market Failures Approach to Business Ethics. Dans A. BRINK, & J. D. RENDTORFF, *Studies in Economic Ethics and Philosophy* (éd. Sp, Vol. 9, pp. 69-89). Dodrecht: Springer-Verlag.
- HEATH, J. (2012). Institutions et bénéfices coopératifs. Dans D. ANCTIL , D. ROBICHAUD, & P. TURMEL, *Penser les institutions: une introductions aux défis contemporains de la philosophie politique* (pp. 15-45). Québec: Presses de l'Université Laval.
- HERMAN, B. (1981). On the Value of Acting from the Motive of Duty. *The Philosophical Review*, 90(3), pp. 358-382.
- HINDE, R. (2011). *Changing How We Live: Society from the Bottom-Up*. Nottingham: Spokesman.
- HINDE, R. (2013). Ce dont nous avons besoin. *La Revue Durable*(48), pp. 35-40.
- HINTIKKA, J. (1969). Deontic Logic and its Philosophical Morals. Dans J. HINTIKKA, *Models for Modalities: Selected Essays* (pp. 184-214). Dodrecht: Reidel.

- HIRSCHMAN, A. O. (2013 (1977)). *The Passions and the Interests: Political Arguments for Capitalism before Its Triumph* . Oxford: Princeton University Press.
- HOBBS, T. (1991 (1651)). *Leviathan*. Cambridge, UK: Cambridge University Press.
- HOEKSTRA, K. (1997). Hobbes and the Foole. *Political Theory*, 25(5), pp. 620-654.
- HOLLIS, M. (1977). *Models of Man: Philosophical Thought on Social Action*. Cambridge: Cambridge University Press.
- HOLLIS, M., & NELL, E. J. (1975). *Rational Economic Man: A Philosophical Critic of Neo-Classical Economics*. Cambridge: Cambridge University Press.
- HOMANN, K. (2014). *Sollen und Können: Grenzen und Bedingungen der Individualmoral*. Vienna : Ibero Verlag - European University Press .
- HOMANN, K., KOSLOWSKI, P., & LUETGE, C. (2007). *Globalization and Business Ethics*. Burlington (US): Ashgate.
- HUEMER, M. (2005). *Ethical Intuitionism*. New-York: Palmgrave Macmillan.
- HUME, D. (1973 (1740)). *Traité de la nature humaine*. (A. Leroy, Trad.) Paris: Aubier.
- HURSTHOUSE, R. (1999). *On Virtue Ethics*. Oxford: Oxford University Press.
- INNALA, H. P. (2000). On the Non-Neutrality of Deontic Logic. *Logique & Analyse*, 43(171/172), pp. 393-410.
- JACKSON, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
- JAFFRO, L. (2016). Irrationalité pratique et contrôle de soi par anticipation. *Philonsorbonne*, 10, pp. 131-152.
- JAQUET, F. (2018). Evolution and Utilitarianism. *Ethical Theory and Moral Practice*, 21, pp. 1151–1161.
- JAQUET, F., & COVA, F. (2019). Internalisme et externalisme: le problème de la motivation morale. Dans O. DESMONS, S. LEMAIRE, & P. TURMEL, *Manuel de métaéthique* (pp. 335-366). Paris: Hermann.
- JAQUET, F., & NAAR, H. (2019). *Qui peut sauver la morale?: Essai de métaéthique*. Paris: Les Editions d'Ithaque.

- JENSEN, A., & RUSHTON, P. (2005). Wanter: More Race Realism, Less Moralistic Fallacy. *Psychology, Public Policy and Law*, 11(2), pp. 328-336.
- JOYCE, R. (2001). *The Myth of Morality*. Cambridge UK; New-York: Cambridge University Press.
- KAHNEMAN, D. (2011). *Thinking Fast and Slow*. New-York: Farras, Straus and Giroux.
- KÄMPFEN, V. (2014, Octobre 31). *Réformer la fiscalité et maintenir la prospérité*. Consulté le Avril 02, 2019, sur Fédération des Entreprises Romandes : <https://www.fer-ge.ch/web/fer-ge/-/reformer-la-fiscalite-et-maintenir-la-prosperte?inheritRedirect=true#.VJBeq3v7Mpo>
- KANT, E. (1841 (1793)). De l'origine du mal dans la nature humaine. Dans E. KANT, *La religion dans les limites de la raison*. Paris: Librairie de Ladrangue.
- KANT, E. (1986 (1793)). Sur le lieu commun: Il se peut que ce soit juste en théorie, mais, en pratique, cela ne vaut point. Dans E. KANT, *Oeuvres philosophiques, tome III*. Paris: Pléiades.
- KANT, E. (2007 (1785)). *Fondements de la métaphysique des Mœurs*. Paris: Hatier.
- KANT, E. (2016 (1785)). *Grundlegung zur Metaphysik der Sitten*. Berlin: Suhrkamp.
- KELSEN, H. (1991). *General Theory of Norms*. (M. Hartney, Trad.) Oxford: Clarendon Press.
- KENDE, P. (1959). L'intérêt personnel dans le système d'économie socialiste. *Revue économique*, 10(3).
- KORSGAARD, C. (1986). Skepticism about Practical Reason. *The Journal of Philosophy*, 83(1), pp. 5-25.
- KORSGAARD, C. (1996). *The Sources of Normativity*. Cambridge : Cambridge University Press.
- KRAUT, R. (1998). Egoism and Altruism. Dans E. CRAIG, *Routledge Encyclopedia of Philosophy*. London; New-York: Routledge.
- LARMORE, C. (2008). *The Autonomy of Morality*. Cambridge: Cambridge University Press.
- LARRÈRE, C. (1996). Intérêt. Dans M. CANTO-SPERBER, *Dictionnaire d'éthique et de philosophie morale* (Vol. 1, pp. 928-936). Paris: PUF.

- LAZARI-RADEK de , K., & SINGER, P. (2012). The Objectivity of Ethics and the Unity of Practical Reason. *Ethics*, 123, pp. 9-31.
- LEMAIRE, S. (2004). Internalisme. Dans M. CANTO-SPERBER, *Dictionnaire d'éthique et de philosophie morale* (pp. 936-940). Paris: Presses Universitaires de France.
- LEMAIRE, S. (2007). Une défense de l'hédonisme psychologique. Dans R. LEFEBRE, *Le Plaisir, Réflexions antiques, approches modernes*. (pp. 199-214.). Rouen; Havre : Publications des Universités de Rouen et du Havre .
- LEMAIRE, S. (2008). *Les désirs et les raisons: de la délibération à l'action* . Paris: Vrin.
- LEMMON, E. J. (1962). Moral Dilemmas. *Philosophical Review*, 71(2), pp. 139-158.
- MACKIE, J. L. (1977). *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.
- MANDEVILLE, B. (1991 (1729)). Troisième dialogue: l'amour propre et la politesse. Dans *La fable des abeilles, deuxième partie*. Paris: Vrin.
- MANDEVILLE, B. (1998 (1714)). *La fable des abeilles, première partie*. Paris: Vrin.
- MANIQUET, F. (2018). L'articulation entre les démarches positive et normative dans quelques branches de l'économie. *Revue philosophique de Louvain*, 116(2), pp. 187-213.
- MANTEL, S. (2017, Décembre). Worldly Reasons: An Inquiry Ontological Inquiry Into Motivating Considerations and Normative Reasons. *Pacific Philosophical Quarterly*, 98(S1), 5-28.
- MARKOVITS, J. (2014). *Moral Reason*. Oxford: Oxford University Press.
- MARTIN, A. (2019). On Respecting Animals, or Can Animals be Wronged Without Being Harmed? *Res Publica*, 25(1), pp. 83-99.
- MASLOW, A. H. (2004). *L'accomplissement de soi: de la motivation à la plénitude*. (É. Borgeaud, Trad.) Paris : Eyrolles .
- MCCLOSKEY, D. (2006). *The Bourgeois Virtues: Ethics for an Age of Commerce* . London: University of Chicago Press .
- MCDOWELL, J. (1979). Virtue and Reason. *Monist*, 62, pp. 331-350.

- MCMAHON, C. (1981). Morality and the Invisible Hand. *Philosophy and Public Affairs*, 10(3), pp. 247-277.
- MCNAUGHTON. (1988). *Moral Vision*. Oxford: Basil Blackwell.
- MELE, A. (2009). Weakness of Will and Akrasia. *Philosophical Studies*, 150(3), pp. 391-404 .
- MERCIER, S. (2014). *L'éthique dans les entreprises* (éd. troisième). Paris: Editions La Découverte.
- MILLER, A. (2013). *Contemporary Metaethics: An Introduction* (éd. 2nd). Cambridge: Polity Press.
- MIRRELES, J. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies*, 175(38), pp. 175-208.
- MONGE, R., STRUDLER, A., & KIM, T. (2015). Bounded Ethicality and The Principle that "Ought" Implies "Can". *Business Ethics Quarterly*, 25(3), pp. 341-361.
- MONGIN, P. (1999). Normes et jugements de valeur en économie normative. *Information sur les sciences sociales*, 34(8), pp. 521-533.
- MONGIN, P. (2018). Les origines de la distinction entre le positif et le normatif en économie. *Revue philosophique de Louvain*, 116(2), pp. 151-186.
- MOORE, E. C. (1957). The Moralistic Fallacy. *Journal of Philosophy*, 54(2), pp. 29-42.
- MOORE, G. E. (1976). Review of Franz Brentano's The Origin of the Knowledge of Right and Wrong. Dans L. MCALISTER, *The Philosophy of Franz Brentano*. London: Duckworth.
- MOORE, G. E. (1993 (1903)). *Principia Ethica*. Cambridge: Cambridge University Press.
- MURPHY, M. (2002). *An Essay on Divine Authority* . Ithaca: Cornell University Press.
- NAGEL, T. (1970). *The Possibility of Altruism*. Princeton: Princeton University Press.
- NAGEL, T. (1986). *The View from Nowhere*. New-York: Oxford University Press.
- NAGEL, T. (1997). *The Last Word*. New-York: Oxford University Press.
- NAGEL, T. (2008). The Value of Inviolability. Dans P. BLOOMFIELD, *Morality and Self-Interest* (éd. Editions Kindle, pp. 102-114). Oxford: Oxford University Press.

- NELSON, M. (1995). Is It Always Fallacious to Derive Values from Facts? *Argumentation*, 9, pp. 553-562.
- NELSON, M. (2003). Who Needs Valid Moral Arguments? *Argumentation*, 17, pp. 35-42.
- NELSON, M. (2007). More Bad News for the Logical Autonomy of Ethics. *Canadian Journal of Philosophy*, 37(2), pp. 203-216.
- NELSON, M. (2015). What the Utilitarian Cannot Think. *Ethical Theory and Moral Practice*, 18(4).
- NÉRON, P.-Y. (2010). La moralité implicite du marché. *Les ateliers de l'éthique*, 5(1), pp. 4-22.
- NÉRON, P.-Y. (2016). Coopération, compétition et moralité: La main invisible du marché et la division du travail moral. *Raison publique : revue semestrielle de philosophie sociale et politique*(16), pp. 187-207.
- NIDA-RÜMELIN, J. (2011). *Die Optimierungsfalle: Philosophie einer humanen Ökonomie*. München : Btb-Verlag .
- NOOTEBOOM, B. (2014). *How Markets Work and Fail, and What to Make of Them*. Cheltenham: Edward Elgar.
- NUROCK, V. (2011). *Sommes-nous naturellement moraux?* . Paris: PUF .
- O'CONNOR, M. P., & THOMPSON, M. (2009). Information Technology and Communication System Development: Information and Knowledge. Dans G. BARBIROLI, *Principles of Sustainable Development: Encyclopedia of Life Support Systems* (pp. 56-78). Oxford: EOLSS/UNESCO.
- OGIEN, R. (2001). Le Rasoir de Kant. *Philosophiques*, 28(1), pp. 9-25.
- O'NEILL, J. (1990). Property in Science and the Market. *The Monist: Property Rights*, 73(4), pp. 601-620.
- O'NEILL, O. (2015). *Constructing Authorities: Reason, Politics and Interpretation in Kant's Philosophy* . Cambridge (UK) : Cambridge University Press.
- PAOLY, P. (2012). *Comprendre l'économie politique: manuel à l'usage des curieux*. Vevey: Editions de l'Aire.
- PARFIT, D. (2011). *On What Matters* (Vol. 2). Oxford: Oxford University Press.

- PHAM, T. K. (2008). L'hypothèse d'utilité relative dans l'analyse: enjeux et conséquences. *Revue d'économie politique*, 118(04), pp. 541-572.
- PIES, I. (2014). Der Wirtschaftsethische Imperativ lautet : Denkfehler vermeiden ! Dans H. F. WOLF, *Ökonomische Moral oder moralische Ökonomie: positionen zu den Grundlagen der Wirtschaftsethiki* (pp. 16-51). München, Deutschland : Verlag Karl Alber.
- PIGDEN, C. R. (1990). Ought-Implies-Can: Erasmus, Luther and R.M. Hare. *Sophia*, 29(1), pp. 2-30.
- PINK, T. (2004). Moral Obligation. Dans A. O. (ed.), *Modern Moral Philosophy* (pp. 159-186). Cambridge: Cambridge University Press.
- PINKER, S. (1997). *How the Mind Works*. New-York: W. W. Norton & Company.
- PINKER, S. (2002). *The Blank Slate: the Modern Denial of Human Nature* . London: Penguin Books.
- POTTER, N. T. (1997). The Synthethic a priori Proposition of Kant's Ethical Philosophy. *Jahrbuch für Recht Und Ethik* , 5.
- PRICHARD, H. A. (1912). Does Moral Philosophy Rest on a Mistake? *Mind*, 21, pp. 21-37.
- PRIOR, A. N. (1960). The Autonomy of Ethics. *Australasian Journal of Philosophy*, 38(3), pp. 198-206 .
- QUILLIET, B. (2007). *L'acharnement théologique: histoire de la grâce en Occident, IIIe-XXIe siècle*. Paris : Fayard .
- QUINN, P. (1978). *Divine Commands and Moral Requirements* . Oxford: Oxford University Press.
- RACHELS, J. (2000). Naturalism. Dans LaFolette, *The Blackwell Guide to Ethical Theory* (pp. 74–91). Oxford: Blackwell Publishing.
- RAILTON, P. (1986). Moral Realism. *The Philosophical Review*, 95(2), pp. 163-207.
- RAILTON, P. (2016). Le réalisme moral. *Les ateliers de l'éthique*, 11(2-3), pp. 171-212.
- RAND, A. (1967). *Capitalism: The Unknown Ideal*. New-York: Penguin.
- RASMUSSEN, D. C. (2014). *The Pragmatic Enlightenments: Recovering the Liberalism of Hume, Smith, Montesquieu and Voltaire*. Cambridge: Canbridge University Press.

- RAWLS, J. (2006). *Paix et démocratie: Le droit des peuples et la raison publique* . Paris : La découverte .
- RAWLS, J. (2009). *Théorie de la justice*. (C. Audard, Trad.) Paris: Points.
- RAYNAULD, J., & STRINGER, Y. (2003). *Problèmes et politiques économiques*. Sainte-Foy (Québec): Presses de l'Université du Québec.
- REINACH, A. (2004 (1913)). *Les Fondements a priori du droit civil*. (R. d. Calan, Trad.) Paris : Vrin.
- ROCHELEAU-HOULE, D. (2013). L'expressivisme moral: une conclusion de la critique darwinienne de Sharon Street? *Les cahiers d'Ithaque* , pp. 25-41.
- RODRIGUES, J. (2013). Between Rules and Incentives: Uncovering Hayek's Moral Economy. *The American Journal of Economics and Sociology*, 72(3), pp. 565-592.
- ROOJEN van, M. (2018). Moral Cognitivism vs. Non-Cognitivism. (E. D. Zalta, Éd.) *Stanford Encyclopedia of Philosophy*.
- ROSATI, C. S. (2016). Moral Motivation. (E. N. Zalta, Éd.) *Stanford Encyclopedia of Philosophy*.
- ROSS, W. D. (1930). *The Right and the Good*. Oxford: Oxford University Press.
- RUSE, M. (1993). Une défense de l'éthique évolutionniste. Dans J.-P. CHANGEUX, *Fondements naturels de l'éthique* (pp. 35-64). Paris: Odile Jacob.
- SAILER, S. (2002, Octobre 30). Q&A: *Steven Pinker of 'Blank Slate'*. Consulté le Septembre 12, 2019 , sur United Press International : https://www.upi.com/Odd_News/2002/10/30/QA-Steven-Pinker-of-Blank-Slate/26021035991232/
- SAYRE-MCCORD, G. (1986). Deontic Logic and The Priority of Moral Theory. *Noûs*, 20(2), pp. 179-197.
- SCANLON, T. (1998). *What We Owe to Each Other*. Cambridge (MA) and London : Harvard University Press.
- SCANLON, T. (2014). *Being Realistic About Reasons*. Oxford : Oxford University Press.
- SCHEFFLER, S. (1992). *Human morality*. New-York: Oxford University Press.

- SCHROEDER, M. (2005). Realism and Reduction: The Quest for Robustness. *Philosophers' Imprint*, 5(1), pp. 1-18.
- SCHROETER, F. (2005). Normative Concepts and Motivation. *Philosophers' Imprint*, 5(3), pp. 1-23.
- SCHULTHESS, D. (1986). Une conception de l'objectivité du mal. Dans R. KAEHR, & J. HAINARD, *Le mal et la douleur* (pp. 171-180). Neuchâtel: Musée d'ethnographie .
- SCHULTHESS, D. (1998). Reflexions sur la motivation économique. Dans R. DAMIEN , & A. TOSEL, *L'action collective: coordination, conseil, planification* (Vol. 12, pp. 247-257). Besançon: Annales littéraires de L'université de Franche-Comté.
- SEARLE, J. (1964). How to Derive "Ought" from "Is". *The Philosophical Review*, 73, pp. 43-58.
- SEN, A. (1977). Rational Fools: A Critique of the Behavioral Foundations of Economic Theory. *Philosophy and Public Affairs*, 6(4), pp. 317-344.
- SÈVE, R. (1993). L'Éthique comme besoin. Dans J.-P. CHANGEUX, *Fondements naturels de l'éthique* (pp. 105-115). Paris: Odile Jacob.
- SHAFER-LANDAU, R. (2005). *Moral Realism: A Defense*. Oxford : Oxford University Press.
- SHIRINE, S. (2008). La notion d'intérêt général chez Adam Smith: de la richesse des nations à la puissance des nations. *Géoéconomie*, 45(2), pp. 55-71.
- SIEBERT, H. (2011). Incentives, Efficiency and Equity in a Market Economy. Dans *Social and Ethical Aspects of Economics* (éd. 2nd). Vatican City : Pontifical Council for Justice and Peace .
- SIMMEL, G. (1987). *La philosophie de l'argent*. (t. d. Ivernel, Trad.) Paris: PUF.
- SIMON, H. (1957). *Models of Man: Social and Rational. Mathematical Essays on Rational Behavior in a Social Setting*. New York : Willey .
- SINNOTT-ARMSTRONG, W. (1984). "Ought" Conversationally Implies "Can". *The Philosophical Review*, 93(2), pp. 249-261.
- SLEMROD, J. B. (2000). *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*. Cambridge: Harvard University Press.

- SMALL, E. (2011). The New Noah's Ark: beautiful and useful species only. *Biodiversity*, 12(4), pp. 232-247.
- SMITH, M. (1994). *The Moral Problem*. Oxford; Cambridge: Wiley-Blackwell.
- SMITH, M. (2004). Descriptivisme. Dans M. CANTO-SPERBER, *Dictionnaire d'éthique et de philosophie morale* (pp. 492-499). Paris: Presses Universitaires de France.
- SMITH, M. (2009). The Motivation Argument for Non-Cognitivism. Dans C. PIGDEN, *Hume on Motivation and Virtue* (pp. 105-120). London: Palgrave Macmillan.
- SPENCER, H. (1879). *The Data of Ethics*. London : Williams and Norgate.
- SPINOZA, B. (1988 (1677)). *L'Éthique*. (B. Pautrat, Trad.) Paris: Editions du Seuil.
- STEVENSON, C. L. (1944). *Ethics and Language*. New Haven: Yale University Press.
- STOCKER, M. (1971). "Ought" and "Can". *Australasian Journal of Philosophy*, 49(3), pp. 303-316.
- STOCKER, M. (1979). Desiring the Bad: An Essay in Moral Psychology. *The Journal of Philosophy*, 76(12), pp. 738-753.
- STRANDBERG, C., & BJÖRKLUND, F. (2013). Is Moral Internalism Supported by Folks Intuitions? *Philosophical Psychology*, 26(3), pp. 319-335.
- STREET, S. (2006). A Darwinian Dilemma for Realist Theories. *Philosophical Studies*, 127, pp. 109-166.
- STROUD, S., & SVIRSKY, L. (2019). Weakness of Will. Dans Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.
- STURGEON, N. (1988). Moral Explanations. Dans G. SAYRE-MCCORD, *Essays in Moral Realism* (pp. 229-255). New-York: Cornell University Press.
- SUGDEN, R. (1981). *The political economy of public choice: An Introduction to Welfare Economics*. Oxford: M. Robertson.
- TAPPOLET, C. (2011). La normativité des concepts évaluatifs. *Philosophiques*, 38(1), pp. 157-176.
- THOMPSON, M. (1995). The Representation of Life. Dans R. HURSTHOUSE, G. LAWRENCE, & W. QUINN, *Virtues and Reasons* (pp. 247-296). Oxford: Clarendon Press.

- TIMMONS, M. (1985). Kant and the Possibility of Moral Motivation. *The Southern Journal of Philosophy*, 23(3), pp. 377-398.
- TIROLE, J. (2016). *Economie du bien commun*. Paris, France: Presses Universitaires de France.
- TIROLE, J., & BENABOU, R. (2006). Incentives and Prosocial Behavior. *The American Economic Review*, 96(5), pp. 1652-1678.
- TOMLIN, P. (2008). Envy, Facts and Justice: A Critique of the Treatment of Envy in Justice as Fairness. *Res Publica*, 14(2), pp. 101-116.
- UNESCO. (1986). Manifeste de Séville sur la violence. *Conférence générale de l'UNESCO*. Paris.
- URI, G., MEIER, S., & REY-BIEL, P. (2011). When (and Why) Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4), pp. 191-210 .
- VELLEMAN, D. J. (1998). Is Motivation Internal to Value. Dans C. FEHIGE, & U. WESSELS, *Preferences*. Berlin: Walter de Gruyter.
- VERBURG, R. (2015). Bernard Mandeville's Vision of the Social Utility of Pride and Greed. *European Journal of the History of Economic Thought*, 22(4), pp. 662-691.
- VERBURG, R. (2018). *Greed, Self-Interest and the Shaping of Economics* . New-York: Routledge .
- von WRIGHT, G. H. (1963). *Norm and Action: A logical Enquiry* . London: Routledge and Kegan Paul.
- WALLACE, J. R. (1998). Moral Motivation. Dans E. CRAIG, *Routledge Encyclopedia of Philosophy* (pp. 522-528). London and New-York: Routledge.
- WEDGWOOD, R. (2006). The Meaning of 'Ought'. Dans R. SHAFER-LANDAU, *Oxford Studies in Metaethics* (Vol. 1, pp. 127-160). Oxford: Oxford University Press.
- WEDGWOOD, R. (2007). *The Nature of Normativity* . Oxford : Oxford University Press.
- WENNSTRÖM, J. (2017). *Market, State, and Morality: Two Studies of How Left and Right Undermined Moral Motivation in the Swedish School System*. Linköping (Sweden): Department of Management and Engineering (Thesis).

- WHITE, R. (2005). "Ought" Implies "Can": Kant and Luther, A Contrast. Dans G. MCDONALD ROSS , & T. MCWALTER, *Kant and His Influence* (pp. 1-72). New-York: Continuum International Publishing Group.
- WILLIAMS, B. (1972). *Morality* (éd. Editions Kindle). New-York: Cambridge University Press.
- WILLIAMS, B. (1973). *Problems of the Self* . Cambridge: Cambridge University Press .
- WILLIAMS, B. (1981). Internal and External Reasons. Dans B. WILLIAMS, *Moral Luck: Philosophical Papers 1973-1980*. Cambridge : Cambridge University Press.
- WILLIAMS, B. (1981). *Moral Luck* . Cambridge: Cambridge University Press.
- WILLIAMS, B. (1994). Égoïsme et altruisme. Dans B. WILLIAMS, *La fortune morale: moralité et autres essais* (J. Lelaidier, Trad., pp. 155-176). Paris : Presses Universitaires de France.
- WILLIAMS, B. (1994). Moralité . Dans B. WILLIAMS, *La fortune morale: moralité; et autres essais* . Paris : Presses Universitaires de France.
- WILLIAMS, B. (1995). *Making Sense of Morality and other philosophical papers (1982-1993)*. Cambridge : Cambridge University Press.
- WILLIAMS, B. (2006). *Ethics and the Limits of Philosophy* (éd. Première édition 1985). Abingdon: Routledge.
- WILLIAMS, B. (2014). Vertus et vices. Dans M. CANTO-SPERBER, *Dictionnaire d'éthique et de philosophie morale* (Vol. 2, pp. 2019-2024). Paris : Presses Universitaires de France .
- WILSON, F. (1982). Mill's Proof That Happiness Is the Criterion of Morality. *Journal of Business Ethics*, 1(1), pp. 59-72.
- WOLF, J.-C. (2002). Ist es rational, Egoist zu sein? Dans J.-C. WOLF, *Ethik und Politik ohne Gewissheiten* (pp. 85-105). Freiburg (Schweiz): Universitätsverlag Freiburg Schweiz.
- WOLF, J.-C. (2004). Ethischer Egoismus. *Erwägen, Wissen, Ethik*, 15(4), pp. 513-590.
- WOLFELSPERGER, A. (2010). Le Désintéressement malveillant: Ou le chapitre manquant du Traité de l'homme économique de Jon Esalter. *Revue française de sciences politiques*, 60(4), pp. 769-785.
- WRIGHT, C. (1985). Review of Blackburn 1984. *Mind*.

